Investigating social influences on and changes in health-related-behaviour clustering during mid-adulthood, using data from two British birth cohort studies

Claire Mawditt

A thesis submitted for the degree of Doctor of Philosophy

Department of Epidemiology and Public Health

University College London

2017

Declaration

I, Claire Mawditt, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

[redacted]

Claire Mawditt

Abstract

Introduction: This thesis examines how four health-related behaviours (HRBs) – smoking, alcohol, diet, physical activity – cluster within individuals during mid-adulthood, the relationship between socio-economic position (SEP) in pre-adolescence and mid-adulthood and HRB cluster membership in mid-adulthood and the extent to which mid-adulthood SEP influences change in HRB cluster membership during mid-life.

Methods: The research used datasets from two British birth cohort studies: The National Child Development Study (born in 1958) and the British Birth Cohort study (born in 1970). Latent variable modelling was employed to: (i) identify cross-sectional membership of clusters, who shared patterns of HRBs: (ii) examine changes in cluster membership during mid-life: and iii) validate the relationship between SEP in pre-adolescence and mid-adulthood and HRB cluster membership.

Results: Three distinct clustered patterns of HRBs were identified and subsequently labelled: 'Risky' (1–9%), 'Moderate Smokers' (20–30%) and 'Mainstream' (68–77%). The Mainstream cluster was characterised by more health-promoting behavioural patterns, i.e. not smoking, frequent fruit and vegetable consumption, less frequent consumption of chips and fried food, being more physically active, although frequent consumption of sweet foods was common. HRB cluster patterns were largely consistent across cohort and gender groups, with some differences in prevalence. More disadvantaged SEP in pre-adolescence predicted more disadvantaged SEP in mid-adulthood which increased the probability of membership of the 'Risky' and 'Moderate Smokers' clusters compared to the 'Mainstream' cluster. HRB cluster membership was found to be relatively stable during mid-life, although there was evidence of transitions to more health-promoting clusters. These transitions were not influenced by mid-adulthood SEP.

Discussion: Consistent findings for the two cohorts imply HRB clustering and their social patterning persists across time and provides a person-centred understanding that can inform interventions to improve HRBs. The contemporaneous influence of mid-adulthood SEP on cluster membership provides optimism, suggesting that mid-adulthood lifestyles may be modifiable.

Acknowledgments

I would firstly like to thank my primary and secondary supervisors, Noriko Cable and Amanda Sacker, who have both dedicated a considerable amount of time and expertise to my PhD. Secondly, I would like to thank the remaining members of my supervisory panel, Annie Britton and Yvonne Kelly, who have provided valuable insights throughout. I feel privileged to have had such a committed and knowledgeable supervisory panel, this project would not have been possible without their invaluable support, advice and feedback.

I would also like to thank members of the International Centre for Lifecourse Studies (ICLS) for their professional and friendly support over these years. Being a PhD student in the ICLS has provided me with many learning opportunities and has greatly benefitted my professional development. I have been privileged to share my PhD experience with fellow ICLS students whose support and encouragement is greatly appreciated. I thank Professor Ingrid Schoon who generously gave her time to examine by MPhil to PhD transfer and helped focus and improve this project.

Many thanks to the Economic and Social Research Council for funding my PhD studentship as well as an Overseas Institutional Visit to Penn State University and a fellowship in the House of Lords library. I am grateful to staff at Penn State University and the House of Lords library who made these placements thoroughly enjoyable and worthwhile learning experiences.

Finally, I would like to thank my family and friends for their unwavering support and encouragement. I especially thank my partner, my parents and my sister, for giving me the strength and determination to carry on. I dedicate this thesis to you.

Table of contents

List of tables

List of figures

List of abbreviations

HRB	Health-Related Behaviour
SEP	Socio-Economic Position
NCDS	National Child Development Study (1958 birth cohort)
BCS70	British Birth Cohort Study (1970 birth cohort)
SD	Standard Deviation
SE	Standard Error
CI	Confidence Interval
FIML	Full Information Maximum Likelihood
MLR	Robust Maximum Likelihood
WLSMV	Weighted Least Squares with Robust Standard Errors
CFI	Comparative Fit Index
RMSEA	Root Mean Square Error of Approximation
SEM	Structural Equation Model
LPA	Latent Profile Analysis
LTA	Latent Transition Analysis
CFA	Confirmatory Factor Analysis
EFA	Exploratory factor analysis
LMR	Lo-Mendell Rubin Likelihood Ratio Test
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
aBIC	Adjusted Bayesian Information Criterion

1 Chapter 1: Introduction

This doctoral research focuses on four health-related behaviours (HRBs): smoking, alcohol, diet and physical activity. These four HRBs have been selected due to their known effects on mortality and morbidity in high-income countries, including the United Kingdom (Lim et al., 2013). These four HRBs are a major focus of domestic public health policy because they are major risk factors for a number of non-communicable diseases including cardiovascular disease, cancer and lung disease which are among the leading causes of premature mortality in the United Kingdom (DOH, 2014).

Smoking, heavy alcohol consumption, physical inactivity, high consumption of sugar and fat, and low consumption of fruit and vegetables negatively impact on health (Schuit et al., 2002). In contrast, not smoking, moderate alcohol consumption, being physically active, consuming a higher proportion of fruit and vegetables and a low proportion of fat and sugar, positively impact on health (Byrne et al., 2016; Harrington et al., 2010). From a public health perspective, addressing negative HRBs and promoting positive HRBs is important because they are modifiable (Filippidis et al., 2016).

Policymakers attempting to address negative HRBs and promote positive HRBs are currently working within siloes that consider them as individual and unrelated entities (Buck and Frosini, 2012; Filippidis et al., 2016; Noble et al., 2015). This thesis takes the view that considering these HRBs as unrelated may not be appropriate, given a growing body of empirical evidence (Meader et al., 2016; Noble et al., 2015) which suggests that individuals have multiple HRBs which do not co-occur through chance and therefore cluster. Consequently, this thesis takes a person-centred approach by considering the underlying relationship between HRBs, in order to better understand the formation of health lifestyles in mid-adulthood.

This thesis also identifies the influence of social circumstances on HRB clustering as well as how HRB clustering may change during mid-adulthood. This is based upon existing empirical evidence which suggests that the clustering of HRBs is socially patterned (Meader et al., 2016; Noble et al., 2015), multiple HRBs have the potential to change during mid-life (Mulder et al., 1998) and social circumstances can influence multiple HRB change (Buck and Frosini, 2012).

The thesis builds upon this existing evidence by considering differences in HRB clustering across two British cohorts and examines the influence of social circumstances at different points in the lifecourse on HRB clustering across the two cohorts. Change in HRB clustering during mid-adulthood is also investigated in the earlier-born cohort.

This chapter provides relevant background for the thesis and will be drawn upon throughout. Concepts of multiple HRBs and HRB clustering are introduced, the salience of HRB clustering in mid-adulthood and the social determinants and lifecourse framework are explained, as is the significance of cross-cohort comparisons. The structure of the thesis is also provided.

1.1 Multiple health-related behaviours

Individuals practise multiple HRBs in their everyday lives (Cockerham, 2005). Empirical evidence has identified that, on average, individuals in the UK engage in 2 to 3 negative HRBs out of the 4 listed above, with a very small number of people engaging in 0 or 1 negative HRBs (Buck and Frosini, 2012; Lawder et al., 2010; Poortinga, 2007; Shankar et al., 2010). This is of public health concern, given evidence that suggests that the presence of multiple negative HRBs increases the risk of mortality (Khaw et al., 2008; Kvaavik et al., 2010; Martin-Diener et al., 2014), morbidity (Alageel et al., 2016; Artaud et al., 2016; Chow et al., 2010; Sabia et al., 2012) as well as pre-outcome indicators, such as cardiovascular disease biomarkers (Loprinzi et al., 2016).

1.2 Health-related-behaviour clustering

There are two main areas of investigation within multiple HRB research: co-occurrence and clustering (McAloney et al., 2013). Examination of co-occurrence involves analyses 'of the concurrent but independent, engagement in two or more HRBs' (McAloney et al., 2013: 366) whereas clustering investigates 'underlying associations between co-occurring health-related behaviours' (McAloney et al., 2013: 366). Therefore, the clustering of HRBs 'implies that they are not independent of each other' (Ebrahim et al., 2004: 4). Literature using these different approaches to investigate HRB has been appraised in this thesis. Studies which have used both approaches suggest that the presence of multiple HRBs is not through chance (Laaksonen et al., 2001; Poortinga, 2007; Schuit et al., 2002; Silva et al., 2013; Tobias et al., 2007). For example, the literature appraised in this thesis provides compelling evidence that individuals who smoke and drink alcohol heavily also have poorer (Meader et al., 2016; Noble et al., 2015).

This evidence suggests that these HRBs are inter-related and that the clustering approach is preferable in order to study patterns of health behaviour. This implies that current public health policies, which tackles these four HRBs as individual entities, may not be appropriate and that public health interventions need to take a more sophisticated approach and consider the inter-relatedness of these four HRBs. Moreover, a clustering approach to HRB allows a deeper understanding of the types of people who share these lifestyles which can inform the development of lifestyle interventions that target specific subgroups of the population (Graham et al., 2016).

At the same time, there is still a large amount of variability in the number and nature of HRB patterns being reported across studies investigating HRB clustering (Noble et al., 2015). The literature appraised in this thesis suggests that this lack of consensus is, in part, due to heterogeneity across the studies in terms of the populations under investigation, the measurement of HRBs and the analytical methods employed. Therefore, whilst insightful, this literature provides no *a priori* on which to base a theoretical approach.

Consequently, a data driven clustering approach using latent variable models to investigate multiple HRBs is preferred and is the focus of this doctoral research.

The clustered patterns of multiple HRBs are considered to reflect health lifestyles (Cockerham, 2005; Jones et al., 2011; Maller, 2015; Pampel et al., 2010). The clusters may consist of negative HRBs that are health damaging (e.g. smoking, heavy alcohol consumption, a diet high in sugar and fat and low in fruit and vegetables, and physical inactivity), positive HRBs that are health promoting (e.g. not smoking, light or moderate alcohol consumption, a diet low in sugar and fat and high in fruit and vegetables and low in sugar and fat, and regular physical activity) or a mixture of both.

These clusters of HRBs are considered to provide self-identity and meaning and reflect the consumption patterns of particular social groups (Blue et al., 2014; Cockerham, 2005). Therefore, groups of individuals who share clustered patterns of HRBs practise them within a particular social context located in a particular time and place (Burton-Jeangros et al., 2015). These subgroups dictate the persistence or demise of HRB clusters (Maller, 2015). Within this context, HRB clusters not only reflect social groups, but can also be considered dynamic and prone to change across historical time, generations and the lifecourse (Short and Mollborn, 2015).

Consequently, a comprehensive investigation of HRB cluster patterns in different subgroups of the population can provide a person-centred understanding of multiple HRBs which in turn can inform both individual-level and population-level interventions to improve HRBs that reflect the individual’s experience (Buck and Frosini, 2012; Watts et al., 2015).

1.3 The importance of health-related-behaviour clustering in mid-adulthood

The focus of this doctoral research is HRB clustering in mid-adulthood. There is some evidence to support the notion that individual HRBs are relatively stable during this period of life (Benzies et al., 2008; Mulder et al., 1998). This is in contrast to adolescence and young adulthood (i.e. teens and twenties (Backett and Davison, 1995)) which are considered to be normative periods for behavioural experimentation (Fothergill et al., 2009; Schooling and Kuh, 2002).

HRB stability during mid-life implies that HRBs could be sustained across mid-adulthood which is considered to be an important period in the lifecourse when individuals play a central role in supporting younger and older generations and make a valuable contribution to the workforce and society (Lachman et al., 2015). Prolonged patterns of negative HRBs may adversely impact the next generation, with evidence demonstrating parental behaviours influence those of their children (Brown and Ogden, 2004; Crawley and While, 1996; Edwardson and Gorely, 2010; Gilman et al., 2009; Nash et al., 2005; Pearson et al., 2009; Van Der Vorst et al., 2005). There may also be repercussions for the economy, given that informal caregiving to older generations is most commonly performed by mid-age adults (Lachman et al., 2015) and is increasingly relied upon in developed countries (O'Reilly et al., 2015). From a labour force perspective, poor health in the working-age population, of which a high proportion are middle-aged, costs the UK economy over £100 billion annually (Shreeve et al., 2015).

This also has implications on health outcomes, given evidence that HRBs may continue to track from mid-adulthood into later life (Hamer et al., 2012) and that a persistent lifestyle consisting of multiple negative HRBs during mid-life is associated with earlier mortality (Berstad et al., 2016).

At the same time, some change in HRBs has been observed during mid-adulthood, tending to be in a positive direction – for example, increased fruit and vegetable consumption (Artaud et al., 2016; Backett and Davison, 1995; Benzies et al., 2008; Mulder et al., 1998; Sijtsma et al., 2012), reductions in the number of cigarettes smoked or abstinence from smoking (Artaud et al., 2016; Backett and Davison, 1995; Mulder et al., 1998; Paffenbarger Jr et al., 1993) and reduced alcohol consumption (Backett and Davison, 1995; Benzies et al., 2008; Britton et al., 2015; Meng et al., 2014; Molander et al., 2010; Mulder et al., 1998). Some research has also found HRB change in mid-life to be negative – for example, reductions in physical activity (Allender et al., 2008b; Artaud et al., 2016; Corder et al., 2009; Mulder et al., 1998; Wannamethee et al., 1998).

This line of enquiry is important given that to date, and to my knowledge, there is no research examining the stability of HRB clustering in mid-life in Britain. Furthermore, research indicates that positive change in HRBs during mid-adulthood may prevent early death (Berstad et al., 2016) as well as improve physical functioning (Cooper et al., 2011) and reduce the risk of disability in later life (Artaud et al., 2016).

Investigating the stability of HRB clustering during mid-life may elucidate ‘natural fluctuations’ in HRBs (Mulder et al., 1998) which can inform population-level interventions to improve HRBs in this particular age group. Therefore, the extent to which HRB stability in mid-life is applicable to the clustering of multiple HRBs is both questionable and salient and is subsequently investigated in this work.

1.4 A cross-cohort comparison of two British birth cohort studies

This thesis uses longitudinal prospectively collected data from two British birth cohort studies. The National Child Development Study (NCDS) which follows the lives of 17,514 live births born within the same week in 1958 across England, Scotland and Wales (Power and Elliott, 2006) and the 1970 British Birth Cohort Study (BCS70), following the lives of 16,571 people born across England, Scotland and Wales in one week in 1970 (Elliott and Shepherd, 2006). Further information on these studies is provided in section 5.1.1.

The ability to take a data driven rather than theoretically driven approach to determine the clustering of HRBs is possible due to these two well-conducted cohort studies, the NCDS and the BCS70, that provide large samples containing information on all four HRBs. Moreover, the two cohort studies used in this thesis were selected because both contain rich longitudinal data on HRBs during mid-life and social circumstances at different points in the lifecourse. This information is prospectively, rather than retrospectively, collected, allowing for the temporal ordering of variables across the lifecourse, improving causal inference (Hoyle and Robinson, 2003) and minimising recall bias (Cohen et al., 2010).

A cohort comparison is made possible in these two cohorts due to their similarities in study design and both contain deliberately similar variables (Ekinsmyth et al., 1992). This ability to compare cohorts will be exploited in this thesis. Such a comparison maximises the research potential of these existing cohort studies (Khoury and Evans, 2015) by evidencing the successful replication of the research methodology. This approach allows for a valid comparison of the study results by reducing the likelihood of differences due to differential measurement. Moreover, the cross-cohort comparison controls for cohort effects which are inherent to any single longitudinal study, and indicates the combined effects of period and age (Schoon, 2006). Additionally, there is a period of overlap across the two cohorts. Participants in both cohorts are in mid-adulthood in the early 2000s, with BCS70 participants being in their 30s and NCDS participants being in their 40s. Consequently, similarities in findings at these ages may, to some extent, elucidate period effects (Schoon, 2006).

Findings which are consistent across the two cohort studies provide a stronger argument that the results persist across time and may to some extent be generalised to later cohorts (i.e. mid-age adults within Britain today). At the same time, findings that are different across the two cohorts can increase our understanding of how the unique experiences of these groups of participants born 12 years apart influence the clustering of HRBs and the relationship between HRB cluster membership and social circumstances. In consequence, elucidating cohort differences may serve to generate new hypotheses as to how the unique experiences of later cohorts may influence their health lifestyles.

1.5 Social determinants and lifecourse framework

This thesis takes the view that social circumstances at different points in the lifecourse may be particularly salient in the development of health lifestyles in mid-adulthood. SEP during pre-adolescence (age 8–11, (Maggs et al., 2008)) may be a critical period in the formation of HRB clustering in mid-adulthood. It is conceived that this effect, at least in part, could occur through the accumulation of resources dictating early mid-adulthood SEP which may in turn influence mid-adulthood HRB clustering.

This doctoral research builds upon well-established determinants of health models (Brunner and Marmot, 2006; Dahlgren and Whitehead, 1991) whereby structural factors are considered to determine HRB clustering rather than individual agency. This approach diverts attention from individual choice as the driver of lifestyle and instead highlights the role of disadvantaged social circumstances at key stages in life, captured by their socio-economic position (SEP) consisting of material, occupational and cultural dimensions (Bartley et al., 1999; Sacker et al., 2001), in relation to HRBs. Disadvantaged SEP may prevent participation in positive HRBs that are health promoting whilst at the same time encouraging negative HRBs that are health damaging (Katikireddi et al., 2013; Maller, 2015).

Examining the role of SEP during pre-adolescence in relation to mid-adult HRB clustering could show how social circumstances at an early stage in the lifecourse would link to a particular HRB cluster, deepening our understanding of health inequalities previously observed in the two cohort studies (i.e. the NCDS and the BCS70) used in this thesis (Ferri et al., 2003; Mensah and Hobcraft, 2008; Parsons et al., 2013), via behavioural or social pathways. This line of enquiry is important, given evidence that HRBs may explain between 54% (Laaksonen et al., 2008) and 70% (Beauchamp et al., 2010) of the relationship between measures of SEP and mortality. The implications of this framework would be that policies and interventions to prevent health-damaging HRB clustering in mid-age should take a lifecourse approach by considering social circumstances earlier in life.

1.6 Structure of the thesis

Chapter 2 outlines literature that guided the development of the conceptual model for this thesis and elaborates on the points raised in this section. The model to be tested is described in chapter 3. This is followed by the aims, objectives and research questions which are outlined in chapter 4. An overview of the data, variables and statistical analysis used in this thesis are presented in chapter 5. The three succeeding chapters 6, 7 and 8 relate to each of the research objectives outlined in chapter 4, each providing more detail on the methods used to address that specific objective, including the analytical sample and management of missing data as well as the results and an interim discussion. Chapter 9 summarises result chapters 6, 7 and 8, provides further interpretations and policy implications, outlines areas of future work, gives a brief final reflection and ends with concluding remarks.

2 Chapter 2: Literature review

2.1 Introduction

The literature reviewed in this chapter has guided the formulation of the research aim (outlined in chapter 4) and is divided into five parts. The literature outlined in the first three parts has contributed to the development of each of the three research objectives (outlined in chapter 4). Section 2.2 relates to objective 1 and outlines an appraisal of the body of research investigating HRB clustering and factors that have been found to predict HRB cluster membership. Section 2.3 relates to objective 2, presenting literature pertaining to the relationship between HRBs in adulthood and different aspects of social circumstances in childhood and adulthood. Section 2.4 relates to objective 3 and describes research investigating the stability of HRBs during mid-adulthood. Section 2.5 of this chapter gives a brief summary of the literature reviewed and section 2.6 outlines gaps in knowledge.

2.2 An appraisal of health-related-behaviour clustering research

A growing body of research evidence indicates that four HRBs, smoking, alcohol, diet and physical activity, cluster together and do not co-occur by chance. Much of this work is succinctly described in two recently conducted literature reviews. The first was undertaken by Noble et al. (2015) which identified fifty-six studies conducted internationally exploring the clustering of smoking, nutrition, alcohol and physical activity risk factors. The second was conducted by Meader et al. (2016) who focused on studies conducted in the United Kingdom investigating the clustering of negative HRBs, including alcohol misuse, physical inactivity, poor diet and smoking alongside a wider range of risk behaviours (such as sexual behaviour, drug misuse and not wearing a seatbelt), identifying thirty-seven studies in total.

Through examination of the references lists from these two literature reviews (Meader et al., 2016; Noble et al., 2015) as well as conducting literature searches in several databases,[1] twenty-one studies[2] investigating the clustering of all four HRBs within working-age adult population samples were appraised. A summary of these studies can be found in Table 2.1.

[1] Including Web of Science, PubMed, Ovid, MEDLINE, Cochrane Library, CINAHL, AMED.

[2] Thirteen were identified via literature search engines and reference lists between 2013 and 2017, three of which are also mentioned by Meader et al. (2016). Seven additional studies were identified via the literature review conducted by Noble et al. (2015).

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples									
Author and Year	**Smoking**	**Alcohol**	**Diet**	**Physical Activity**	**Analysis type**	**Sample**	**Total N**	**Associations found**	**Limitations**
Bécue-Bertaut et al. (2008)	Smokers Former smokers Non-smokers	Drinkers Occasional drinkers Former drinkers Non-drinkers	Open ended question 'what did you eat yesterday?'	Physically inactive Walking Sport Running	Cluster analysis	Croatian population (age 18+)	5,048	Seven clusters identified C1 'Drinkers/former drinkers, former smokers, physically inactive' C2 'Drinkers/former drinkers, smokers/former smokers, walking/running' C3 'Drinkers, smokers/former smokers, sport/running' C4 'Non-drinkers, non-smokers, physically inactive, healthy diet [+poor SRH]' C5 'Non-drinkers, non-smokers, physically inactive [+good SRH]' C6 'Occasional drinkers, smokers, walking, healthy diet [+good SRH]' C7 'Occasional drinkers, smokers, walking [+excellent SRH]' SEP Covariates: Education Economic status (i.e. employment)	Lack of information on classification criterion used for health-related behaviour variable categories. Difficult to compare with other studies as a result.

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples									
Author and Year	**Smoking**	**Alcohol**	**Diet**	**Physical Activity**	**Analysis type**	**Sample**	**Total N**	**Associations found**	**Limitations**
Berrigan et al. (2003)	Never/Former smokers combined. Current smokers.	Average number of alcoholic drinks consumed per day and episodes of binge drinking (≥5 drinks in one setting) in previous 12 months. Adherence to recommendations: ≤2 drinks per day (men) ≤1 drinks per day (women) And no episodes of binge drinking in previous 12 months.	Portions of fruit and vegetables consumed per day (via food frequency questionnaire). % of fat in diet (recall of fat consumption in last 24 hours). Adherence to recommendations: ≥5 portions of fruit and vegetables consumed per day. <30% total calories from fat.	Frequency and intensity of leisure-time physical activity. Being physically active: 3 x vigorous activity (METs > 6) per week. Or 5 x moderate activity (METs 3–6) per week.	Prevalence Odds Ratio	American population (age 20+, mean=46)	15,425	Patterns identified: Smoking and alcohol. Smoking, alcohol and diet. All four HRBs cluster (both positive and negative). Covariates: Education Income	Dichotomised health-related behaviour measures. Using prevalence odds ratios to determine clustering.
Bondy and Rehm (1998)	Non-smoker Current smoker	Alcohol consumption in last 7 days. Average frequency of alcohol use in last 12 months.	Percentage of energy from fat.	Level of total physical activity. Energy expenditure units.	Cluster analysis	American population (age 35–64)	6,060	Men = 9 clusters 'Yuppies1' 'Special event drinker' 'Yuppies2' 'Active middle-agers' 'Gourmands' 'Problem drinkers' 'Poor and unwell' 'No check-ups' 'Over-doers' Women = 8 clusters 'Older and wiser' 'Hard drinkers' 'Yuppies' 'Alcohol problems' 'Jocks' 'Weekenders' 'Negative' 'Educated Jocks' SEP Covariate=Income	Large number of clusters. Query application to health-related behaviour policy/intervention?

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples									
Author and Year	**Smoking**	**Alcohol**	**Diet**	**Physical Activity**	**Analysis type**	**Sample**	**Total N**	**Associations found**	**Limitations**
Conry et al. (2011)	Non-smokers (not smoking now and <100 cigarettes in lifetime) Former smokers (not smoking now and >100 cigarettes in lifetime) Current smokers (> 100 cigarettes in lifetime)	Alcohol use disorder identification test used. (AUDIT-C) Score of 1–5 indicates moderate drinking. Score of 6+ indicates hazardous drinking.	Food Frequency Questionnaire. Multiple food groups included. Dietary approaches to stop hypertension (DASH) score used. (1=poor; 5=excellent)	International Physical Activity Questionnaire. IPAQ gives continuous MET scores. Low=little/no exercise Moderate =5 days per week, 30 minutes, 600 METs per week. Vigorous=3 days per week, 1500 METs per week.	Cluster analysis	Irish population (age 18+)	10,278	6 clusters identified 'Former Smokers' 'Temperate' 'Physically Inactive' 'Healthy lifestyle' 'Multiple risk factor' 'Mixed lifestyle' Covariates: Occupational social class	Only one information criterion used to select the number of clusters (BIC).
De Vries et al. (2008)	Non-smokers. Occasional and daily smokers combined.	Dutch Quantity-Frequency-Variability Questionnaire. ≤3 glasses alcohol per day for men = adherence to norm. ≤2 glasses alcohol per day for men = adherence to norm.	Fruit and Vegetable consumption. ≥2 pieces of fruit per day = adherence to norm. ≥200g vegetables per day = adherence to norm.	Frequency of activity for at least 10 minutes. Engaging in 30 minutes activity 5 days a week = adherence to norm.	Latent Class Analysis	Dutch population (age 12+)	9,449	3 clusters identified 'Healthy' 'Unhealthy' 'Poor Nutrition' Covariates: Education	Dichotomised health behaviour measures.

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples									
Author and Year	**Smoking**	**Alcohol**	**Diet**	**Physical Activity**	**Analysis type**	**Sample**	**Total N**	**Associations found**	**Limitations**
Falkstedt et al. (2016)	Daily smoker Non-daily smoker 'Do you currently smoke tobacco on a daily basis?'	'Risky' vs 'non-risky' alcohol consumption. Risky alcohol consumption defined as consuming > 168 g 100% alcohol per week for men and > 108 g 100% alcohol per week for women. Or consuming the equivalent of half a bottle of spirits (35 cl) on the same occasion at least once a month.	Meeting recommended levels of fruit and vegetable consumption. Non-adherence to recommended levels defined as less than 4 servings of fruit and vegetables per day (approximately 400g per day).	Below recommended levels of physically activity. Recommended levels of leisure physical activity defined as ≥150 min of moderate intensity or ≥75 min of high intensity activity per week or a combination of both.	Prevalence Odds Ratio	Swedish population (age 30–65)	24,241	Having none and having all four health-risk behaviours were found to cluster. Information not provided on the clustering of specific HRBs. Covariates: Participant's education. Participant's parents' education.	Dichotomised health behaviour measures. Using prevalence odds ratios to determine clustering. No information on the clustering of specific HRBs.
French et al. (2008)	Current Smoker Non-smoker	Number of drinkers per day in the past week. Unsafe drinker is ≥5 or ≥3 drinks on any day for males or females.	Average number of portions of fruit and vegetables per day. Low consumption is <2 servings or <cups of fruit and vegetables per day.	Number of days of general or moderate physical activity in the past week. Inactive is <5 or <3 days of general or moderate physical activity.	Cluster analysis	Australian population (age 16–69)	8,668	Four clusters identified 'Safe' 'Moderate' 'Risky smokers' 'Risky drinkers' SEP Covariates: Income	Dichotomised health behaviour measures.

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples									
Author and Year	**Smoking**	**Alcohol**	**Diet**	**Physical Activity**	**Analysis type**	**Sample**	**Total N**	**Associations found**	**Limitations**
Heroux et al. (2012)	Current smokers. Non-smokers.	Number of drinks consumed in past week. Non/light drinkers grouped. Heavy drinkers indicated by: ≥8 drinks for women ≥15 drinks for men	3-day diet record. Coded using Cooper Clinic Nutrition and Exercise Evaluation System. Multiple food groups included. Scores categorised 'Unhealthy diet' = 5th quintile. 'Healthy diet' = 1st-4th quintile.	Frequency, distance and time spent exercising. Categorised into 'Physical active' =doing sports or walk/jog 1–16km per week 'Physical inactive' =no sports or walk/jog 0 km per week Cardiorespiratory fitness (measured using treadmill exercise) 'Unfit' indicated by: Lowest 20%	Latent Class Analysis	US population (age 20–84)	13,621	2 clusters identified 'Healthy' 'Unhealthy' (particularly diet) SEP Covariates: None	Dichotomised health behaviour measures.
Laaksonen et al. (2001)	Non-smokers. Occasional and daily smokers combined.	Units of alcohol per week. 'Unhealthy alcohol consumption' indicated by: >8 units per week for men >5 units per week for women	'Unhealthy diet' = scoring 2/3 on unhealthy choices index fresh vegetables <3 times a week using butter on bread drinking whole milk eating >5 slices of bread per day	'Unhealthy physical activity' = Leisure-time physical activity for 30 minutes duration < once per week.	Prevalence Odds Ratio	Finish population (age 20–64)	22,745	Unhealthy behaviours positively associated. Alcohol consumption and diet had an inverse association. Presence of other health behaviours modified strength of associations between 2 behaviours. Direction of alcohol and physical activity depended on presence of other health behaviours. SEP Covariate: Education	Dichotomised health behaviour measures. Using prevalence odds ratios to determine clustering.

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples									
Author and Year	**Smoking**	**Alcohol**	**Diet**	**Physical Activity**	**Analysis type**	**Sample**	**Total N**	**Associations found**	**Limitations**
Maibach et al. (1996)	Average number of cigarettes per day	Average frequency of alcohol consumption. Average number of drinks per drinking day.	Average number of portions of fruit and vegetables per day. Fat consumption. Frequency of eating desserts.	Days per week of 20 minutes vigorous physical activity. Days per week of moderate physical activity and average duration.	Cluster analysis	American population (adults – age range not reported)	2,910	Seven clusters identified 'Physical fantastics' 'Active attractives' 'Tense but trying' 'Decent dolittles' 'Passively healthy' 'Hard living hedonists' 'Non-interested nihilists' SEP Covariates: Household income Education	No covariates included in the analysis. Age range not reported.
Patterson et al. (1994)	Average number of cigarettes per day.	Average number of drinks per week.	Diet quality score (24-hour diet recall and 2-day food diary).	Frequency of leisure-time physical activity.	Cluster analysis	American population (age 21+)	1,463	7 clusters identified 'Health promoting' 'Good diet' 'Fitness' 'Passive' 'Drinker' 'Smoker' 'Hedonic' SEP Covariates: Education Income	Only measures quantity of alcohol consumption (no frequency). Only measures frequency of leisure-time physical activity (no other forms of physical activity considered or duration/intensity of activity).

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples									
Author and Year	**Smoking**	**Alcohol**	**Diet**	**Physical Activity**	**Analysis type**	**Sample**	**Total N**	**Associations found**	**Limitations**
Poortinga (2007)	Non-smokers Light/moderate/ heavy smokers combined	Units of alcohol per week. 'Unhealthy alcohol consumption' indicated by: >8 units on at least one day per week for men >6 units on at least one day per week for women	Fruit and vegetable consumption. 'Unhealthy diet' indicated by: <5 portions or fruit and vegetables per day	Frequency, duration, intensity of physical activity (includes leisure-time, commuting, occupational, domestic) 'Unhealthy physical activity' indicated by: <30 minutes moderate activity 5 days per week in the last 4 weeks.	Prevalence Odds Ratio	UK population (age 16–64)	11,492	Smoking and heavy alcohol consumption cluster strongly with unhealthy diet. Lack of physical activity cluster with unhealthy diet. Inverse association between smoking and physical activity. Inverse association between alcohol and physical activity. SEP Covariates: Occupation Housing tenure Employment.	Dichotomised health behaviour measures. Using prevalence odds ratios to determine clustering.
Schneider et al. (2009)	Non-smokers and occasional smokers combined. Daily smokers.	Alcohol (in grams) consumed over past weekend and last day of week. 'Unhealthy alcohol consumption' indicated by: >10g per day for women >20g per day for men	Fruit and Vegetable consumption. 'Unhealthy diet' indicated by: <250g fruit and vegetables per day	Nature, duration, frequency and intensity of physical activity (sports only) 'Unhealthy physical activity' indicated by: Sporting activity <1hr per week for a whole year.	Cluster analysis	German population (age 50–70)	2,002	5 clusters identified: 'No risk behaviours' 'Physically inactive' 'Fruit and Vegetable avoiders' 'Smokers with risk behaviours' 'Drinkers with risk behaviours' SEP Covariates: Income, education, occupation.	Dichotomised health behaviour measures. No information criterion used to determine number of clusters (R-squared, pseudo F, t-statistics used) Small sample (no power calculation provided). Restricted to those aged 50+.

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples

Author and Year	Smoking	Alcohol	Diet	Physical Activity	Analysis type	Sample	Total N	Associations found	Limitations
Schuit et al. (2002)	Non-smokers Smokers (≥1 cigarette per month)	Glasses of alcohol per day. 'Unhealthy alcohol consumption' indicated by: >3 glasses of alcohol per day for men >2 glasses of alcohol per day for women	Fruit and Vegetable consumption 'Unhealthy diet' indicated by: <350g vegetables and fruit per day.	Duration and intensity of physical activity (includes commuting, leisure, domestic). 'Unhealthy physical activity' indicated by: <30 minutes moderate or vigorous physical activity per day	Prevalence Odds Ratio	Dutch population (age 20–65)	16,789	All behaviours associated with each other. Except heavy alcohol consumption and low physical activity (no association found). Alcohol and smoking had the strongest association. SEP Covariates: Education	Dichotomised health behaviour measures. Using prevalence odds ratios to determine clustering.
Silva et al. (2013)	Non-smoker and former smokers combined. Current smokers.	Alcohol use disorder identification test used. (AUDIT) Score ≥8 indicates problematic alcohol consumption.	Fruit and vegetable consumption. 'Unhealthy diet' indicated by: <5 portions or fruit and vegetables per day	Frequency of physical activity (leisure-time only). 'Unhealthy physical activity' indicated by: <once per week in the last 3 months.	Prevalence Odds Ratio	Brazil population (age 20–59)	1,720	All behaviours associated with each other. Most common clustering was smoking, unhealthy diet, low physical activity. Strongest clustering was smoking, problematic alcohol use and unhealthy diet. SEP Covariates: Occupational class Income Education.	Dichotomised health behaviour measures. Using prevalence odds ratios to determine clustering. Small sample (power calculation is provided).

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples

Author and Year	Smoking	Alcohol	Diet	Physical Activity	Analysis type	Sample	Total N	Associations found	Limitations
Slater and Flora (1991)	Average number of cigarettes smoked per day.	Average frequency of consuming alcoholic drinks.	Dietary intake assessed using 9-item index (classification of unhealthy diet not given)	Index score of positive responses to physical activity involving walking, climbing stairs or vigorous exercise (classification of low physical activity not given)	Cluster analysis	American population (mean age 37, no age range given)	2,502	7 clusters identified (2 discarded – very small) 'Healthful adults' 'Unhealthful adults' 'Worried older adults' 'Healthful talkers' 'Healthful young adults' 'Unhealthful young adults' 'Young athletes' SEP Covariates: Education Income Household size	Dichotomised health-related behaviours used for diet and physical activity. Classification of 'unhealthy' diet / 'low' physical activity not given.
Tobias et al. (2007)	Non-smokers and occasional smokers combined. Daily smokers.	Alcohol use disorder identification test used. (AUDIT) Score ≥8 indicates problematic alcohol consumption.	Fruit and vegetable consumption. 'Unhealthy diet' indicated by: <5 portions or fruit and vegetables per day	Duration, frequency, intensity of physical activity (nature of activity not stated). 'Unhealthy physical activity' indicated by: <150 minutes per week of moderate intensity (3 METs)	Prevalence Odds Ratio	New Zealand population (age 15+)	10,241	Unhealthy behaviours had a stronger association compared to health behaviours. Healthy behaviours associated with one another. Except non-smoking and healthy physical activity. Healthy drinking and healthy physical activity. Unhealthy behaviours associated with one another. Strongest clustering was smoking, unhealthy alcohol and unhealthy diet. SEP Covariates: Deprivation index (9 SES variables)	Dichotomised health behaviour measures. Using prevalence odds ratios to determine clustering.

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples									
Author and Year	**Smoking**	**Alcohol**	**Diet**	**Physical Activity**	**Analysis type**	**Sample**	**Total N**	**Associations found**	**Limitations**
Tseng and Lin (2008)	Average number of cigarettes per day.	Average number of drinks per month.	Frequency of fruit and vegetable consumption (days per week).	Minutes of physical activity per week.	Factor analysis	Taiwanese population (age 15+)	26,755	Two factors identified 'Risky' 'Protective' within each age group (15–24; 25–39; 40–54; ≥55) for men and women SEP Covariates: None	Lack of comparability with other study results due to analysis type. Only gender and age considered in the analysis.
van Nieuwenhuijzen et al. (2009)	Total number of cigarettes per day (continuous)	Number of days per week alcohol consumed (continuous) Total number of glasses consumed a day (range 0–7) Heavy drinking variable (number of days x number of glasses) >4 indicates heavy drinking.	Fruit and Vegetable consumption Number of days per week x number portions (continuous)	Frequency, duration, intensity of physical activity (includes leisure, commuting, domestic, occupation) Exercise Light = 2–4 METs Exercise Moderate = 4–6.5 METs Exercise vigorous = ≥6.5 METs	Confirmatory Factor Analysis	Dutch population (age 12–40)	4,395	3 clusters identified for adults (19–40) 'Health' 'Alcohol' 'Delinquency' Alcohol consumption identified as a separate cluster from smoking, diet and exercise. Negative association found between 'Alcohol' and 'Health' clusters. SEP Covariates: Education	Included other health behaviours in model (sleep, unsafe sex, drug use, unsafe driving) Lack of comparability with other study results due to analysis type. Restricted to those aged ≤40.

Table 2.1 Summary of studies used in the appraisal of HRB clustering research

Cross-sectional studies investigating the clustering of four health behaviours (smoking, alcohol, diet, physical activity) in adults of working-age population samples									
Author and Year	**Smoking**	**Alcohol**	**Diet**	**Physical Activity**	**Analysis type**	**Sample**	**Total N**	**Associations found**	**Limitations**
Verger et al. (2009)	Non-smokers and occasional smokers combined. Daily smokers.	Alcohol use disorder identification test used. (AUDIT) Regular users indicated by: Consumption at least 4–5 days per week. Binge users indicated by: Consuming 6 glasses on the same occasion at least twice a month.	Fruit and green vegetable consumption. 'Unhealthy Diet' indicated by: <Daily or almost daily consumption.	Physical activity frequency. 'Unhealthy Physical Activity' indicated by: Stating 'no' to one question *'Do you regularly engage in some sport or activity that you would view as sport in terms of intensity or duration?'*	Cluster analysis	French population (age 18+)	17,355	Five clusters identified 'Healthy lifestyle' 'Non-consumers of fruit and green vegetables' 'Regular alcohol consumers' 'Smokers' 'Frequent binge drinkers' SEP Covariates: Household income Occupational class	Dichotomised health behaviour measures.
Vermeulen-Smit et al. (2015)	Average number of cigarettes smoked per day in the last 4 weeks. None Less than one a week Less than one a day 1–5 per day 6–10 per day 11–20 per day More than 20 per day	Average frequency of alcohol consumption in last 12 months. And Frequency of drinking ≥5 drinks in one day in last 12 months. Every day Nearly every day 3 to 4 days a week 1 to 2 days a week 1 to 3 days a month less than once a month Average number of drinks per drinking day. Range = 0 – 20.3	Average number of days per week of groups consumed: Breakfast; Vegetables; Fresh fruit; Salty snacks; Sweet snacks; Deep-fried meals; Sweet soft drinks. Ranging from never/<once per week (0) to once per week (7).	Hours per week spent doing physical activity/sport? Range 0–30.	Latent Class Analysis	Dutch population (age 21–67)	5,303	Four clusters identified: Class 1: most healthy, mainly non-smokers, moderate drinkers, active, healthy diet. Class 2: smokers, moderate drinkers, inactive, unhealthy diet. Class 3: smokers, heavy episodic drinkers, active, unhealthy diet. Class 4: smokers, frequent heavy drinkers, active, low fruit. SEP Covariates: Education, employment	

2.2.1 Methodological considerations

The methodologies of the twenty-one studies investigating HRB clustering have been appraised and have led to a number of factors being considered when comparing the study findings and in the development of this doctoral research. These factors are outlined below and relate to the treatment of HRB variables, sample characteristics and the statistical analyses employed.

2.2.1.1 *Dichotomising HRB variables*

In the appraised HRB clustering literature, it appears that measures of the HRBs are commonly dichotomised into healthy and unhealthy categories. This practice leads to a substantial loss of information within each study. For example, studies dichotomised alcohol use into those drinking heavily compared to everyone else (De Vries et al., 2008; Falkstedt et al., 2016; Heroux et al., 2012; Laaksonen et al., 2001; Poortinga, 2007; Schneider et al., 2009; Schuit et al., 2002; Silva et al., 2013; Tobias et al., 2007). This comparison does not consider the differences between those who do not drink and those who drink light or moderate amounts. This is important, given that these groups may differ in terms of their socio-demographic characteristics (Bellis et al., 2016; Marmot, 1997; Power et al., 1998).

Moreover, comparing the findings of these studies is problematic because dichotomised cut points vary, depending on how the HRB is conceptualised by the researchers. For example, some studies combine occasional and daily smokers together (De Vries et al., 2008; Laaksonen et al., 2001) while others combine occasional and non-smokers together (Schneider et al., 2009; Tobias et al., 2007; Verger et al., 2009).

Where possible, HRBs should not be dichotomised and instead treated as continuous or ordered. This approach not only increases statistical power (MacCallum et al., 2002), but more importantly it maximises the rich information in the data, potentially identifying clusters that may have been missed if variables were dichotomised (McAloney et al., 2013).

2.2.1.2 *Statistical analyses used to identify HRB clustering*

Different statistical techniques have been used across the studies to determine clustering. Statistical techniques include prevalence odds ratios (Berrigan et al., 2003; Falkstedt et al., 2016; Laaksonen et al., 2001; Poortinga, 2007; Schuit et al., 2002; Silva et al., 2013; Tobias et al., 2007), cluster analysis (Bécue-Bertaut et al., 2008; Bondy and Rehm, 1998; Conry et al., 2011; French et al., 2008; Maibach et al., 1996; Schneider et al., 2009; Slater and Flora, 1991; Verger et al., 2009), latent class analysis (De Vries et al., 2008; Heroux et al., 2012; Vermeulen-Smit et al., 2015) and confirmatory factor analysis (Tseng and Lin, 2008; van Nieuwenhuijzen et al., 2009). The latter three, i.e. cluster analysis, latent class analysis and confirmatory factor analysis, are data reduction techniques (Hofstetter et al., 2014).

Data reduction techniques may be preferable to prevalence odds ratios because they maximise efficiencies in the data by summarising information on multiple HRB variables into a select few (Hofstetter et al., 2014). Prevalence odds ratios are intended to capture the odds of having one negative HRB, given the prevalence of another negative HRB (Schuit et al., 2002). Prevalence odds ratios compare the observed prevalence of HRBs with the expected prevalence. If the prevalence odds ratio is above one, this indicates the clustering of HRBs[3] (Schuit et al., 2002). This method of analysis only allows researchers to treat HRB variables as dichotomous, potentially leading to a loss of rich information (McAloney et al., 2013). Furthermore, prevalence odds ratios have to be calculated for each potential combination of HRBs (McAloney et al., 2013). Unlike prevalence odds ratios, data reduction techniques do not require each HRB variable to be dichotomised into categories such as 'risky' and 'non-risky'. Instead, data reduction techniques can incorporate a combination of dichotomous, ordinal and continuous variables into the model (Wang and Wang, 2012). Using variables with an array of categories retains valuable information on individual differences that may be present in the data (MacCallum et al., 2002).

These three data reduction techniques (cluster analysis, latent class analysis, confirmatory factor analysis) can be further divided into those that are person-centred or variable-centred (Hofstetter et al., 2014). Cluster analysis and latent class analysis are person-centred approaches whilst confirmatory analysis is variable-centred. Person-centred approaches conceptualise interrelationships between HRBs as clustering together, with HRB patterns being shared by subgroups of the population (McAloney et al., 2013). This implies that an individual's HRBs are represented through belonging to a single group. In contrast variable-centred approaches consider the interrelationship between HRBs to be a single continuum upon which individuals are placed (Hofstetter et al., 2014).

A person-centred, rather than variable-centred, approach may be preferable because unlike other processes, such as ageing or cognitive decline, lifestyle is not an inevitable linear process. Consequently, HRB clustering may be better considered as categorical and therefore person-centred, with each category reflecting distinct groups who are distinguished from one another

[3] For example a POR of 2.5 for smoking and heavy alcohol consumption clustering indicates that individuals who smoke have 2.5 higher odds of heavy alcohol consumption in comparison to individuals who do not smoke.

by different characteristics, rather than a purely linear and variable-centred process which implies a severity continuum (Stapinski et al., 2016).

Whilst latent class analysis and cluster analysis are both considered acceptable person-centred data reduction methods which can be used to identify subgroups who share HRB patterns (McAloney et al., 2013), latent class analysis may be preferable to cluster analysis. Unlike cluster analysis, latent class analysis has a number of statistical procedures to guide the selection of clusters (Wang and Wang, 2012) and is more flexible when working with large sample sizes and non-standardised variables (Hofstetter et al., 2014). Moreover, latent class analysis is considered to have stronger theoretical underpinnings because it is based upon probability modelling rather than cluster analysis which uses distance to merge individuals into clusters (Hofstetter et al., 2014). Finally, latent class analysis is used to identify an unobserved categorical latent variable conceived to explain associations between observed variables (Collins and Lanza, 2010). This latent variable approach minimises potential measurement error which is inherent in observed measures (Hagger-Johnson et al., 2011).

2.2.2 Health-related-behaviour cluster patterns

Despite the different methodologies of the twenty-one studies (see Table 2.1) outlined above (see section 2.2.1), all of them found underlying associations between these four HRBs, identifying distinct behavioural patterns of smoking, alcohol consumption, diet and physical activity. These findings support theories that suggest the collective patterning of HRBs (Blue et al., 2014; Cockerham, 2005).

The studies found that positive[4] HRBs tend to cluster, as do negative[5] HRBs. This supports the concept of binary lifestyles which consist of either positive or negative HRBs (Cockerham, 2005). The HRBs can be distinguished according to participation and maintenance. Some HRBs require active participation and can be hard to maintain (i.e. diet and physical activity) whilst others require restraint and may be difficult to reduce (i.e. alcohol use and smoking) (Borland, 2013e). For example, twenty of the studies appraised identified the clustering of a poor diet and physical inactivity (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Bondy and Rehm, 1998; Conry et al., 2011; De Vries et al., 2008; French et al., 2008; Heroux et al., 2012; Laaksonen et al., 2001; Maibach et al., 1996; Patterson et al., 1994; Poortinga, 2007; Schneider et al., 2009; Schuit et al., 2002; Silva et al., 2013; Slater and Flora, 1991; Tobias et al., 2007;

[4] Broadly defined as not smoking; none, light or moderate alcohol consumption; moderate to high levels of physical activity; low intake of fat and sugar; daily consumption of fruit and vegetables.
[5] Broadly defined as smoking daily; heavy alcohol consumption; low levels of physical activity; high intake of fat and sugar; less than daily consumption of fruit and vegetables.

Tseng and Lin, 2008; van Nieuwenhuijzen et al., 2009; Verger et al., 2009; Vermeulen-Smit et al., 2015). The remaining study by Falkstedt et al. (2016) only provided information on the number of negative HRBs and the corresponding prevalence odds ratios, rather than providing information pertaining to specific combinations of HRBs. Therefore, the extent to which these two HRBs clustered together could not be verified.

The majority of the appraised studies also found the clustering of smoking and heavy alcohol consumption (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Bondy and Rehm, 1998; Conry et al., 2011; De Vries et al., 2008; Heroux et al., 2012; Laaksonen et al., 2001; Maibach et al., 1996; Patterson et al., 1994; Poortinga, 2007; Schneider et al., 2009; Schuit et al., 2002; Slater and Flora, 1991; Tobias et al., 2007; Tseng and Lin, 2008; Verger et al., 2009; Vermeulen-Smit et al., 2015). Again, information pertaining to the clustering of these two specific HRBs could not be obtained from Falkstedt et al. (2016).

However, two studies found no clustering between smoking and alcohol consumption (Silva et al., 2013; van Nieuwenhuijzen et al., 2009). The findings of Silva et al. (2013) suggested the odds of smoking and problematic alcohol use was lower than would be expected randomly (indicated by a prevalence odds ratio <1). The authors suggest that possible reasons for any discrepancies between their findings and those of other HRB cluster studies could be that their study was conducted in a developing as opposed to a fully developed country. The vast majority of the studies included in this appraisal and in the literature reviews of studies investigating HRB clustering (Meader et al., 2016; Noble et al., 2015) are conducted in developed countries. Van Nieuwenhuijzen et al. (2009) found alcohol consumption sat within a separate cluster to smoking, diet and physical activity. In this study, confirmatory factor analysis was used to determine clustering, treating alcohol and smoking variables continuously as opposed to other studies which dichotomised these variables. This may help to explain discrepancies given their different conceptualisations of behavioural patterning. Confirmatory factor analysis is a variable-centred approach whereas latent class and cluster analysis are person-centred approaches (Hofstetter et al., 2014).

Additionally, the appraised studies identified the clustering of multiple negative HRBs, combining HRBs that are considered hard to maintain (i.e. diet and physical activity) with those that are hard to reduce (i.e. alcohol use and smoking) (Borland, 2013e). For example, many of the studies identified that individuals who smoke and drink alcohol heavily also have poorer diets (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Bondy and Rehm, 1998; Conry et al., 2011; French et al., 2008; Laaksonen et al., 2001; Maibach et al., 1996; Patterson et al., 1994;

Poortinga, 2007; Schuit et al., 2002; Silva et al., 2013; Slater and Flora, 1991; Tobias et al., 2007; Verger et al., 2009; Vermeulen-Smit et al., 2015) and lower levels of physical activity (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Conry et al., 2011; French et al., 2008; Laaksonen et al., 2001; Silva et al., 2013; Tobias et al., 2007; Verger et al., 2009).

However, the clustered patterns of these negative HRBs were not found universally. Three studies did not identify clustering of smoking, alcohol consumption and fruit and vegetable consumption (De Vries et al., 2008; Schneider et al., 2009; Verger et al., 2009). Instead, these studies found that low fruit and vegetable consumption sat within a separate cluster to alcohol and smoking behaviour. These studies all used data reduction techniques to determine clustering. One used latent class analysis (De Vries et al., 2008), the other two used cluster analysis (Schneider et al., 2009; Verger et al., 2009). Notably, these three studies (De Vries et al., 2008; Schneider et al., 2009; Verger et al., 2009) also used lower thresholds to identify participants with 'healthy' fruit and vegetable consumption in comparison to the rest of the studies included in this appraisal. For example, these studies found 71% (De Vries et al., 2008), 29% (Schneider et al., 2009) and 18% (Verger et al., 2009) of their population samples did not have a 'healthy' fruit and vegetable consumption. In constrast, three studies (Poortinga, 2007; Silva et al., 2013; Tobias et al., 2007) using higher thresholds to define 'healthy' fruit and vegetable consumption found 76% (Poortinga, 2007), 46% (Tobias et al., 2007) and 81% (Silva et al., 2013) of their population samples did not have a 'healthy' fruit and vegetable consumption. The use of a higher threshold to define 'healthy' fruit and vegetable consumption has the potential to underestimate the influence of poor diet on cluster formation. This illustrates how dichotomising HRB variables can be problematic when comparing findings.

Further discrepancies in study findings pertaining to the clustering of multiple negative HRBs are apparent. For example, Laaksonen et al. (2001) identified that the clustering of heavy alcohol consumption, smoking daily and low fruit and vegetable consumption was only present in women (indicated by a prevalence odd ratio above 1). The authors suggested that in the overall sample the clustering of smoking, alcohol and physical activity was stronger than the clustering of alcohol, smoking and fruit and vegetable consumption. This contrasts with other studies that find no association between smoking, alcohol consumption and physical activity (Conry et al., 2011; Heroux et al., 2012; Poortinga, 2007; Schneider et al., 2009; Schuit et al., 2002; Silva et al., 2013; Tobias et al., 2007).

Again, the utilisation of different statistical techniques across the studies makes comparing their findings difficult. For example, research conducted by Schneider et al. (2009) and Conry et al. (2011) found physical inactivity to sit within a different cluster to that of heavy alcohol consumption and smoking daily, which contrasts with the findings of Laaksonen et al. (2001). Notably, these studies utilised cluster analysis (a data reduction technique) to determine clustering, whereas Laaksonen et al. (2001) used prevalence odds ratios. Moreover, Laaksonen et al. (2001) used a lower threshold to define 'healthy' physical activity compared to other studies that found no association between smoking, alcohol consumption and physical activity (Conry et al., 2011; Heroux et al., 2012; Poortinga, 2007; Schneider et al., 2009; Schuit et al., 2002; Silva et al., 2013; Tobias et al., 2007).

Alongside HRB patterns indicating that positive HRBs tend to cluster, as do negative ones, the studies also indicate complexity, identifying inverse associations between positive and negative HRBs. This provides some evidence that HRB clustering extends beyond a binary concept of health lifestyles (Cockerham, 2005). Amongst the appraised studies there was an association between heavy alcohol consumption and higher levels of physical activity (Bécue-Bertaut et al., 2008; Bondy and Rehm, 1998; Laaksonen et al., 2001; Maibach et al., 1996; Patterson et al., 1994; Poortinga, 2007; Slater and Flora, 1991; Vermeulen-Smit et al., 2015). This research also suggests higher levels of physical activity in smokers and former smokers compared to non-smokers (Bécue-Bertaut et al., 2008; Conry et al., 2011; Maibach et al., 1996; Patterson et al., 1994; Poortinga, 2007) and that smokers and heavy drinkers were less likely to consume sweet snacks (Bécue-Bertaut et al., 2008; Maibach et al., 1996; Vermeulen-Smit et al., 2015). Moreover, some studies have found an association between high levels of physical activity and a poor diet (Bécue-Bertaut et al., 2008; Maibach et al., 1996; Patterson et al., 1994).

The inverse associations between positive and negative HRBs highlighted above have also been evidenced elsewhere. For example, studies indicate a relationship between heavy alcohol consumption and higher levels of physical activity (French et al., 2009; Leasure et al., 2015). Scholars suggest that these behaviours may become mutually dependent in particular social contexts (Blue et al., 2014), such as drinking alcohol after participating in sport (French et al., 2009). There is also evidence of a relationship between reduced sugar intake and smoking (Crawley and While, 1996; Méjean et al., 2011; O'Doherty et al., 2011; Whichelow et al., 1991), although some have found this only to be the case for light and moderate smokers as opposed to heavy smokers (Iredale et al., 2016). It is conceived that smoking may have an effect on dietary preferences (Colditz et al., 1991; Lampure et al., 2014), by influencing taste buds (Iredale et al., 2016). Eating less sugar has also been found to be associated with heavy alcohol consumption when compared to abstainers (Méjean et al., 2011).

One explanation for these inverse associations between positive and negative HRBs comes from qualitative research which suggests that individuals 'balance out' positive and negative HRBs (Backett and Davison, 1995), known as compensatory health beliefs (Knäuper et al., 2004). It may therefore be possible that people who drink alcohol heavily and/or smoke may be aware that these HRBs could be damaging for health and compensate by partaking in HRBs conceived to be health promoting, such as doing more exercise and/or eating less sugar. For example, the inverse patterns of alcohol consumption and sugar may reflect a replacement of sugar intake with alcohol use (Colditz et al., 1991).

As mentioned in section 2.2.1, whilst the inverse associations between HRBs elucidated in the appraised HRB cluster studies are insightful, comparing their findings directly is difficult because different statistical techniques are used to determine HRB clustering. Moreover, dichotomising HRB measures into healthy and unhealthy categories reduces comparability across the studies and may be too simplistic, leading to a loss of rich information which could elucidate more complex HRB clustered patterns that extend beyond a binary concept of health lifestyles (Cockerham, 2005).

2.2.3 Predictors of health-related-behaviour cluster membership

Age, gender and socio-economic position (SEP) are three predictors that are related to HRB clustering (Meader et al., 2016; Noble et al., 2015). The influence of these predictors on HRB clustering can either be in relation to cluster membership (i.e. the same HRB cluster patterns are found to exist across population subgroups but prevalence rates differ) or in relation to the nature of the HRB cluster patterns (i.e. HRB cluster patterns are not equivalent across population subgroups) (Collins and Lanza, 2010).

All of the authors claimed their samples were representative of the populations from which they came. The samples are relatively large, ranging from 1,463 (Patterson et al., 1994) to 26,755 (Tseng and Lin, 2008). These large samples suggest adequate statistical power to detect HRB clustering and the predictive effect of covariates, such as age, gender and SEP. However, Silva et al. (2013) was the only study in this review to mention undertaking a power calculation.

2.2.3.1 *Age*

Of the twenty-one appraised studies investigating the clustering of HRBs, fourteen have explored age differences according to HRB cluster membership (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Bondy and Rehm, 1998; Conry et al., 2011; Patterson et al., 1994; Poortinga, 2007; Schneider et al., 2009; Schuit et al., 2002; Slater and Flora, 1991; Tobias et al., 2007; Tseng and Lin, 2008; van Nieuwenhuijzen et al., 2009; Verger et al., 2009; Vermeulen-Smit et al., 2015) and four according to the nature of the HRB cluster patterns (Laaksonen et al., 2001; Schuit et al., 2002; Tseng and Lin, 2008; van Nieuwenhuijzen et al., 2009).

Studies investigating the relationship between age and HRB cluster membership suggest a higher membership of older participants in clusters characterised by multiple positive HRBs in comparison to younger participants, more likely to be members of HRB clusters characterised by multiple negative HRBs (Berrigan et al., 2003; Conry et al., 2011; Patterson et al., 1994; Poortinga, 2007; Schneider et al., 2009; Tseng and Lin, 2008; Verger et al., 2009; Vermeulen-Smit et al., 2015).

However, inconsistencies in the relationship between age and HRB cluster membership are apparent in relation to physical activity. For example, ten studies identified an HRB cluster characterised by physical inactivity (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Bondy and Rehm, 1998; Conry et al., 2011; Patterson et al., 1994; Schneider et al., 2009; Slater and Flora, 1991; Tseng and Lin, 2008; Verger et al., 2009; Vermeulen-Smit et al., 2015). Three of these studies found a higher membership of older people in this physically inactive HRB cluster

(Berrigan et al., 2003; Patterson et al., 1994; Slater and Flora, 1991), whereas five studies found a higher membership of younger people in HRB clusters characterised by physical inactivity (Bécue-Bertaut et al., 2008; Conry et al., 2011; Tseng and Lin, 2008; Verger et al., 2009; Vermeulen-Smit et al., 2015). By contrast, two studies found no age difference in membership of this physically inactive HRB cluster (Bondy and Rehm, 1998; Schneider et al., 2009).

Four studies investigated age differences in the nature of the HRB patterns (Laaksonen et al., 2001; Schuit et al., 2002; Tseng and Lin, 2008; van Nieuwenhuijzen et al., 2009). Three of these studies found HRB patterns differed according to age (Laaksonen et al., 2001; Schuit et al., 2002; Tseng and Lin, 2008). For example, Schuit et al. (2002) found smoking and heavy alcohol consumption clustered strongly in participants aged 20–29 whereas smoking and low fruit and vegetable consumption clustered strongly in participants aged 50–59. Tseng and Lin (2008) identified that, for men, HRB patterns characterised by multiple negative HRBs were more prevalent in those under 55 and HRB patterns characterised by multiple positive HRBs were more prevalent in those over 55. Laaksonen et al. (2001) found that, in men, the relationship between smoking and diet differed according to age as did alcohol and diet, although the authors do not elaborate on the nature of this relationship. For women, Laaksonen et al. (2001) found little difference in the nature of HRB cluster patterns between age groups. The remaining study found no difference in the nature of HRB cluster patterns amongst younger (19–24) and older (25–40) adults (van Nieuwenhuijzen et al., 2009).

These results highlight the complexity of health-related-behaviour clustering prevalence and the nature of the cluster patterns across age groups. Some studies included individuals as young as 12 years old in their samples (De Vries et al., 2008; van Nieuwenhuijzen et al., 2009) and others up to 84 years old (Heroux et al., 2012), which may contribute to inconsistent findings across the studies.

Inconsistencies in the relationship between age and HRB clustering are also highlighted by the authors of the two systematic literature reviews of research investigating HRB clustering (Meader et al., 2016; Noble et al., 2015). Both reviews indicate that, whilst younger age does appear to predict membership of clusters characterised by multiple negative HRBs, this finding is less consistent across studies in comparison to the other two common predictors of HRB clustering, namely gender and SEP.

However, before turning attention to gender and SEP as predictors of HRB clustering, another factor strongly related to age is considered, namely participant cohort (i.e. year of birth). Notably, whilst many of the appraised studies consider participant age on HRB clustering, none consider the effect of cohort.

2.2.3.2 *Cohort*

The lack of consideration to potential cohort differences in HRB clustering amongst the appraised studies is an important gap in the evidence base. Cohort is considered to influence lifestyle formation because HRBs change over historical time and people's lives play out within a specific historical context (Burton-Jeangros et al., 2015). An example is given by Wadsworth and Kuh (1997):

'It is possible to be born into a period of high prevalence of parental smoking and to have lived middle life in a time of much reduced likelihood of smoking' (Wadsworth and Kuh, 1997: 864).

This perspective is supported by empirical evidence finding cohort differences in the prevalence of smoking, alcohol consumption, diet and physical activity among British adults (Elliott et al., 2007; Schoon and Parsons, 2003; Whitley et al., 2014). Previous work comparing individual HRBs of people born in 1958 and 1970 (Schoon and Parsons, 2003) found some HRBs were better amongst those born in 1970, e.g. eating chips less frequently for both genders and fewer women smokers (Schoon and Parsons, 2003), the latter finding being consistent with declines in the prevalence of smoking over the past 50 years (RCP, 2012). Other HRBs were worse in the later-born cohort, e.g. eating fruit less frequently for both genders (Schoon and Parsons, 2003).

The influence of cohort on HRB change over time is further considered in section 2.4.3 of this chapter.

2.2.3.3 *Gender*

As mentioned above (see section 2.2.3.1), the findings from two literature reviews of HRB clustering research (Meader et al., 2016; Noble et al., 2015) implicate gender as a predictor of HRB clustering. This is not surprising, given that gender differences in HRBs are inextricably linked to gender-specific roles within the family and society (Bartley, 2016b; Devine, 2005; Schoon and Parsons, 2003). Negative HRBs, such as smoking and alcohol consumption, are often associated with masculinity (Fleming and Agnew-Brune, 2015). Historically a higher proportion of men smoked because it was not socially acceptable for women to smoke

(Schoon and Parsons, 2003). Moreover, drinking alcohol heavily and avoiding healthy foods were historically considered to be more normative among men (Fleming and Agnew-Brune, 2015). Some positive HRBs, such as physical activity, are often perceived as unfeminine (Fleming and Agnew-Brune, 2015).

These two literature reviews of evidence regarding HRB clustering suggest that men are more likely to belong to clusters characterised by negative HRBs compared to women. However, one of the authors suggests this relationship between gender and the presence of multiple HRBs is relatively weak (Meader et al., 2016).

Amongst the twenty-one appraised studies investigating HRB clustering (outlined in Table 2.1), sixteen assessed gender differences in HRB cluster membership (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Conry et al., 2011; French et al., 2008; Laaksonen et al., 2001; Maibach et al., 1996; Patterson et al., 1994; Poortinga, 2007; Schneider et al., 2009; Schuit et al., 2002; Silva et al., 2013; Slater and Flora, 1991; Tobias et al., 2007; Tseng and Lin, 2008; Verger et al., 2009; Vermeulen-Smit et al., 2015). All of the studies found a higher membership of women compared to men in clusters characterised by positive HRBs in respect to smoking, diet and alcohol consumption, although the relationships were not always consistent in terms of physical activity. Some studies found a higher membership of women in physically inactive clusters (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Patterson et al., 1994; Poortinga, 2007; Schuit et al., 2002; Silva et al., 2013; Slater and Flora, 1991; Tobias et al., 2007; Verger et al., 2009), others found a higher membership of men in inactive clusters (Conry et al., 2011; French et al., 2008; Tseng and Lin, 2008; Vermeulen-Smit et al., 2015). Moreover, one study found no gender difference in membership of the cluster labelled 'Physically Inactives' (Schneider et al., 2009).

Seven studies also examined gender differences in the nature of HRB cluster patterns (Bondy and Rehm, 1998; Laaksonen et al., 2001; Poortinga, 2007; Schuit et al., 2002; Tobias et al., 2007; Tseng and Lin, 2008; van Nieuwenhuijzen et al., 2009). Two of these studies found the clustering of positive HRBs was stronger amongst women (Tseng and Lin, 2008; van Nieuwenhuijzen et al., 2009), although for one of the studies this was only found amongst adolescents (age 16–18), not adults (age 19–24; age 25–40) (van Nieuwenhuijzen et al., 2009). One study selected different numbers of HRB clusters for men (=9) and women (=8), indicating that HRB patterns were not equivalent across gender groups (Bondy and Rehm, 1998). Three studies found a stronger clustering of negative HRBs, particularly heavy alcohol consumption and smoking daily (indicated by a higher prevalence odds ratio), in women compared to men

(Laaksonen et al., 2001; Poortinga, 2007; Tobias et al., 2007). The remaining study found no gender differences in the nature of HRB cluster patterns (Schuit et al., 2002).

Whilst inconsistencies are apparent, this appraisal of HRB cluster research does indicate that gender has the potential to predict both HRB cluster membership and differences in the nature of the HRB cluster patterns. The view that HRBs differ according to gender is also supported by other empirical evidence suggesting that, compared to men, women consume less alcohol (Britton et al., 2015; Meng et al., 2014), are less likely to smoke (Jarvis, 1994) and those that do smoke consume fewer cigarettes per day (Allen et al., 2016). Compared to men, women are less physically active (Blaxter, 1990; Hunt et al., 2001; Sullivan et al., 2013). Women tend to have healthier diets than men (Wardle et al., 2004) although there is research to suggest that women consume more sugar in comparison to men (Colditz et al., 1991; Méjean et al., 2011; Worsley et al., 2012).

There is also research to suggest historical shifts in HRB and gender relations over time. For example, research undertaken in a British cohort born in 1970 at age 42 suggested that women were found to do substantially less physical activity in comparison to men (Sullivan et al., 2013). However, gender differences in physical activity were not found for women of the same age born 12 years earlier in 1958 (Parsons et al., 2006). Moreover, there is evidence suggesting a convergence of smoking (HSCIC, 2014b; McCartney et al., 2011) and alcohol consumption (Elliott et al., 2007; Keyes et al., 2011; Meng et al., 2014; Purshouse et al., 2017; Schoon and Parsons, 2003; Slade et al., 2016) amongst men and women over time. An increase in alcohol consumption was observed amongst women born in 1970 in comparison to women born in 1958 (Elliott et al., 2007; Schoon and Parsons, 2003). These results suggest a complex relationship between gender and cohort, thus adding further support to the argument raised in section 2.2.3.2 that there is a need to consider the impact of cohort, alongside gender, on HRB clustering (Pavalko and Caputo, 2013).

Alongside empirical evidence, the mechanisms through which gender influences HRB cluster patterns have been made explicit. For example, it is proposed that the finding of higher levels of physical inactivity and sugar consumption amongst women may be due to the competing time demands that disproportionately affect women as a consequence of their increased participation in the labour market in recent decades whilst continuing to undertake domestic chores (Chou et al., 2004; Nomaguchi and Bianchi, 2004). This in turn could be associated with a fall in home-cooked meals and rises in the availability of processed food (POST, 2015, Swinburn et al., 2011), increasingly relied upon due to such restrictions on time (Devine, 2005; Worsley et al., 2012).

Furthermore, the suggestion of a convergence in smoking (HSCIC, 2014b; McCartney et al., 2011) and alcohol consumption (Elliott et al., 2007; Keyes et al., 2011; Meng et al., 2014; Purshouse et al., 2017; Schoon and Parsons, 2003; Slade et al., 2016) amongst men and women over time may be linked to improvements in the economic conditions of women in relation to their educational attainment, participation in the labour force and delays in child birth leading to more permissive attitudes towards women's drinking (Keyes et al., 2011; Slade et al., 2016) and increased opportunities for women to consume alcohol and targeted marketing of alcohol products at women (Bosque-Prous et al., 2015; Keyes et al., 2011).

Together this work suggests a potential complex relationship between gender, cohort and HRB clustering, both in terms of membership and the nature of the HRB cluster patterns. Plausible mechanisms through which gender and cohort may influence HRBs are proposed, relating to historical shifts in HRB and gender relations over time.

2.2.3.4 *Socio-economic position (SEP) in adulthood*

The two literature reviews of studies investigating HRB clustering suggest that HRB clustering is socially patterned, finding SEP to be a consistent predictor of HRB clustering (Meader et al., 2016; Noble et al., 2015). These two literature reviews (Meader et al., 2016; Noble et al., 2015) found a clear cross-sectional relationship between disadvantaged SEP in adulthood and membership of clusters characterised by negative HRBs.

Seventeen studies, out of the twenty-one studies appraised, considered the relationship between HRB cluster membership and SEP (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Bondy and Rehm, 1998; Conry et al., 2011; De Vries et al., 2008; Falkstedt et al., 2016; French et al., 2008; Maibach et al., 1996; Patterson et al., 1994; Poortinga, 2007; Schneider et al., 2009; Schuit et al., 2002; Silva et al., 2013; Slater and Flora, 1991; Tobias et al., 2007; Verger et al., 2009; Vermeulen-Smit et al., 2015). Four studies explored the influence of SEP on the nature of HRB cluster patterns (De Vries et al., 2008; Falkstedt et al., 2016; Laaksonen et al., 2001; Schuit et al., 2002).

Fourteen studies identified advantaged SEP to be associated with membership of clusters characterised by positive HRBs and disadvantaged SEP to be associated with membership of clusters characterised by negative HRBs (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Bondy and Rehm, 1998; Conry et al., 2011; De Vries et al., 2008; Falkstedt et al., 2016; French et al., 2008; Patterson et al., 1994; Poortinga, 2007; Schneider et al., 2009; Schuit et al., 2002; Silva et al., 2013; Slater and Flora, 1991; Vermeulen-Smit et al., 2015). In contrast, Tobias et al. (2007) found the reverse, i.e. a higher prevalence of individuals in disadvantaged circumstances in clusters characterised by positive HRBs and a higher prevalence of individuals in advantaged circumstances in clusters characterised by negative HRBs. Maibach et al. (1996) found no relationship between SEP and HRB cluster membership. Patterson et al. (1994) identified 'smoker lifestyle', 'drinker lifestyle' and 'hedonic lifestyle' clusters, the latter cluster characterised by both smoking and alcohol consumption. The authors (Patterson et al., 1994) noted that the 'drinker lifestyle' was associated with higher family income and the 'smoker lifestyle' was associated with lower family income.

Notably, all of the studies used different measures of SEP. For example, education (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; De Vries et al., 2008; Falkstedt et al., 2016; Laaksonen et al., 2001; Maibach et al., 1996; Patterson et al., 1994; Schneider et al., 2009; Schuit et al., 2002; Silva et al., 2013; Slater and Flora, 1991; van Nieuwenhuijzen et al., 2009; Vermeulen-Smit et al., 2015), occupational grade (Conry et al., 2011; Poortinga, 2007; Schneider et al., 2009; Silva et al., 2013; Verger et al., 2009) and income (Berrigan et al., 2003; Bondy and Rehm, 1998; French et al., 2008; Maibach et al., 1996; Patterson et al., 1994; Schneider et al., 2009; Silva et al., 2013; Slater and Flora, 1991; Verger et al., 2009). These differences in the measurement of SEP make cross-study comparisons difficult and may explain some inconsistencies in the results.

However, despite this variation in the measurement of SEP across the studies, a clear association prevails in the majority of the studies. This consistency provides compelling evidence of a relationship between SEP and HRB cluster membership.

Four of the twenty-one studies tested for SEP differences in the nature of HRB clustered patterns (De Vries et al., 2008; Falkstedt et al., 2016; Laaksonen et al., 2001; Schuit et al., 2002). All four of these studies used education as the only marker of SEP. Relying on this individual variable may be inadequate, given that SEP is considered to be multi-faceted (Pavalko and Caputo, 2013) and includes material, occupational and cultural dimensions (Bartley et al., 1999). One study categorised educational attainment into three groups: high, middle and low (De Vries et al., 2008). The other study (Schuit et al., 2002) used five educational categories. Conversely, Laaksonen et al. (2001) measured years of education, collapsed into three categories. The remaining study incorporated information on the education levels of participants as well as those of their parents in order to capture four educational trajectories: stable high, stable low, upwardly mobile and downwardly mobile (Falkstedt et al., 2016).

Out of the four studies investigating SEP differences in HRB cluster patterns, three found that, whilst education predicted HRB cluster membership, the nature of the HRB clusters did not differ according to educational attainment (De Vries et al., 2008; Falkstedt et al., 2016; Schuit et al., 2002). Conversely, Laaksonen et al. (2001) found that the combination of smoking and physical activity differed according to education in men. Among women, the combinations of smoking and alcohol and smoking and diet differed according to education. However, the authors of the latter study (Laaksonen et al., 2001) do not elaborate on the nature or strength of these associations.

The findings of a clear relationship between HRB cluster membership and SEP in the appraised studies is consistent with research comparing HRBs in two British birth cohorts, born in 1958 and 1970. This research suggests that disadvantaged adulthood SEP increases the likelihood of smoking, poor diet and physical inactivity for both men and women (Schoon and Parsons, 2003).

Together this body of research supports the social determinants of health models (Brunner and Marmot, 2006; Dahlgren and Whitehead, 1991) which suggest that HRBs are influenced by structural factors in adulthood.

2.3 The relationship between health-related behaviour in adulthood and dimensions of socio-economic position in childhood and adulthood

This section of the literature review relates to objective 2 and presents evidence from studies investigating the relationship between HRBs in adulthood and different aspects of social circumstances in childhood and adulthood.

2.3.1 Socio-economic position in childhood and adult health-related behaviours

The two literature reviews suggest that HRB clustering is socially patterned (Meader et al., 2016; Noble et al., 2015). However, to date there is a lack of research investigating the role of social circumstances in childhood and HRB clustering in adulthood. Only one of the twenty-one studies considered how socio-economic position (SEP) early in life shapes HRB clustering in adulthood (Falkstedt et al., 2016). The study was based on a sample of Swedish working-age adults and found that disadvantaged childhood SEP (measured by parental education) and disadvantaged adulthood SEP (measured by participant education) were together predictive of membership of clusters characterised by three or four health-damaging behaviours (Falkstedt et al., 2016). Whilst insightful, this study was subject to a number of methodological limitations outlined in section 2.2.1, measuring HRBs with dichotomised variables, using prevalence odds ratios to detect HRB clustering and including education as the only measure of SEP. Moreover, parental education was based on retrospective accounts from participants as opposed to being measured prospectively which may be subject to recall bias (Cohen et al., 2010). This could lead to an underestimation of the effects of childhood SEP, given the better measurement of adulthood SEP (Power et al., 2005).

Investigating the role of SEP in childhood on HRB clustering is an important area of enquiry. Disadvantaged SEP in childhood has been found to be positively associated with health-damaging HRBs among British adults, including smoking (Andersson and Maralani, 2015; Bann et al., 2016; Clouston et al., 2015; Lacey et al., 2010), heavy alcohol consumption (Bann et al., 2016; Clouston et al., 2015; Maggs et al., 2008), poor diet (Bann et al., 2016; Power and Hertzman, 1997) and physical inactivity (Bann et al., 2016; Juneau et al., 2014; Pereira et al., 2014). Moreover, research undertaken with US twins suggests that the clustering of smoking and alcohol consumption in adulthood can be partially explained by shared environmental factors in childhood (Sudharsanan et al., 2016).

Based on this evidence and upon well-established determinants of health models (Brunner and Marmot, 2006; Dahlgren and Whitehead, 1991) whereby structural factors are considered to determine HRBs, it is possible that childhood SEP could determine adult HRB clustering via a behavioural pathway. This implies a direct effect of childhood SEP on adulthood lifestyles by embedding some HRBs through regular participation (Blue et al., 2014; Fennis et al., 2015) which are carried through into mid-adulthood (James et al., 1997; Parry, 2013; Patrick and Nicklas, 2005; Savage et al., 2007; Schooling and Kuh, 2002) and establishing attitudes and beliefs towards other HRBs they are yet to experience (Cohen et al., 2010).

It is also possible that the effect of childhood SEP on adulthood HRB may occur via a social pathway, based on research which has found childhood SEP strongly influences adulthood SEP amongst participants from two British birth cohorts, born in 1958 and 1970 (Anders and Dorsett, 2017; Breen and Goldthorpe, 2001; Bukodi and Goldthorpe, 2012). This implies an indirect effect, the influence of childhood SEP on adulthood HRB occurring through adulthood SEP, which may reinforce HRBs embedded in childhood (Schooling and Kuh, 2002) and influence HRBs at the same age.

Other empirical evidence also lends support to the existence of both the behavioural and social pathways. Studies have identified that childhood SEP has a direct association with adulthood HRBs, despite a large proportion of this effect occurring indirectly via adulthood SEP (Elhakeem et al., 2015; Kamphuis et al., 2013; Pudrovska and Anishkin, 2013; Schooling and Kuh, 2002; van de Mheen et al., 1998; Watt et al., 2009; Yang et al., 2008). However, other studies only suggest a social pathway, finding that the effect of childhood SEP on HRBs in adulthood is fully explained by adulthood SEP (Kestila et al., 2015; Kvaavik et al., 2012; Paavola et al., 2004).

2.3.2 The relationship between material, cultural and occupational dimensions of socio-economic position and adult health-related behaviours

This thesis conceives there to be three dimensions of SEP – namely, material, occupational and cultural – given the consensus amongst leading scholars (Bartley et al., 1999; Cohen et al., 2010; Duncan and Magnuson, 2012; Pavalko and Caputo, 2013). There is evidence to suggest that these dimensions may differentially shape HRBs in adulthood (Bartley et al., 1999; Sacker et al., 2001; Stait and Calnan, 2016).

Below is an outline of literature suggesting that the effect of material, occupational and cultural dimensions of SEP on HRBs may occur through physical access, psychosocial stress and social group habitus. Physical access to HRBs could be restricted by a lack of material resources, psychosocial stress may be the result of material disadvantage and/or lower grade occupations, and social group habitus potentially shapes HRBs through exposure to particular cultural resources and networks.

2.3.2.1 *Material dimension*

The *material* dimension of SEP may restrict physical access to positive HRBs that are likely to promote health and wellbeing by dictating the purchasing power for those HRBs, such as buying fresh fruit and vegetables (Jones et al., 2014; Kelly et al., 2016; Pechey et al., 2013) and leisure-time physical activity (Beenackers et al., 2012; Chinn et al., 1999; Kelly et al., 2016; Parry, 2013). Disadvantaged circumstances could also trigger individual psychosocial stress, increasing the likelihood of engaging in negative HRBs that are health damaging as a coping method (Krueger and Chang, 2008). For example, smoking (Hoek and Smith, 2016; Lindström et al., 2013; Pampel et al., 2010) and concurrent smoking and heavy alcohol consumption (Twyman et al., 2016). Children in materially disadvantaged households may have greater exposure to adults who partake in negative HRBs to cope with economically challenging circumstances (Cohen et al., 2010; McKenna et al., 2016).

Some pathways between SEP and HRBs, particularly those relating to psychosocial stress, are associated with relative as opposed to absolute material disadvantage (Schnittker and McLeod, 2005). However, absolute material circumstances may also play a role. Partaking in health-damaging HRBs may contribute to increased absolute material deprivation, given evidence that taxes on cigarettes and alcohol have the potential to be regressive by disproportionately affecting those in lower income brackets (Dragone et al., 2015; Garrett et al., 2015; Townsend et al., 1994; Vandenberg and Sharma, 2015; Warner et al., 1995), further reducing their ability to engage in other health-promoting HRBs.

Commonly used indicators to capture the material dimension of SEP are living in social housing (Lacey et al., 2014; Sacker et al., 2001), overcrowding (Elliott and Lawrence, 2014; Lacey et al., 2014; Morgan and Baker, 2006; Sacker et al., 2001; Schoon, 2006), receiving free school meals in childhood (Elliott and Lawrence, 2014), lack of car ownership in adulthood (Morgan and Baker, 2006; Sacker et al., 2001), and receiving state benefits associated with disadvantage (Elliott and Lawrence, 2014; Lacey et al., 2014; Sacker et al., 2001; Schoon, 2006), as well as household income (Duncan and Magnuson, 2012; Elliott and Lawrence, 2014).

Living in social housing is associated with greater material disadvantage compared to other forms of housing tenure (Lupton et al., 2009). Overcrowding is a well-established indicator of material disadvantage (Morgan and Baker, 2006). Car ownership has been identified as highly correlated with income (Phillimore et al., 1994). Historically, free school meals (Gorard, 2012) and state benefits (e.g. unemployment or family income supplementary benefit) (Lacey et al., 2014) have been allocated on the basis of financial hardship. Household income is a material measure which elucidates gradients in society (Lynch et al., 2004). These measures are considered to adequately capture the economic circumstances of households (Duncan and Magnuson, 2012; Elliott and Lawrence, 2014; Lacey et al., 2014; Schoon, 2006).

2.3.2.2 *Occupational dimension*

The *occupational* dimension of SEP closely associates with employment relations that determine job strain (low job control and high job demand) (Bartley et al., 1999; Sacker et al., 2001). Those in lower occupational grades have less control over their working conditions (Gallie et al., 2016). Exposure to job strain tends to be higher in lower occupational grades (Marmot et al., 1991). Empirical research has found that higher job strain is associated with increased odds of adopting multiple negative HRBs (Heikkilä et al., 2012; Lallukka et al., 2008) and that the relationship between job strain and poor health outcomes is explained to some extent by negative HRBs (Brunner et al., 2007; Kivimäki et al., 2013; Nyberg et al., 2015), although this has not been found universally (Magnusson Hanson et al., 2016). In consequence, positive associations between the occupational dimension of SEP and negative HRBs (Brunner et al., 2007; Kivimäki et al., 2013; Lallukka et al., 2008; Nyberg et al., 2015) may be explained by maladaptive coping methods used to manage stress related to insecure employment relationships (Pampel et al., 2010).

The National Statistics Socio-Economic Classification (NS-SEC) was designed to adequately capture occupational grade and employment relations (ONS, 2010a). The stress associated with employment relations is experienced at the individual level (Sacker et al., 2001). Thus, the occupational SEP of the individual is often considered, rather than the highest occupational grade in the household (Sacker et al., 2001). The types of benefits participants receive through their occupation, such as a pension scheme, could also capture employment relations by reflecting more stable employment and better working conditions. For example, pension schemes are more likely to be offered to those in higher occupational grades (Ginn and Arber, 1993).

2.3.2.3 *Cultural dimension*

The *cultural* dimension of SEP, referred to as the norms and cultures of distinct social groups (Bartley, 2016a), could also shape habitual behaviours of individuals in childhood and adulthood. Bourdieu's theory of 'habitus' claims the regular adoption of behaviour could become habitual and intuitive and inform 'dispositions' to participate in particular HRBs (Blue et al., 2014). A collection of dispositions shared by social groups would become the habitus of that group who reproduce and reinforce the habitus from which they came (Cockerham, 2005). Parents with higher levels of education have acquired knowledge and skills that promote HRBs beneficial for health (Burton-Jeangros et al., 2015) and likely adhere to a particular social group habitus of behaviour (Bartley, 2016a).

Whilst education is a multi-faceted measure of SEP (Cockerham, 2005), academic and vocational qualifications are considered a type of 'institutionalised cultural capital' (i.e. official qualifications) (Kamphuis et al., 2015). Empirical research has identified associations between parental education and adulthood smoking (Gagné et al., 2015; Hayward and Gorman, 2004) and physical activity (Pereira et al., 2014). Moreover, receiving an elite education has been found to be associated with a reduced likelihood of smoking and increased physical activity in mid-adulthood (Bann et al., 2016), which may in part be explained by their 'educational' habitus (Georg, 2016). Empirical evidence suggests a strong relationship between education, social circles and HRBs, with research suggesting 10% of the relationship between HRBs and education can be accounted for by the exposure to particular social networks (Cutler and Lleras-Muney, 2010). In addition, parental interest in education could be considered to fall under 'incorporated cultural capital' (i.e. knowledge, skills and beliefs) (Kamphuis et al., 2015). Such cultural beliefs are likely to be transferred from a parent to a child (Abel and Frohlich, 2012; Georg, 2016; Singh-Manoux et al., 2005). For example, the value placed on education by parents is transmitted to their children and predictive of their educational attainment (Bartley, 2004).

Furthermore, education in adulthood dictates social group habitus through its influence on social norms, friendship choices and social interactions (Bann et al., 2016; Bartley, 2016a), contributing to an individual's identity which is strongly linked to lifestyle (Fennis et al., 2015; Geronimus et al., 2016; Kelly et al., 2016; Reid et al., 2010). For example, there are strong links between diet and identity (Hedegaard, 2016). It is suggested that amongst more disadvantaged status groups there is a stronger culture of participating in health-damaging HRBs, such as smoking (Pampel et al., 2010). This is considered to be the product of 'unequal power and privilege' (Geronimus et al., 2016: 109), tied to an identity of non-conformity and toughness (Pampel et al., 2010).

The Cambridge scale captures patterns of social relationships (Prandy, 1999; Rose, 2008). This measure indicates social distance between the participant and individuals in other occupations by measuring their friendship choices (Bartley, 2016a; Sacker et al., 2001), which could be used as an indicator of the cultural dimension of SEP.

2.3.2.4 *The interrelationship between material, occupational and cultural dimensions of socio-economic position*

The three dimensions of SEP mentioned above, material, occupational and cultural, are considered highly correlated with each other (Abel et al., 2011; Cable, 2014; Hayward and Gorman, 2004). For example, whilst education can provide health-promoting knowledge, material resources are also needed in order take advantage of health-promoting opportunities (Herd et al., 2007).

In childhood, material and cultural dimensions intertwine. Educational attainment is strongly related to income (Bukodi and Goldthorpe, 2012), therefore parents with lower educational attainment are more likely to experience financial hardship. Households with fewer material resources are considered to have poorer home learning environments which impact on children's educational attainment (Duncan et al., 2010; Goodman and Gregg, 2010). Parental education is highly related to the educational attainment of their children, partly through higher quality parent-child interactions (Harding et al., 2015). However, the quality of parent-child interactions is likely to be poorer for parents struggling financially (Duncan et al., 2010). In adulthood, dimensions of SEP are inextricably linked. For example, education is strongly predictive of occupational and material circumstances in adulthood (Bukodi and Goldthorpe, 2011; Kamphuis et al., 2015) which can in turn enhance the ability to obtain further qualifications (Abel et al., 2011; Bukodi and Goldthorpe, 2011).

2.3.3 Sensitive periods in the lifecourse and adult health-related behaviours

Pre-adolescence (age 8–11 (Maggs et al., 2008)) may be a ‘sensitive period’ (Cable, 2014; Mishra et al., 2011) for the formation of HRBs in adulthood. Compared to younger peers, children in pre-adolescence are more conscious of their identity (West et al., 2010) and how they differ from others in terms of their social background (Leahy, 1981). During pre-adolescence a superficial perspective on the differences between SEP groups, which is based on differences in possessions or appearance, may be superseded by a deeper understanding of people that are attached to certain SEP groups (Leahy, 1981).

Individuals are likely to settle into their adult roles during mid-adulthood (considered to start in the third decade of life (Backett and Davison, 1995)). At this point in life, resources (e.g. skills and knowledge) that were acquired in childhood are likely to re-emerge, referred to as a ‘pull from the past’ (Fennis et al., 2015: 2). Mid-adulthood may be a period in which there is a refocusing on past experience triggered by a growing awareness of finite time and key events, such as having children (Backett and Davison, 1995).

Mid-adulthood SEP reflects adulthood SEP attainment, which has occurred through the accumulation of resources from earlier in life (Hayward and Gorman, 2004). This is a period in the lifecourse when individuals reach occupational maturity, see the economic returns of previous educational achievement, and occupational status tends to be highest (Bukodi and Goldthorpe, 2011; Herd et al., 2007). Behavioural differentials according to SEP may be particularly pertinent in mid-life, given that disparities in SEP and associated HRBs have been found to be wider in mid-adulthood compared to older ages (Pavalko and Caputo, 2013; Willson et al., 2007). Research also suggests that health outcomes at older ages are associated with SEP in mid-adulthood via an increased prevalence of health-damaging HRBs (heavy alcohol consumption, smoking and physical inactivity) amongst disadvantaged SEP groups (Hessel and Avendano, 2016).

2.3.4 The relationship between pre-adolescent socio-economic position and mid-adulthood health-related behaviours

Pre-adolescent SEP dictates exposure to particular patterns of HRBs, such as diet and physical activity, which are routinely performed (Blue et al., 2014; Fennis et al., 2015) and thus likely to be carried into mid-adulthood (James et al., 1997; Parry, 2013; Patrick and Nicklas, 2005; Savage et al., 2007; Schooling and Kuh, 2002).

Whilst adolescence and young adulthood (i.e. teens and twenties (Backett and Davison, 1995)) may be ages at which individuals begin to engage in health-damaging HRBs, such as heavy alcohol consumption and smoking, these are considered to be normative periods of behavioural experimentation (Fothergill et al., 2009; Schooling and Kuh, 2002) whereby the uptake of health-damaging HRBs increases but then behaviours tend to improve (Dutra et al., 2017; Wiium et al., 2015). Moreover, HRBs at these ages are increasingly influenced by external factors, such as peer groups and popular media, relative to the social circumstances of the individual's family (Vallejo-Torres et al., 2014; West, 1997; Weyers et al., 2010).

With a re-emphasis on past experiences in mid-adulthood, it is plausible that mid-adulthood HRB clustering is more likely to be affected by pre-adolescent SEP. Family SEP could shape children's attitudes and beliefs on the appropriateness of patterns of HRBs they are yet to experience, such as smoking and alcohol consumption, which are later observed in mid-adulthood (Cohen et al., 2010). Children view parents as role models, normalising their HRBs (Cohen et al., 2010; Pampel et al., 2010; Wickrama and Wickrama, 2010). Cultural resources, such as social values and norms, are in part acquired through social learning that varies according to SEP (Abel, 2008). Social learning theory (Bandura, 1991) claims that behaviour is learnt through environmental observations which are then internalised and imitated later in life.

2.3.5 The mediating effect of mid-adulthood socio-economic position

As mentioned above (see section 2.3.3), mid-adulthood SEP reflects adulthood SEP attainment. This attainment is conceived to occur through the accumulation of resources in childhood (Hayward and Gorman, 2004) which strongly predicts SEP and HRBs in adulthood (Cohen et al., 2010; Hayward and Gorman, 2004). Therefore, mid-adulthood SEP may reinforce HRBs embedded earlier in life through regular participation (Blue et al., 2014). The ability to partake in health-promoting, and avoid health-damaging, HRBs will be determined by physical access, psychosocial stress and social group habitus pathways described above. Disadvantaged SEP in childhood has been found to predict disadvantaged SEP in adulthood for men and women in two British cohorts, born in 1958 and 1970 (Anders and Dorsett, 2017; Breen and Goldthorpe, 2001; Bukodi and Goldthorpe, 2012). Below is a description of how material and cultural dimensions of SEP in pre-adolescence may influence material, occupational and cultural dimensions of SEP in mid-adulthood.

2.3.5.1 *The potential pathway between the material dimension of SEP in pre-adolescence and in mid-adulthood*

Direct transmission of material resources from pre-adolescence to mid-adulthood may occur. Advantaged SEP in pre-adolescence implies higher household resources which are considered to largely determine those generated in adulthood (Carvalho, 2012; Hayward and Gorman, 2004). Children from disadvantaged households with fewer resources have less chance of accruing them (Heinz and Marshall, 2003). For example, material resources can accumulate via the intergenerational transfer of income and household wealth, implying that material resources in childhood are likely to be transmitted to material resources in adulthood (Davies and Shorrocks, 2000; De Nardi, 2004).

2.3.5.2 *The potential pathway between the cultural dimension of SEP in pre-adolescence and the material and cultural dimensions of SEP in mid-adulthood*

Familial cultural resources may predict cultural and material resources in adulthood. Both parental education (Harding et al., 2015; Howe et al., 2013a) and parental interest in their child's education (Harding et al., 2015) have been found to be predictive of the educational outcomes of their children. Educational outcomes shape material resources in adulthood through higher incomes (Cutler and Lleras-Muney, 2010). Moreover, educational habitus (Georg, 2016) experienced during in pre-adolescence can shape mid-adulthood social norms, friendship choices and social interactions (Bann et al., 2016; Bartley, 2016a).

2.3.5.3 *The potential pathway between the material dimension of SEP in pre-adolescence and the cultural dimension of SEP in mid-adulthood*

Material resources in early life can determine cultural resources at older ages (Hayward and Gorman, 2004). Material disadvantage in childhood reduces the ability to buy books, or pay for school trips or extra-curricular activities, which may impede learning and educational attainment (Amato, 1994; Horgan, 2007). Financial constraints may dictate the quality of schooling. Despite education being free in the United Kingdom, families who are unable to afford to move to an affluent area will be excluded from catchment areas with better comprehensive schools (Singleton et al., 2011; Smithers and Robinson, 2010). Living in overcrowded accommodation can be disruptive to children's education by reducing their ability to concentrate and do homework (Solari and Mare, 2012). Familial financial worries may also be a barrier to entering or effectively engaging in higher education, preventing their child from obtaining a degree (Greenbank and Hepworth, 2008; Raque-Bogdan and Lucas, 2016).

2.3.5.4 *The potential pathway between material and cultural dimensions of SEP in pre-adolescence and the occupational dimension of SEP in mid-adulthood*

Material and cultural resources can also influence occupational outcomes (Greenbank and Hepworth, 2008; Halleröd and Gustafsson, 2011). From a material perspective, work may be entered at a younger age for those from poorer households in order to earn money (Lynch and O'riordan, 1998). Culturally speaking, research in the United States has found that first-generation college (i.e. undergraduate) students from 'blue collar' backgrounds (i.e. parents without a bachelor's degree) tend to have lower career aspirations than those whose parents went to college (Martinez et al., 2009; Raque-Bogdan and Lucas, 2016). Duncan and Magnuson (2012) suggest that parents who may not have been exposed to higher education and whose occupational experience has been within lower-grade jobs (associated with less autonomy and self-direction than higher occupational grades) will place more emphasis on discipline and conformity as skills their children will need in the future, compared to parents in higher occupational grades who may instead encourage the development of autonomy in their children.

2.4 The stability of health-related behaviours during mid-adulthood

Research investigating HRB clustering has provided useful insights. As mentioned above in section 2.2, the HRB clustering literature suggests that clustered patterns of HRBs are likely to exist, that these HRB patterns are shared by subgroups of the population and that they may be socially patterned. However, to date, research investigating HRBs and their relationship with SEP has tended to assume stability of HRBs over time. This is articulated by Oude Groeniger and van Lenthe (2016):

'The underlying assumption is that health behaviours are fairly stable over time, and that the initial baseline measurement is a good indicator for lifelong exposure' (Oude Groeniger and van Lenthe, 2016: 1).

A similar observation was made by Laaksonen et al. (2002), who investigated the interrelatedness of multiple HRBs over time and suggested that examining HRB stability was worthwhile, given that, whilst some studies had found HRBs to be stable over time, this was not always the case.

On this basis, a review of the literature has been undertaken in order to identify whether there is sufficient evidence to support the commonly held assumption of stability in HRBs during mid-life and whether this assumption should be challenged in regard to HRB clustering.

Below is a review of literature pertaining to change in multiple HRBs over time. Change in HRBs over time may relate to age, period and/or cohort effects (Jiang et al., 2013; Meng et al., 2014; Vedøy, 2014) and the spillover effect (Poortinga et al., 2013) of change in one HRB on another (Prochaska et al., 2008a). As mentioned in section 2.2.3.2, cohort comparisons of participants of the same age control for age and highlight a combination of period and cohort effects (Schoon, 2006). Moreover, using information from one cohort at two different ages controls for cohort effects and indicates the combined effects of period and age (Schoon, 2006).

This review has been separated according to studies investigating HRB change during mid-adulthood (indicating age effects), studies that consider HRB over historical time (indicating period effects), evidence from research utilising birth cohort studies (indicating cohort effects), and research investigating the interrelatedness of change amongst multiple HRBs (indicating spillover effects). The final consideration is of research implicating gender and SEP as predictors of change in HRBs.

2.4.1 Health-related behaviour change during mid-adulthood

To date, research on the clustering of smoking, alcohol, diet and physical activity has been largely cross-sectional with no examinations of HRB clustering longitudinally amongst British men and women in mid-age. Whilst there has been research investigating multiple HRB stability amongst individuals during childhood, adolescence and young adulthood (Childs and Sullivan, 2013; Chung et al., 2005; de Winter et al., 2016; White et al., 2009), there is little evidence focusing on individuals during mid-adulthood.

Four studies have been identified investigating stability of all four HRBs across an individual's lifecourse (Benzies et al., 2008; de Winter et al., 2016; Mulder et al., 1998; Paavola et al., 2004). Two of these studies investigated HRB clustering (Benzies et al., 2008; de Winter et al., 2016), one used a sample of adolescents (de Winter et al., 2016) and the other used a sample of Swedish women in mid-life (Benzies et al., 2008). Whilst all of these studies were different in their design and sample, all four suggested that smoking behaviour had greater continuity over time compared to alcohol, diet and physical activity (Benzies et al., 2008; de Winter et al., 2016; Mulder et al., 1998; Paavola et al., 2004).

This is line with behaviour change theory which suggests smoking is particularly addictive because nicotine acts directly on the brain leading to noticeable physiological changes and thus over time a strong conditioning effect occurs, whereby the body begins to anticipate these physiological changes (Borland, 2013b). Moreover, whilst alcohol could also be considered an addictive behaviour, it has a different temporal rhythm to smoking (Room, 2004). It is common for smoking occasions to take place multiple times throughout the day, whereas, for most people, alcohol consumption tends to occur less frequently (Room, 2004). Moreover, whilst smoking does not tend to act as a cue for alcohol consumption, drinking alcohol is often a cue to smoke (Room, 2004).

Amongst the two studies focusing on mid-aged adults (Benzies et al., 2008; Mulder et al., 1998), both found alcohol to be more stable than diet and physical activity. The first study employed a variable-centred data reduction technique (i.e. principal components analysis) to investigate the interrelatedness of multiple HRBs during mid-adulthood amongst women (Benzies et al., 2008). The authors found two distinct constructs subsequently labelled 'healthy eating' and 'addictions'; these correlated weakly with each other and most women remained stable within each dimension (pearson correlation between time 1 and time 2, 'healthy eating'=0.7, 'addictions'=0.8). Change that did occur within each dimension tended to be positive (i.e. increased consumption of fruit and vegetables and reductions in smoking and alcohol consumption). The second study (Mulder et al., 1998) focused on a sample of middle-aged men and found that HRBs were relatively stable (35% physical activity, 50% diet, 65% alcohol, 71% smoking). The study found change in a positive direction (30% physical activity, 26% diet, 21% alcohol, 18% smoking) and negative direction (35% physical activity, diet 25%, 15% alcohol, 11% smoking), with 11% of the sample changing two or more HRBs over a four-year period. HRB change was related to age. Participants in the two older age groups (age range 50–70) were more likely to reduce or stop smoking, have healthier diets and reduce their alcohol consumption in comparison to the younger age groups (age range 30–49), indicating positive HRB change. However, physical activity amongst participants in the two older age groups was more likely to decline than among the younger participants.

Although the sample in the first study (Benzies et al., 2008) was considered to be representative of Swedish women in mid-life, the sample could be considered small (N=569). Therefore, the extent to which the results can be generalised to other populations is poor. Similarly, whilst the results of the second study investigating HRB stability in mid-life (Mulder et al., 1998) are insightful, comparability is limited, given that this research was based on a sample of Dutch men in mid-life located in two northern provinces of the Netherlands.

Findings from studies investigating stability of behavioural patterns for multiple HRBs in mid-adulthood (Benzies et al., 2008; Mulder et al., 1998) is supported by empirical research investigating individual HRBs, which suggests that during mid-age there is a degree of stability in smoking (Fidler et al., 2013; Mathew et al., 2016), alcohol consumption (Britton et al., 2010; Johnstone et al., 1996; Kaplan et al., 2012), diet (Harrington et al., 2014; Parsons et al., 2006) and physical activity (Mertens et al., 2016; Pereira et al., 2014). From a theoretical perspective, it is possible that little change occurs in HRBs during this period of the lifecourse, compared to the experimental and more turbulent periods of adolescence and young adulthood (Fothergill et al., 2009; Schooling and Kuh, 2002). Mid-adulthood is generally considered a period of increased regularity and routine (Backett and Davison, 1995). For example, it is a time when child rearing and employment are most common (Lachman et al., 2015).

However, the studies also suggest that some change in multiple HRBs during mid-adulthood may occur (Benzies et al., 2008; Mulder et al., 1998). Although cross-sectional research investigating HRB clustering according to age (outlined fully in section 2.2.3.1) cannot detect stability, it does provide some useful insights into the relationship between HRB clustering and age, suggesting a higher membership of older participants in clusters characterised by health-promoting HRBs compared to younger participants who are more likely to be members of clusters characterised by health-damaging HRBs (Berrigan et al., 2003; Conry et al., 2011; Patterson et al., 1994; Poortinga, 2007; Schneider et al., 2009; Tseng and Lin, 2008; Verger et al., 2009; Vermeulen-Smit et al., 2015). This suggests that, as people age, their HRB cluster patterns tend to improve.

Other empirical evidence lends itself to the hypothesis that during middle age individuals tend to improve their diets (Artaud et al., 2016; Backett and Davison, 1995; Benzies et al., 2008; Mulder et al., 1998; Parsons et al., 2006; Sijtsma et al., 2012), reduce or abstain from smoking (Artaud et al., 2016; Backett and Davison, 1995; Mulder et al., 1998; Paffenbarger Jr et al., 1993) decrease alcohol consumption (Backett and Davison, 1995; Benzies et al., 2008; Britton et al., 2010; Meng et al., 2014; Molander et al., 2010; Mulder et al., 1998) and increase their levels of physical activity (Backett and Davison, 1995; Lissner et al., 1996). As mentioned in section 2.3.3, mid-adulthood is associated with a growing awareness of finite time and key events, such as having children (Backett and Davison, 1995). Thus, as people age, their future health becomes more salient and motivation to adopt health-promoting HRBs and stop health-damaging HRBs may increase, either to prevent or respond to chronic illness (Backett and Davison, 1995).

On the other hand, it may also be the case that HRBs worsen during mid-life. This period in the lifecourse is associated with increased responsibilities (Lachman et al., 2015). Such responsibilities may restrict the ability to partake in health-promoting HRBs (Chou et al., 2004; Cutler et al., 2003; Lachman and James, 1997; Worsley et al., 2012) by depleting both the motivation and time required to actively participate in positive HRBs (Borland, 2013a; Kelly et al., 2016). This may explain observed declines in physical activity during mid-adulthood (Allender et al., 2008b; Artaud et al., 2016; Corder et al., 2009; Meyer et al., 2016; Mulder et al., 1998; Wannamethee et al., 1998) and research suggesting that activity levels may increase again after retirement (Feng et al., 2016; Menai et al., 2014). Competing time demands in mid-life are also linked to poorer diet, due to an increased consumption of convenience, rather than home-cooked, meals and snacks which are often high in sugar and fat (Chou et al., 2004; Cutler et al., 2003; Worsley et al., 2012). Moreover, changing negative HRBs is challenging and therefore depleted time and energy levels due to such responsibilities are likely to impinge on the ability to refrain from negative HRBs (Borland, 2013b). This may to some extent explain the relative stability of smoking and alcohol consumption during this period of life (Benzies et al., 2008; Mulder et al., 1998).

Finally, it should also be noted that some research has elucidated complexity in HRB stability for alcohol and physical activity. For example, Kerr et al. (2002) identified that stability in alcohol consumption differs depending on the level of alcohol consumed, being more stable for abstainers and moderate drinkers compared to heavy drinkers. In regard to physical activity, Lissner et al. (1996) found in a sample of middle-aged women that similar proportions increased (19.8%) and decreased (20.9%) their physical activity over a six-year period. Due to such inconsistencies and a lack of evidence concerning the stability of HRB clustering during mid-life, it appears that the assumption of absolute stability in multiple HRBs over time may not be valid (Laaksonen et al., 2002; Oude Groeniger and van Lenthe, 2016) and that there is a need for further investigations into the stability of HRB clustering during mid-life.

2.4.2 Health-related behaviour change over historical time

General trends over historical time have been observed at a population level for these four HRBs in the United Kingdom and other developed countries. Since the 1950s there have been declines in smoking (RCP, 2012) and since the early 1990s there have been increases in the consumption of fruit and vegetables (Ezzati et al., 2015) and leisure-time physical activity (An et al., 2016; Stamatakis et al., 2007). At the same time, there have been increases in sugar consumption (Chang et al., 2017; Singh et al., 2015; WHO, 2015) and alcohol consumption (Meng et al., 2014).

Research investigating multiple HRBs has observed a decline in the prevalence of individuals reporting to have all four health-damaging HRBs (i.e. smoking, heavy alcohol consumption, low consumption of fruit and vegetables, physical inactivity) in the United Kingdom (Alageel et al., 2016; Buck and Frosini, 2012) and other developed countries (DeRuiter et al., 2014; Ding et al., 2015). Two studies investigating change in all four HRBs during the 1990s and early 2000s (DeRuiter et al., 2014; Ding et al., 2015) identified that the most change occurred for those who were physically inactive smokers. In contrast, research conducted in the United Kingdom (Buck and Frosini, 2012), studying change between 2003 and 2008, found that the biggest decrease was for individuals with co-occurring heavy alcohol consumption, low fruit and vegetable consumption and low levels of physical activity.[6] This is supported by research investigating change between 2003 and 2013 in the United Kingdom finding decreases in cigarette smoking, low physical activity, low fruit and vegetable intake and excessive alcohol consumption (Alageel et al., 2016). In contrast, research conducted in the United States found that between 1996 and 2007 the proportion who reported to have health-promoting HRBs in relation to smoking, diet and physical activity reduced (Ford et al., 2010).

Furthermore, research detecting different drinking typologies in a UK population sample observed that those characterised by heavier alcohol consumption had higher than average rates of smoking in comparison to the population average, but the prevalence of smoking in all of the typologies reduced between 1978 and 2010 (Purshouse et al., 2017).

2.4.3 Cohort and health-related behaviour change

The above literature suggests that both circumstances unique to those in mid-age as well as wider population trends over historical time may influence HRB stability during mid-life. A review of research using prospectively collected data from British birth cohorts, controls for potential cohort differences and can therefore be useful in further elucidating the effects of historical time and/or age (Schoon, 2006), although the latter two cannot be distinguished from one another.

[6] Clustering was not investigated by Buck & Frosini (2012). See chapter 1, section 2, for an outline of the two approaches.

Cross-comparisons of birth cohorts in terms of smoking behaviour (Kemm, 2001; Schulze and Mons, 2005) have identified that overall smoking prevalence has decreased over successive cohorts born after the 1920s, although women's smoking prevalence increased between the 1920s and 1960s before decreasing. These studies also found the age at which smokers quit has fallen across subsequent cohorts over time.

Parsons et al. (2006) considered the stability of diet and physical activity using the National Child Development Study, a cohort of British people born in 1958, at ages 33 and 42. They used Spearman correlation coefficients to assess the continuity of these behaviours and determined that it was low to moderate for diet (ranging from 0.23 to 0.49) and physical activity (0.31 men; 0.23 women). During this nine-year period, fruit and vegetable consumption increased, sweet food consumption decreased, fried food consumption increased and there was little change in levels of physical activity.

Another study (Pereira et al., 2014) used the same cohort to examine physical activity over a longer time period (ages 33, 42 and 50) and found continuity, stating that 52% of those who were inactive at age 33 were also inactive at age 42, and 48% of those inactive at age 42 were also inactive at age 50.

Mishra et al. (2006) investigated changes in diet and alcohol consumption over a ten-year period amongst an earlier cohort of middle-aged British adults who were born in 1946. Similarly to Parsons et al. (2006), the authors found marked changes in diet during mid-age. More specifically, they identified a decline in meat, potatoes and sweet foods and increases in fruit and vegetable consumption. The authors also identified increases in alcohol consumption. The latter finding is supported by research which found increases in alcohol consumption in mid-age adults from the 1970 British Birth Cohort Study (Elliott et al., 2007). This increase in alcohol consumption among cohorts of mid-age adults has also been identified elsewhere (Britton et al., 2015; Ilomäki et al., 2010; Kraus et al., 2015; Meng et al., 2014; Molander et al., 2010).

2.4.4 Interrelatedness of change of health-related behaviours

As mentioned in section 2.2, HRBs are considered to be mutually dependent (Blue et al., 2014). Therefore, change in one HRB may lead to change in another (Prochaska et al., 2008a), through behavioural spillover effects (Poortinga et al., 2013)

Empirical evidence suggests that change in smoking is highly related to changes in diet, alcohol and physical activity (Brown et al., 2016; DeRuiter et al., 2014; Laaksonen et al., 2002; Noonan et al., 2016; Prättälä et al., 1998; Tian et al., 2016; Unger, 1996; Yusufov et al., 2016). For example, Laaksonen et al. (2002) found that changes in diet, alcohol and physical activity in adults over a 7-year period involved changes in smoking behaviour. In a population sample of adults in England who were followed for one week, smokers who commenced a quit attempt consumed less alcohol within the same period (Brown et al., 2016). DeRuiter et al. (2014) used latent variable modelling to identify change in HRBs over a 12-year period in a representative population sample of working-age adults and found that increases in physical activity and decreases in alcohol consumption were both associated with lower tobacco use but not with each other. Prättälä et al. (1998) found that, in adults over a 15-year period, the consumption of vegetables increased more for non-smokers and ex-smokers than for smokers and Yusufov et al. (2016) found that positive changes in smoking were related to positive changes in diet.

Moreover, in a sample of young women, heavy alcohol consumption predicted the adoption of smoking (McDermott et al., 2009). Heavy alcohol consumption was predictive of persistent smoking and the re-uptake of smoking after quitting (McDermott et al., 2009). The authors also found being physically active was associated with remaining an ex-smoker (McDermott et al., 2009). A literature review of studies conducted in human laboratories concludes that nicotine and alcohol have a reciprocal relationship, with each substance influencing the craving and use of the other (Verplaetse and McKee, 2017).

The relationship between changes in physical activity, diet and alcohol consumption is less consistent. DeRuiter et al. (2014) found no relationship between changes in physical activity and alcohol consumption. In contrast, Conroy et al. (2015) identified that on days when individuals are physically active they tend to drink more alcohol. Furthermore, increased physical activity has been identified as an effective strategy in reducing alcohol consumption amongst those with problematic alcohol use (Stoutenberg et al., 2016). In terms of diet and alcohol use, individuals with an unhealthy diet at baseline have been found to drink less alcohol 7 years later (Laaksonen et al., 2002). Some studies have identified that increases in physical activity are associated with an improved diet (Fleig et al., 2015; Parsons et al., 2006). In contrast, another study found no association suggesting that diet and physical activity may not be interrelated over time (Woolcott et al., 2013).

2.4.5 Gender and health-related behaviour change

There is evidence to suggest that change in HRB clustering may vary according to gender. Research investigating change in multiple HRBs over time in the United Kingdom found that the proportion of individuals participating in all four health-damaging HRBs (i.e. smoking, heavy alcohol consumption, low fruit and vegetable intake, physical inactivity) reduced more markedly for women compared to men between 2002 and 2008 (Buck and Frosini, 2012). However, the study also identified significant increases in the proportion of women with low fruit and vegetable consumption and heavy alcohol consumption compared to men in the same time period.

The above is indicative of other work that suggests that there has been a convergence in smoking (HSCIC, 2014b; McCartney et al., 2011) and alcohol consumption amongst men and women over time (Elliott et al., 2007; Keyes et al., 2011; Meng et al., 2014; Purshouse et al., 2017; Schoon and Parsons, 2003; Slade et al., 2016).

This convergence is considered to be largely due to an increase in the uptake of smoking (McCartney et al., 2011; Schooling and Kuh, 2002) and heavier alcohol consumption (McCartney et al., 2011; Purshouse et al., 2017) amongst women rather than marked reductions in these health-damaging HRBs amongst men. Research suggests that whilst the uptake of smoking amongst women occurred later in historical time compared to men the decline in smoking has also been slower (McCartney et al., 2011; Schooling and Kuh, 2002). Moreover, whilst alcohol consumption has increased for both men and women in subsequent cohorts born before 1980 (Kerr et al., 2009; Kraus et al., 2015; Meng et al., 2014), amongst drinkers born after 1980 there is evidence to suggest that consumption levels have declined for men but have continued to increase for women (Meng et al., 2014). This suggests a complex and dynamic relationship between smoking, alcohol consumption and cultural constructions of gender over time (McCartney et al., 2011).

Gender differences in the dietary patterns have also been identified in a cohort of British middle-aged adults born in 1946 (Mishra et al., 2006) whereby changes in dietary patterns over a ten-year period differed for men and women. Most notably, sweet food consumption reduced for women but not men (Mishra et al., 2006). Prättälä et al. (1998) identified changes over time in the association between dietary intake and never, ex and current smoking status with never and ex-smokers consuming more vegetables compared to current smokers. The authors also found that amongst men differences in consumption of vegetables between ex-smokers and current smokers increased over time whereas for women these differences increased and then decreased.

Whilst body mass index (BMI) is not considered to be an HRB it is often used as a proxy for diet (Campostrini and McQueen, 2003). Research investigating the relationship between smoking, alcohol and obesity (BMI≥30) has identified gender differences (Lahti-Koski et al., 2002). This research found alcohol abstinent women and heavy drinking men were more likely to be obese and that male smokers were more likely to be obese than never smokers but this was not the case for women.

Parsons et al. (2006) found in a cohort of British adults born in 1958 that there were significant decreases in physical activity levels for women between the ages of 33 and 42 ($p<0.01$), which were not found for men ($p>0.05$). In a cohort of British middle-aged adults born in 1946, longitudinal patterns of physical activity were found to differ for men and women (Silverwood et al., 2011). The study results suggested more women belonged to patterns characterised by higher levels of physical activity than men. Furthermore, changes in physical inactivity and unhealthy diet have been found to be associated with one another in men but not women (Laaksonen et al., 2002).

2.4.6 Socio-economic position and health-related behaviour change

Scholars suggest that the action required to change specific HRBs may differ according to SEP (Borland, 2013c). As mentioned above, whilst some HRBs will require active participation (e.g. diet and physical activity) and others will require restraint (e.g. smoking and alcohol) (Borland, 2013e), in both cases changing health habits is considered challenging and increased access to particular resources is considered to make HRB change easier (Borland, 2013c; Dixon and Banwell, 2009).

The influence of SEP on HRB change highlights how individual agency to change lifestyle is heavily determined by structural factors (Cockerham, 2005). As mentioned in section 2.3.2, it is theorised that material and occupational dimensions of SEP are linked to psychosocial stress (Pampel et al., 2010; Sacker et al., 2001). Stress is considered a major contributor to the persistence of health-damaging HRBs, such as smoking and heavy alcohol, and a lack of participation in health-promoting HRBs, such as a healthy diet and physical activity (Borland, 2013c). Increased exposure to stress reduces motivation and capacity for self-restraint (Borland, 2013c). Moreover, scholars argue that factors such as psychosocial stress are considered to operate at multiple levels (i.e. macro, meso and micro) and hence have a strong proximal influence on the context in which people live and thus on their HRBs (Short and Mollborn, 2015).

Behaviour change theory states that changing HRBs through reduction or restraint requires the introduction of a substitute of some kind (Borland, 2013a). This is consistent with Pampel et al.

(2010) who suggest that increased access to material resources gives smokers attempting to stop the opportunity to choose alternative pleasurable activities. This is supported by the findings of a systematic review of literature investigating the barriers and facilitators of HRBs in mid-life, which found that financial cost associated with health-promoting HRBs was a barrier to their uptake and continuation (Kelly et al., 2016). Moreover, this systematic review found that having one health-promoting HRB encouraged the uptake and maintenance of others.

Furthermore, Chinn et al. (1999) identified that adults in the United Kingdom occupying more advantaged social positions cited ‘internal’ barriers to physical activity, such as time and motivation, whereas those occupying more disadvantaged social positions cited ‘external’ barriers to physical activity, such as lack of economic resources and transport, implying a possible double burden (i.e. internal and external) of disadvantaged SEP on HRB change in mid-life. Pampel et al. (2010) theorised that health-damaging HRBs may be used as a way to cope with material hardship and therefore those with lower incomes may be managing this additional burden when it comes to cessation. These mechanisms may help to explain why HRB change interventions appear to be less effective amongst those with a low income (Bull et al., 2014).

The decision to sustain or change a particular health-damaging HRB may differ by SEP. Some suggest that higher educational attainment will not necessarily increase exposure to health messages but may lead individuals to respond to them due to more efficient information processing (Cutler and Lleras-Muney, 2010; Schooling and Kuh, 2002). This may, to some extent, explain the finding that individuals from more disadvantaged SEP are less likely to reduce their alcohol consumption as a health precaution (Britton and Bell, 2015). Others have found that beliefs on the relationship between diet and health were weaker amongst individuals with lower levels of education (Miura and Turrell, 2014).

Employment relations have also been found to influence change in HRBs. One study found that amongst physically active individuals, those who had higher levels of job strain at baseline were more likely to become physically inactive over time compared to those with lower job strain (Fransson et al., 2012). Experimental research suggests that different levels of nicotine dependence are associated with different levels of biological responses to stress (Morris et al., 2016), thus if nicotine dependence were to increase over time the stress reduction associated with smoking may diminish.

Self-identity and cultural norms are implicated in the uptake and maintenance of health-promoting HRBs (Kelly et al., 2016). Fennis et al. (2015) postulate that HRBs are a crucial element of personal identity and that these are often ingrained and automatic. Higher earnings may allow access to new or luxurious HRBs, which may be health promoting or damaging, changing tastes and preferences of the collective who can afford them (Cutler and Lleras-Muney, 2010). Moreover, whilst individuals in more advantaged SEP groups may be the first to adopt health-promoting HRBs in response to health messages these can also become ingrained as part of a global lifestyle and social group identity (Bartley, 2016a; Pampel et al., 2010), which individuals subsequently strive to maintain (Fennis et al., 2015). It is theorised that over time HRBs synonymous with the most advantaged SEP group are imitated by other SEP groups, known as 'diffusion' (Dixon and Banwell, 2009). This may help to explain historical trends in women's smoking behaviour, with advantaged SEP women being the first to start and the first to stop (Schooling and Kuh, 2002). It may also explain the uptake of HRBs appear to match those of the social group to which an individual attains in adulthood as opposed to the social group from which they were part of in childhood (Borodulin et al., 2012; Karvonen et al., 1999).

However, the inclination to hold onto self-identity (Fennis et al., 2015) can also lead to the maintenance of health-damaging HRBs. Pampel et al. (2010) suggests that health-damaging HRBs, such as smoking, are tied up with a collective identity amongst individuals in more disadvantaged social circumstances of toughness and non-conformity. This may contribute to explaining why, despite being aware of the harms of smoking (Blaxter, 1990), a higher proportion of individuals within more disadvantaged SEP groups smoke in comparison to the rest of the population (Jefferis et al., 2004).

Whilst declines in smoking have been observed over time at a population level (RCP, 2012), this has not been uniform across SEP groups leading to an increase in SEP inequality in smoking across subsequent cohorts (Jefferis et al., 2004; Martikainen et al., 2013; Vedøy, 2014). Research using data from a cohort of British adults born in 1958 found that between age 23 and 41 participants from more disadvantaged SEP backgrounds (captured via four categories of occupational social class) were not only more likely to smoke but less likely to quit than their more advantaged peers (Jefferis et al., 2004). In a sample of US adults aged 50–75 the likelihood of both quitting smoking and sustaining cessation was greater for those with higher educational attainment (Margolis, 2013). Similarly, in a sample of young women lower educational attainment was associated with continued smoking over a ten-year period

(McDermott et al., 2009). Using a sample of Swedish middle-aged women, Benzies et al. (2008) found that higher levels of education predicted change in smoking during a four-year period. Moreover, a study of Italian adults found that successful smoking cessation was related to higher levels of education and the absence of financial difficulties (Coppo et al., 2017).

However, Vedøy (2014) did not find evidence that advantaged SEP was related to an increase in smoking cessation, instead advantaged SEP was associated with a lower likelihood of smoking initiation. Laaksonen et al. (2008) found a complex interplay between gender, SEP and smoking. The investigators identified that among men there was little influence of educational differences on smoking initiation or cessation whereas in less educated women there was an increase in smoking initiation, not observed amongst more educated women.

In a cohort of British middle-aged British adults born in 1946, increases over a ten-year period in fruit and vegetable consumption were more likely for participants with advantaged SEP (higher education and non-manual class) (Mishra et al., 2006). Interestingly, both this study and research conducted on a cohort of British civil servants (Britton et al., 2010) found that participants with advantaged SEP were more likely to increase their alcohol consumption compared to those with disadvantaged SEP.

Trends suggest an increase in leisure-time physical activity over time but a reduction in domestic and occupational physical activity due to technological advances (Stamatakis et al., 2007). This has implications for SEP differences in physical activity over time given that historically lower grade occupations involved higher levels of activity due to the manual nature of the work (Allender et al., 2008a) and domestic activities are associated with manual occupational grades (Stamatakis et al., 2007). Moreover, empirical research has found that moving from lower to higher occupational grades over time is associated with higher levels of physical activity (Borodulin et al., 2012) and that increases in leisure-time physical activity over time are more likely among higher occupational grades and decreases are more likely amongst lower occupational grades (Seiluri et al., 2011). In a sample of US adults aged 50–75, whilst educational differences in commencing physical activity were explained by health status the likelihood of sustaining physical activity over time was not (Margolis, 2013). Instead the study physical activity found to be greater for those with higher educational attainment (Margolis, 2013).

There is also evidence to suggest that decreases in the proportion of individuals with multiple health-damaging HRBs differs according to SEP (Buck and Frosini, 2012; Ding et al., 2015). In a sample of British adults between 2003 and 2008, participants with disadvantaged SEP were found to be less likely to improve their HRBs over time (Buck and Frosini, 2012). The investigators found that an observed decrease in participants with multiple health-damaging HRBs over time was disproportionately represented by those with advantaged SEP. The same relationship was observed for Australian adults between 2002 and 2012 (Ding et al., 2015). Prättälä et al. (1998) found significant associations between changes in dietary intake and smoking status (current, ex, never) over time were largely explained by educational level, whereas, Mulder et al. (1998) found in a sample of middle-aged men from the Netherlands that SEP did not influence change in multiple HRBs over a four-year period.

To date and to my knowledge, there appears to be no research investigating the influence of SEP on stability in HRB clustering during mid-adulthood.

2.5 Summary of the literature review

An appraisal of the literature (outlined in Table 2.1) investigating HRB clustering, suggested that smoking, alcohol consumption, diet and physical activity, do not co-occur by chance and therefore cluster. HRB cluster patterns appear complex consisting of both positive and negative HRBs. Age, gender and SEP are implicated in the formation of HRB clusters, although gender and SEP are considered to be more consistent predictors of HRB clustering in comparison to age (Meader et al., 2016; Noble et al., 2015). A complex relationship between gender and cohort in the formation of HRB clusters may exist, both factors have the potential to predict HRB cluster membership as well as differences in the nature of the HRB cluster patterns.

The literature reviewed suggests material, occupational and cultural pathways through which HRBs in mid-adulthood are influenced by social circumstances during pre-adolescence and mid-adulthood. Moreover, this review indicates that whilst HRB clustering appears relatively stable during mid-adulthood it does have the potential to change and that social circumstances may influence change in HRBs over time through differential access to material, occupational and cultural resources.

The appraisal uncovered inconsistencies across the studies in terms of the HRB cluster patterns and the relationship between age, gender, SEP and HRB clustering. These inconsistencies are likely due to differences in methodology, thus making comparability across the studies poor. Consequently, these methodological differences have been considered in this thesis.

The first methodological consideration is the measurement of HRBs, where dichotomising HRBs should be avoided as this may reduce comparability across studies and leads to a loss of rich information. Secondly, advanced person-centred data reduction techniques, such as latent class analysis, appear to be preferable in comparison to using prevalence odds ratios or cluster analysis to identify HRB clustering. In comparison to cluster analysis, latent class analysis has stronger theoretical underpinnings and statistical procedures to guide model selection and has the ability to incorporate measurement error. Thirdly, large representative samples are required in order to provide adequate statistical power to elucidate HRB patterns and detect differences according to gender, cohort and SEP. Fourthly, the review found that studies tend to use an inadequate number of SEP indicators, which may not fully capture the different dimensions of SEP (i.e. material, occupational and cultural). These measures are often included as observed variables which are prone to measurement error. Using advanced data reduction techniques, such as factor analysis, may circumvent this by detecting a latent variable which can incorporate information from multiple SEP variables and takes into account measurement error. Finally, the only study focusing on SEP in childhood and HRB clustering used retrospective information on SEP in childhood. Ideally, prospectively collected data should be used in order to establish reliable effects between social circumstances earlier in life and HRBs in mid-adulthood.

2.6 Gaps in knowledge

This literature review has highlighted some important gaps in the evidence base. Firstly, there is a lack of research examining the clustering of HRBs across cohorts and only one study, conducted in Sweden, has been found that considers the influence of SEP in childhood on HRB clustering in mid-adulthood. There is a clear need for further research on the effects of childhood SEP on HRB clustering within a British context.

Secondly, there are very few studies investigating multiple HRBs and HRB clustering longitudinally and to my knowledge none have focused on British men and women in mid-adulthood. It is also apparent that most studies investigating multiple HRBs in mid-adulthood assume that they are stable during this period, only measuring them at one time point. Notably, it appears that no studies have considered the extent to which SEP in mid-adulthood influence change in HRB clustering during mid-life.

In light of the above, it is considered that this doctoral project makes novel and valid contributions to existing evidence in three important ways: 1) in elucidating distinct patterns of HRBs shared by subgroups from two British cohorts born 12 years apart; 2) in identifying the role of pre-adolescent SEP in predicting HRB cluster membership in mid-adulthood; 3) in demonstrating the stability of HRB cluster membership during mid-life.

3 Chapter 3: Conceptual model

Based on the reviewed literature (outlined fully in chapter 2), a conceptual model has been developed for this thesis and can be viewed as a diagram in Figure 3.1**.**

HRB clustering

This model conceptualises four HRBs: smoking, alcohol, diet and physical activity, as clustering together and being interrelated. The model therefore incorporates multiple HRBs. HRB clusters may consist of those that are negative and considered to be health damaging (e.g. smoking, heavy alcohol consumption, a diet high in sugar and fat and low in fruit and vegetables and physical inactivity), positive and considered to be health promoting (e.g. not smoking, light or moderate alcohol consumption, a diet low in sugar and fat and high in fruit and vegetables and regular physical activity) or a mixture of both. These HRB clusters are considered to represent health lifestyles which are shared by subgroups of mid-age adults in the British population. The HRB clusters provide self-identity and meaning and reflect the consumption patterns of particular social groups. Therefore, these health lifestyles hold symbolic value and reflect the social context of subgroups who share them, which is both place and time specific. It is possible that both gender and cohort will influence the formation of HRB clusters and are thus considered as effect modifiers. The same HRB cluster patterns may emerge for men and women in both cohorts and gender/cohort differences are only apparent in terms of cluster membership (e.g. more women and those in the later-born cohort belong to clusters characterised by multiple positive HRBs). At the same time, it may be that the nature of the HRB cluster patterns differs according to cohort and gender. For example, given gender convergence for smoking and alcohol consumption over time, gender differences in HRB cluster patterns may be less apparent in the later-born cohort.

Change in HRB clustering over time

The model focuses on change in HRB clustering during mid-adulthood. It is considered likely that HRB clustering during in mid-life will be relatively stable because this period in the lifecourse is characterised by increased regularity and routine. For example, this is a time when individuals tend to settle into mid-adult roles, such as raising children and employment. At the same time, it is plausible that some change in HRB clustering will occur during this period of the lifecourse. For example, as individuals gain more responsibilities during mid-adulthood accompanied by an increased awareness of finite time and future health, their motivation to adopt positive HRBs (e.g. eating more fruit and vegetables and partaking in physical activity) and stop negative HRBs (e.g. smoking and heavy alcohol consumption) may increase.

Consequently, their lifestyles change for the better. Conversely, lifestyles may worsen during mid-life. For example, the increased responsibilities associated with mid-adulthood may lead to competing demands depleting the time and motivation required to participate in positive HRBs (e.g. leisure-time physical activity and preparing healthy meals) and stop negative HRBs (e.g. smoking and alcohol consumption).

Moreover, the place and time specific context in which HRB clusters are embedded is considered to dictate the extent to which particular HRBs continue or cease to exist. General trends observed in the United Kingdom in the latter half of the 20thcentury are likely to play a role, for example declines in smoking prevalence. This model also perceives HRBs as being interrelated and therefore it is considered likely that change in one HRB, may lead to change in others.

Therefore, HRB clusters are considered in this model to be dynamic and may change as a consequence of factors associated with year of birth (cohort) and age (mid-adulthood) as well as place and period (1990s and early 2000s in the United Kingdom).

<u>The relationship between HRB clustering and material, cultural and occupational dimensions of socio-economic position at different stages in the lifecourse</u>

The model incorporates a social determinants and lifecourse framework by considering the role of social circumstances, captured by individual and household SEP, during pre-adolescence and mid-adulthood in shaping mid-adulthood HRBs. It is conceived possible that SEP experienced during pre-adolescence may directly influence HRBs in mid-adulthood through a behavioural pathway. Pre-adolescence SEP is considered particularly salient because it is a period in which children become more aware of their own identity and how they differ from others in terms of their social background.

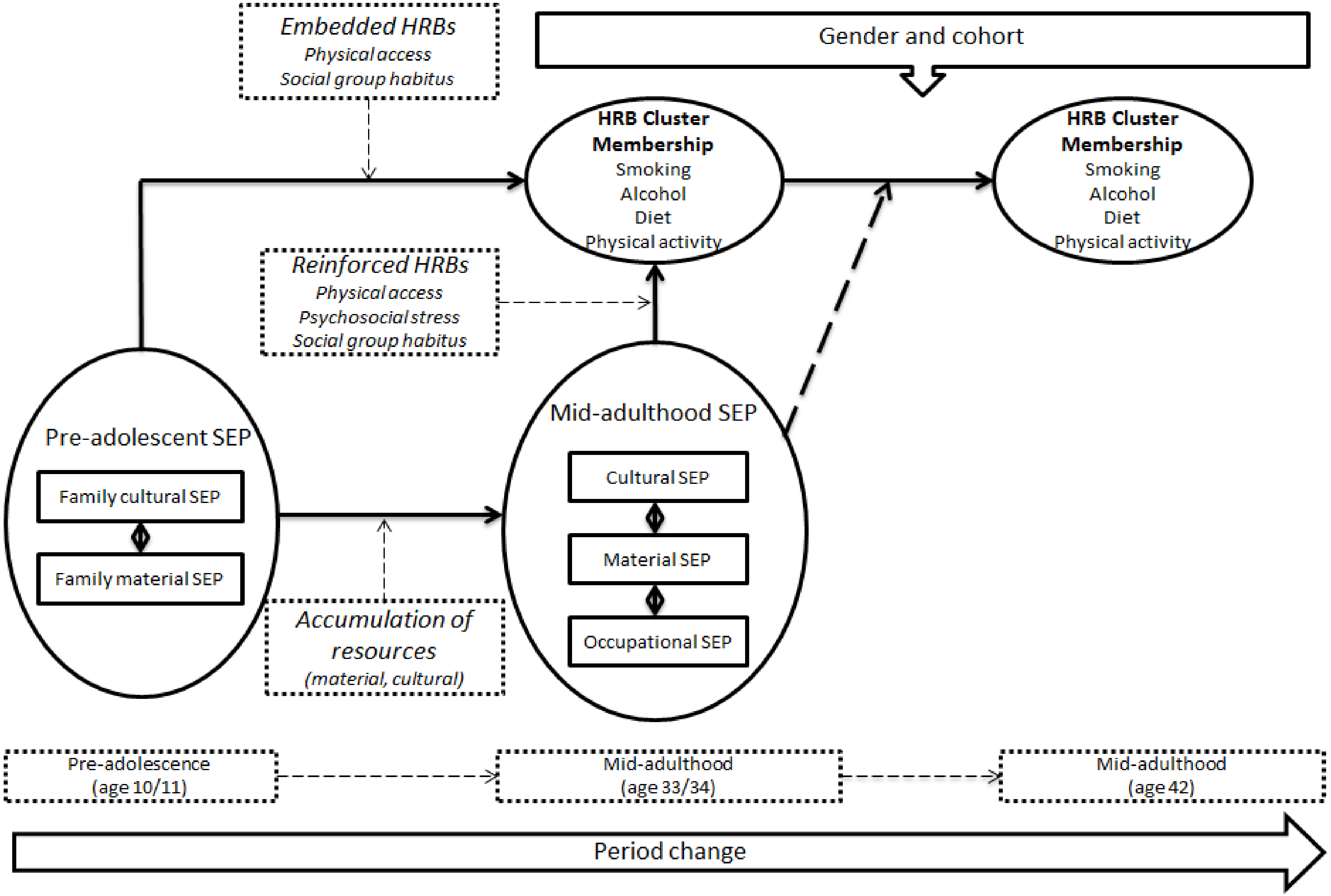
Pre-adolescent SEP will embed some HRBs through regular participation and establishing attitudes and beliefs towards other HRBs they are yet to experience. Exposure to particular HRBs will vary according to pre-adolescent SEP via differentials in the social group habitus of the family (determined by cultural resources) and physical access (determined by material resources). These material and cultural dimensions of SEP in pre-adolescence are considered to be highly interrelated. The pull of pre-adolescent SEP will re-emerge in mid-adulthood following the normative experimental stages of adolescence and young adulthood who are more heavily influenced by external factors compared to household SEP and when individuals are considered to have settled into their mid-adulthood roles.

It is also conceived possible in this model that pre-adolescent SEP will influence mid-adulthood HRBs via a social pathway. Pre-adolescent SEP will set children on lifelong SEP trajectories, shaping their SEP in mid-adulthood. Access to material and cultural resources in pre-adolescence are considered to accumulate and consequently dictate access to material, cultural and occupational resources in mid-adulthood. This accumulation is conceived to occur through transmission of material and cultural resources on later material, cultural and occupational resources.

Consequently, mid-adulthood SEP will reinforce HRBs embedded in pre-adolescence through the accumulation of resources and subsequently influence HRB cluster membership at the same age. The pathway through which mid-adulthood SEP will dictate HRB cluster membership is conceived to occur via physical access (determined by material resources) to positive HRBs, psychosocial stress (determined by material and occupational resources) leading to differential uptake of negative HRBs and social group habitus (determined by cultural resources) shaping overall lifestyle and encompassing these four HRBs. The material, cultural and occupational dimensions of SEP in mid-adulthood are considered to be highly interrelated.

Finally, this model considers possible differentials in HRB change during mid-life according to SEP. Those in more advantaged social circumstances will have increased access to material, occupational and cultural resources, making HRB change easier in comparison to those in more disadvantaged circumstances.

Figure 3.1 The conceptual model to be tested in this thesis



Note: Solid arrows represent tested paths. Dashed arrows represent hypothesised mechanisms. Gender and cohort are considered to modify HRB clustering.

4 Chapter 4: Aims, objectives and hypotheses

4.1 Aims

The aim of this doctoral project is to investigate the patterning, formation and change over time of HRBs.

4.2 Objectives

The research aim described above is addressed through three research objectives listed below.

1. To identify clustered patterns of four HRBs: smoking, alcohol, diet and physical activity, using the data from two British birth cohort studies at one point in mid-adulthood and compare the results for men and women across those cohorts. More specifically, using the 1958 National Child Development Study (NCDS) at age 33 and the 1970 British Birth Cohort Study (BCS70) at age 34, to investigate:
    a. Cohort and gender differences in cluster membership.
    b. Cohort and gender differences in cluster patterns.
2. To investigate the predictive effect of pre-adolescent SEP at age 11 (NCDS) and age 10 (BCS70) on mid-adulthood HRB cluster membership at age 33 (NCDS) and 34 (BCS70). More specifically:
    a. The effect of pre-adolescent material and cultural dimensions of SEP on mid-adulthood HRB cluster membership.
    b. The extent to which these effects are explained by material, occupational and cultural dimensions of mid-adulthood SEP at age 33 (NCDS) and 34 (BCS70).
3. To assess stability in HRB cluster membership between age 33 and age 42 in the NCDS[7] and the role of dimensions of SEP at age 33 in predicting change in HRB cluster membership.

[7] Only information from the earlier-born cohort was used to address objective 3 due to a lack of equivalence in HRB measurement at ages 34 and 42 in the BCS70.

4.3 Hypotheses

Hypotheses corresponding to the three research objectives are as follows:

Objective 1:

- Distinct clustered patterns of HRB will be identified, some will consist of multiple positive HRBs, others will be characterised by multiple negative HRBs. Based on the appraised literature (see section 2.2.2), there will be a cluster of non-smokers with healthier diets and lower levels of alcohol consumption in comparison to clusters consisting of people who smoke (=multiple positive HRBs) and a smaller cluster consisting of heavy smokers and drinkers with poorer diets but higher levels of physical activity than the cluster consisting of non-smokers (=multiple negative HRBs).
- HRB cluster patterns and/or membership will differ according to cohort and gender. More women and those in the later-born cohort will belong to clusters characterised by multiple positive HRBs and more men and those born earlier will belong to clusters characterised by multiple negative HRBs. Gender convergence over time for smoking and alcohol consumption will mean gender differences in HRB cluster patterns will be less apparent in the later-born cohort.

Objective 2:

- More disadvantaged pre-adolescent SEP will predict membership of mid-adulthood HRB clusters characterised by multiple negative HRBs.
- More disadvantaged mid-adulthood SEP will predict membership of mid-adulthood HRB clusters characterised by multiple negative HRBs.
- The relationship between pre-adolescent SEP and mid-adulthood HRB cluster membership will be partially mediated by mid-adulthood SEP.

Objective 3:

- HRB cluster membership between ages 33 and 42 will be relatively stable with a high probability (>70%) of individuals remaining in the same cluster. There will be a smaller probability (approximately 20%) of moving from a cluster characterised by multiple negative HRBs to a cluster characterised by multiple positive HRBs.[8]
- Age 33 SEP will predict change in HRB clustering during mid-adulthood. More advantaged SEP at age 33 will be associated with change from a smoking to non-smoking cluster.

---

[8] Probabilities based on the findings of Mulder et al (1998) and Benzies et al (2008), see section 2.4.1.

5 Chapter 5: Methods

This chapter describes the methods used to achieve the research objectives outlined in chapter 4. This includes a more detailed account of the data taken from two British birth cohort studies (first introduced in section 1.4), a description of the research variables and an overview of the statistical analytical approach.

5.1 Data

5.1.1 An introduction to the two birth cohort studies

5.1.1.1 *The National Child Development Study (NCDS).*

The National Child Development Study (NCDS) follows the lives of 17,514 live births born within the same week in 1958 across England, Scotland and Wales (Power and Elliott, 2006). The parents of these participants were followed up in order to ascertain the child's development at ages 7, 11 and 16. Immigrants born in the same week, identified through their schools, were added to the sample at age 7, 11 and 16. Participants have been followed throughout their adulthood at ages 20, 23, 33, 42, 46, 50 and 55. The NCDS is a multidisciplinary study which has collated information from a number of sources on the educational, developmental, medical, economic and social aspects of participants' lives (Power and Elliott, 2006).

Data were taken from different data collection sweeps in order to address the three research objectives outlined in chapter 4. To achieve objective 1, information from participants at age 33 (CLS, 2008a) was used. For objective 2, information at birth (CLS, 2014) and ages 7 (CLS, 2014), 11 (CLS, 2014) and 33 (CLS, 2008a) were included. Finally, to address objective 3, information was taken from age 33 (CLS, 2008a) and 42 (CLS, 2008b).

5.1.1.2 *The 1970 British Birth Cohort Study (BCS70).*

The 1970 British Birth Cohort Study (BCS70) follows the lives of 16,571 people born across England, Scotland and Wales in one week in 1970(Elliott and Shepherd, 2006). Information was initially collected from their parents at age 5, 10 and 16. The participants have been followed during adulthood at ages 26, 30, 34, 38, 42 and 46. Like the NCDS, the BCS70 is a multidisciplinary study which has collated information in relation to many aspects of participants' lives, making it a rich data resource (Elliott and Shepherd, 2006).

Again, data from different sweeps was used to address research objectives 1 and 2 outlined in chapter 4. For objective 1, information was taken from participants at age 30 (CLS, 2016c) and 34 (CLS, 2016b). To achieve objective 2, information from ages 5 (Butler, 2016), 10 (CLS, 2016a), 30 (CLS, 2016c) and 34 (CLS, 2016b) were included. Data from the BCS70 was not used to address objective 3 due to a lack of equivalence in HRB measurement at ages 34 and 42.

5.1.2 Study attrition

Similarly to all longitudinal epidemiological studies, the two cohorts had missing data at each data collection sweep (i.e. unavailable at one sweep but then returned for the next), missing data due to a loss of participants over time (i.e. lost in all subsequent sweeps) and non-response to individual items at each sweep.

Table 5.1, presents the response rates for the relevant data collection sweeps of each study. Ferri et al. (2003) conducted a cross-cohort comparison of the two birth cohorts. The authors found the response rates of these cohorts remained high over time and showed relatively little bias. However, the authors do indicate that more men than women and those with the least education have a higher propensity to leave the study over time.

Analytical samples were derived corresponding to each of the three research objectives (see chapter 4). Information on the derivation of each analytical sample and the management of missing data pertaining to that sample are provided in the methods sections of chapters 6, 7 and 8, corresponding to research objectives 1, 2 and 3 respectively.

Table 5.1 Sample sizes and response rates for relevant data collection sweeps in the NCDS and BCS70

NCDS					BCS70				
Year	Age	Collection method*	Sample achieved	% of eligible sample	Year	Age	Collection method*	Sample achieved	% of eligible sample
1958	Birth	Mother; medical records	17,416	99%	1970	Birth	Mother; medical records	16,571	96%
1965	7	Parents; cohort member	15,425	92%	1975	5	Parents; cohort member	13,071	78%
1969	11	Parents; teacher; cohort member	15,337	92%	1980	10	Parents; teacher; cohort member	14,874	86%
1991	33	Cohort member; spouse/partner;	11,407	70%	2000	30	Cohort member;	11,261	66%
2000	42	Cohort member;	11,419	70%	2004	34	Cohort member; spouse/partner	9,665	74%

Note: NCDS source is Power & Elliott (2006) BCS70 source is Elliott & Shepherd (2006). *The respondents listed are those that provided data to achieve each research objective.

5.1.3 Ethical considerations

In both cohorts, data were collected for participants during childhood by trained staff who interviewed the cohort member's parents. In mid-adulthood information was obtained from the participants via home interviews and self-completed questionnaires. Informed consent was obtained from the participants' parents for data collected in childhood and from the participants themselves for data collected in mid-adulthood (Shepherd, 2012a, Shepherd, 2012b). Consent was based on participants' parents and participants agreeing to be interviewed and completing questionnaires, after receiving information on the study and the choice to opt out (Shepherd, 2012a, Shepherd, 2012b).

In both cohorts, ethical approval for data collected from 2000 onwards (i.e. age 42 in the NCDS and age 30 and 34 in the BCS70) was obtained from the appropriate research ethics committee (Shepherd, 2012a, Shepherd, 2012b). Pre-2000 data collection was subject to internal ethical review (Shepherd, 2012a, Shepherd, 2012b).

Both the NCDS and BCS70 are managed by the Centre for Longitudinal Studies. The data pertaining to these two studies can be accessed via the UK Data Archive and is freely available for researchers to download. The data has been carefully anonymised to protect the confidentiality of study participants prior to being used in this doctoral work. Therefore, ethical approval was not required in order to conduct research on these data.

5.2 Variables

5.2.1 HRB measures

5.2.1.1 *Age 33 NCDS and age 34 NCDS*

Four HRBs – smoking, alcohol, diet and physical activity – were measured using six variables: cigarette smoking, alcohol consumption, fruit and vegetable consumption, chips and fried food consumption, sweets, chocolate, biscuits and cakes consumption, physical activity frequency. The alcohol measure is based upon UK government guidelines active in 1991 (when NCDS participants were 33) and 2004 (when BCS70 participants were 34) for 'safe' weekly consumption (DOH, 1995). The measures of smoking, diet and physical activity are pragmatically determined according to the survey questions and the distribution of the variables. As mentioned above, HRB measures were similar, although not identical, in the NCDS at age 33 and BCS70 at age 34. Appendix 5.1 describes the questionnaire items for each cohort and the process of cohort harmonisation.

Smoking

Participants were asked if they smoked cigarettes and the average number smoked per day (range 0–80) those who reported not smoking cigarettes were coded as 0. Those reporting smoking occasionally (BCS70 only, n=645, 6.4%), were also coded as 0.

Alcohol

Participants alcohol consumption was measured according to average drinking frequency and the number of alcoholic beverages consumed in the previous week. Beverage categories were combined to provide the total number of units consumed (1 unit=8g ethanol, range 0–210 units). This total was categorised according to consumption frequency and quantity, reflecting gender-specific UK guidelines for 'safe' weekly consumption that were in effect at the time the data were collated (DOH, 1995). Participants reporting 0 units in the previous week were coded as 'no units' alongside never and infrequent drinkers. Men reporting 1–21 units and women reporting 1–14 units were coded as 'within limits' as were frequent drinkers, reporting 0 units in the previous week. Men reporting >21 units and women reporting >14 units were coded as 'above limits'.

Leisure-time physical activity

Participants were asked whether they regularly took part in leisure-time physical activity, defined as 'at least once a month, for most of the year', and the frequency: 'every day', '4–5 days per week', '2–3 days per week', 'once a week', '2–3 times a month', 'less often'. Responses with sparse data were combined to a neighbouring point, creating four categories: '≤3 times a month', 'Once a week', '2–3 days a week', '4–7 days a week'.

Diet

Diet was indicated by the average frequency of consumption of six variables: 'fruit'; 'vegetables'; 'chips'; 'fried foods'; 'sweets or chocolate'; 'biscuits' (NCDS), and 'biscuits or cakes' (BCS70). In both studies, participants were asked if they consumed these foods 'more than once a day', 'once a day', '3–6 days a week', '1–2 days a week', 'less than 1 day a week' or ' never'. An additional 'occasional' category, present in the BCS70, was combined with 'less than 1 day a week'.

Principal Components Analysis (PCA) is considered an appropriate data reduction method to identify common patterns of food consumption based upon a correlation matrix (Thorpe et al., 2016). PCA was used to reduce the six diet variables into three composite variables: 'fruit and vegetables', 'chips and fried food' (hereafter fried food) and 'sweets, chocolate, biscuits or cakes' (hereafter sweet food). Eigenvalues >1 and factor loadings >0.2 guided the selection of these three components (Thorpe et al., 2016). Frequency scores for each individual diet variable (range 0 to 5) were added together, creating scores ranging from 0 (never) to 10 (more than once a day).

5.2.1.2 *Age 42 NCDS*

The derivation of the four HRBs at age 42 was very similar to that at age 33 in the NCDS.[9] The same six variables were derived: cigarette smoking, alcohol unit consumption, leisure-time physical activity frequency and the consumption frequency of three food groups: fruit and vegetables, chips and fried food, sweets, chocolate, biscuits and cakes. Again, the alcohol measure was based upon UK government guidelines active in 2000 (when NCDS participants were 42) for 'safe' weekly consumption (DOH, 1995). The measures of smoking, diet and physical activity were pragmatically determined, dependent on the survey questions as well as the distribution of the variables.

Slight differences in the wording of the questionnaire items at ages 33 and 42 in the NCDS meant that the variables required harmonisation across the two time points. Appendix 5.2 describes the questionnaire items at each time point and outlines how the six HRB variables were harmonised.

Smoking

Participants were asked if they smoked cigarettes and the average number smoked per day (range age 42=0–70) those who reported not smoking cigarettes were coded as 0. Those reporting to smoke occasionally (age 42 only, n=484, 3.8%), were also coded as 0.

Alcohol

Alcohol consumption was measured according to average drinking frequency and the number of alcoholic beverages consumed in the previous week. At age 42 the consumption of beer/stout/lager/ale/cider (shortened hereafter to 'beer') was incorrectly coded for some participants as units instead of pints (2 units = 1 pint) (Elliott et al., 2007). Moreover, at this sweep incorrect instructions were given to 'beer' drinkers who consumed cans rather than pints of beer (i.e. they were incorrectly advised that 1 large can = 4 * 1 half pint and 1 small can = 2 * 1 half pint whereas the correct conversion was 1 large can = 2 * 1 half pint and 1 small can = 1 * 1 half pint) (CLS, 2004). In consequence, there is potential for an over-

---

[9] Data relating to age 42 in the BCS70 was not included due to HRB variables being different at age 34 and 42 and these variables could not be reasonably harmonised.

estimation of ‘beer’ consumption at age 42. Appendix 5.3 describes the methods employed to correct for this measurement error.

All alcoholic beverage categories (see Appendix 5.2 for a list of beverages included at each age) were combined to provide the total number of units consumed (1 unit=8g ethanol, range age 33=0–210 units; age 42=0–277 units) and categorised according to consumption frequency and quantity, reflecting UK ‘safe’ weekly consumption guidelines (DOH, 1995) for the years of data collection (age 42=2000). Participants reporting 0 units in the previous week were coded as ‘no units’ alongside never and infrequent drinkers. Men reporting 1–21 units and women reporting 1–14 units were coded as ‘within limits’ as were frequent drinkers, reporting 0 units in the previous week. Men reporting >21 units and women reporting >14 units were coded as ‘above limits’.

Leisure-time physical activity

Participants were asked whether they regularly took part in leisure-time physical activity, defined as ‘at least once a month, for most of the year’, and the frequency: ‘every day’, ‘4–5 days per week’, ‘2–3 days per week’, ‘once a week’, ‘2–3 times a month’, ‘less often’. Responses with sparse data were combined, creating four categories: ‘≤3 times a month’, ‘Once a week’, ‘2–3 days a week’, ‘4–7 days a week’.

Diet

Diet was indicated by the average frequency of consumption of six variables: ‘Fruit’, ‘vegetables’, ‘chips’, ‘fried foods’, ‘sweets or chocolate’ and ‘biscuits or cakes’. In both studies, participants were asked if they consumed these foods ‘more than once a day’, ‘once a day’, ‘3–6 days a week’, ‘1–2 days a week’, ‘less than 1 day a week’ or ‘ never’. An additional ‘occasional’ category, present at age 42, was combined with ‘less than 1 day a week’. Again, similarly to age 33, a principal components analysis was undertaken which suggested collapsing the six diet variables into three composite variables: ‘fruit and vegetables’, ‘chips and fried food’ (hereafter fried food) and ‘sweets, chocolate, biscuits or cakes’ (hereafter sweet food). Frequency scores (range 0 to 5) were added together, creating a score ranging from 0 (never) to 10 (more than once a day).

5.2.2 Socio-economic position (SEP)

5.2.2.1 *Pre-adolescent SEP*

Indicators capturing material and cultural dimensions of SEP in pre-adolescence (age 11 in the NCDS and age 10 in the BCS70) were selected. These selections were both theoretically driven (i.e. based on the conceptual framework outlined in chapter 3) and pragmatically driven (i.e. based on the variables available during pre-adolescence).

Material dimension of SEP

Indicators of the material dimension of SEP were taken from age 11 in the NCDS and age 10 in the BCS70 and were as follows: housing tenure, overcrowding, receiving free school meals, being in receipt of state benefits and where available household income. These measures captured the economic circumstances of the household during pre-adolescence.

In order to aid convergence in statistical models and improve interpretability of the results, nominal variables (housing tenure, receiving benefits) were dichotomised. For both cohorts living in council rented housing was distinguished from those not living in council housing (i.e. home owners, private renters, other). Being in receipt of benefits was dichotomised according to three benefit types associated with disadvantage (unemployment benefit, income support and family income supplement). Receiving free school meals remained a binary variable. All variables were coded 1 if disadvantaged and 0 if advantaged.

Although dichotomous in nature, these indicators were not considered to only capture individuals at the bottom of the social hierarchy, i.e. separating individuals experiencing the most disadvantaged from the rest of the sample. This was indicated by cross-tabulations suggesting a varied distribution of the indicators in each cohort.

Overcrowding in pre-adolescence was available as an ordinal variable in both cohorts, consisting of five categories: ‘up to 1 person per room’, ‘1 person per room’, ‘>1 to 1.5 people per room’, ‘>1.5 to 2 people per room’, ‘>2 people per room’.

A seven-category ordinal measure of family gross weekly income at age 10 in the BCS70 was also included to better reflect the gradient of material circumstances amongst this cohort (a continuous measure of net family income was not available). Parents were not directly asked their income level, instead they were given seven income bands to choose from, these were: ‘<£35 per week’, ‘£35–49 per week’, ‘£50–99 per week’, ‘£100–149 per week’, ‘£150–199 per week’, ‘£200–249 per week’, ‘≥£250 per week’. Family income at age 11 in the NCDS was not available.

Cultural dimension of SEP

Parental education and parental interest in their child’s education were selected to capture the cultural dimension of SEP during pre-adolescence.

Parental education was not available at age 11 in the NCDS or age 10 in the BCS70. Information was therefore taken from earlier ages. In the NCDS, mother's education was taken from birth and father's education was taken from age 7. In the BCS70, mother's and father's education was taken from age 5.

In both cohorts, parental education was measured according to the amount of time parents were in school, dichotomised into whether parents remained in education past minimum school-leaving age; this was either > 14 or > 15 in the NCDS (depending on the age of the parent) and > 15 in the BCS70 (Galindo-Rueda, 2003). Parents who did not stay were coded as 1, those who did were coded as 0.

Using information from an earlier time point was justified on the assumption that parental education was unlikely to change substantially during the participant's childhood. This assumption was verified by cross-tabulating parental education at ages 7 (NCDS) and 5 (BCS70) with parental education at age 16 in both cohorts (excluding information for participants whose mother or father was a different person at each age). This analysis found that in both cohorts only a small proportion of mothers (NCDS=13%; BCS70=9%) and fathers (NCDS=14%; BCS70=10%) moved from the 'didn't participate in education past minimum school-leaving age' to the 'did participate in education past minimum school-leaving age' category. Given that this analysis referred to a longer duration (i.e. from age 5 or 7 to age 16) than that assumed (i.e. from age 5 or 7 to age 10 or 11), the level of stability in the educational status of parents during the participant's childhood was considered adequate.

Parental interest in their child's education was measured at age 11 in the NCDS and age 10 in the BCS70, ascertained via the child's school teacher. In the NCDS, the categories were: 'over concerned', 'very interested', 'some interest', 'little interest', coded 1–4 respectively. In the BCS70, the categories were: 'very interested', 'moderately interested', 'very little interest', 'uninterested', coded 1–4 respectively.

5.2.2.2 *Mid-adulthood SEP*

Similarly to pre-adolescence, both theoretically and pragmatically based indicators capturing material, occupational and cultural dimensions of SEP were selected in mid-adulthood.

Material dimension of SEP

Indicators for the material dimension of SEP in mid-adulthood were: receiving benefits, housing tenure, not owning a car, overcrowding and household income all of which captured the purchasing power of the household. The pre-adolescent indicator receiving free school meals, not applicable to adults, was replaced here with car ownership on the basis that during these later periods of data collection in Britain (NCDS=1991; BCS70=2000) it was normative for households to own a car (ONS, 2010b).

Again, nominal variables (housing tenure, receiving benefits) were dichotomised (coded 1 if disadvantaged). Living in council rented housing was distinguished from other tenure categories (i.e. home owners, private renters, other). Individuals receiving any one of three benefits associated with disadvantage (income support, unemployment benefit and housing benefit) were separated from the rest of the sample. Information on car ownership in mid-adulthood was not available at age 34 in the BCS70 and was therefore taken from age 30, under the assumption that there would be little change in car ownership during this four-year period.

Similarly to the material dimension in pre-adolescence, together these dichotomous indicators were considered to capture the social circumstances of participants at different levels of the social hierarchy, rather than only separating the most disadvantaged from the rest of the sample. This was demonstrated by cross-tabulations identifying a varied distribution of the indicators in each cohort.

Overcrowding was again included as an ordinal variable. A continuous version of the variable was discounted in both cohorts after it was found to have a non-normal distribution (with right skewed distributions indicated by positive skewness coefficients and heavy tailed distributions indicated by kurtosis coefficients of >3, (Kirkwood and Sterne, 2003)) and no appropriate transformation (examined using the 'gladder' command in Stata version 14, (StataCorp, 2014)). The continuous variable was therefore divided into four categories: 'up to 1 person per room', '1 person per room', '>1 to 1.5 people per room', '>1.5 people per room'. Unlike in pre-adolescence there were a smaller number of participants living in housing with more than 2 people per room in mid-adulthood therefore this small proportion of participants were incorporated into the '>1.5 people per room' category. This is line with trends of households becoming smaller over time in the United Kingdom (Druckman and Jackson, 2008).

To capture the family unit (Corna, 2013), total household weekly net income was measured on a continuous scale, accounting for the cohort members take home pay and that of any spouse or cohabiting partner. Income was log transformed due to its long right-tailed distribution and was adjusted for the size of the household using the OECD square root method (Anyaegbu, 2010).

Occupational dimension of SEP

In mid-adulthood, the occupational dimension of SEP was represented by employment relations. This was operationalised using the National Statistics Socio-Economic Classification (NS-SEC) (ONS, 2010a) which can be used to assign individuals into one of three ordinal categories. Appendix 5.4 outlines the Office for National Statistics procedure of collapsing the nominal eight- and five-category versions of the NS-SEC in order to construct an ordinal three-category version of the NS-SEC: 'Higher managerial, administrative and professional', 'Intermediate' and 'Routine and manual', coded 1, 2 and 3 respectively.

Participants were allocated to an NS-SEC category using employment type (i.e. employer, self-employed, employee) and the most recent versions of the social occupational classification at the time of data collection (SOC90 in the NCDS; SOC00 in the BCS70), both of which have the same conceptual basis and correspond to one another but the latter has been updated to reflect changes in occupations over time (Elias et al., 1999). Employment relations were conceived to be experienced at an individual level (Sacker et al., 2001) and therefore captured by the occupational SEP of each participant rather than the highest in the household. In both cohorts, participants who were not employed at the time of data collection were asked about their most recent employment experience. Individuals who had never been employed (NCDS unemployed=39, permanently sick/disabled=57, homemakers=303, other=42; BCS70 unemployed=40, permanently sick/disabled=143, homemakers=474, other=67) were classified as missing on the basis that their HRBs would not be influenced by employment relations but via material and cultural dimensions of SEP.

The types of employee benefits participants received through their occupation were also included to capture employment relations, reflecting more stable occupations with better employment conditions (Ginn and Arber, 1993). In the NCDS, participants indicated whether their employer offered four types of employee benefits these were: an employer pension scheme, the chance to buy company shares, access to a company car and private medical insurance (coded 1 if offered and 0 if not offered). In the BCS70, two variables were included. The first related to whether participants were offered a pension scheme by their employer (coded 1 if they were offered the scheme and 0 if they were not). For those who indicated that they were offered the pension scheme, the second variable captured whether the participant was a member of their employer's pension scheme (coded 1 if they were part of the scheme and 0 if they were not).

Cultural dimension of SEP

In mid-adulthood, the cultural dimension of SEP was captured by the participant's highest held qualification. This was considered an indicator of 'institutionalised cultural capital', theorised to encompass all official qualifications (Kamphuis et al., 2015). In both cohorts, education was based on broad classifications of academic and vocational qualifications (Schoon, 2006). The variable had five categories: 'no qualifications', 'CSE 2–5/NVQ1', 'O Level/ NVQ2', 'A Level/NVQ3', 'Diploma or higher qualification below degree/NVQ4', 'Degree or higher/NVQ5 or 6', coded 0–5 respectively.

A continuous variable on the age the cohort member left full-time education was also included. This variable was only available at age 42 in the NCDS. There were a small number of participants who reported having left full-time continuous education after age 33 (n=10) and these were assigned to the 'still in full-time education' category. The number of participants reported to still be in full-time continuous education at age 42 was surprisingly high (n=206). Cross-tabulation of this variable at age 42 with the variable capturing age left full-time continuous education at age 33 found that the majority of these participants classified themselves as employed at age 33 (n=175). Moreover, at age 23 participants were asked the year in which they commenced full-time employment. It may be that these individuals returned to education between ages 23, 33 and 42. However, given that we are interested in participant's continuous educational experiences up to the age of 33, participants reporting being in education after this age or who did not have information at age 42 were recoded according to the age at which they started employment which was asked at age 23 (n=1,341) or if this information was unavailable were coded as missing (n=454). In the NCDS, this variable ranged from 14 to 33. In the BCS70, the small number who reported to still be in full-time continuous education at age 34 (n=8) were coded as missing due to incongruence. A cross-tabulation with employment found that the majority of these participants reported being employed and none reported being students.

The Cambridge scale, a validated measure of the social distance between the participant and individuals in other occupations (Prandy and Lambert, 2003), was also selected to capture patterns of social relationships thus informing the cultural dimension of SEP. The theoretical premise of the scale is that individuals are more likely to socialise with people who share a similar social position to them (Prandy and Lambert, 2003). The scale is based upon information from four surveys, which together reached more than 7,000 individuals, living in five British regions,[10] between the late 1960s and mid-1970s (Prandy, 1990). During these surveys individuals were asked about the occupations of up to four 'people with whom you are friendly' (Prandy, 1990). The survey provided information on patterns of social interaction between occupational units which were used to construct a scale representing the relative position of the occupational unit within the order of social interaction and stratification. A more detailed description of the ranking of occupational units in the scale can be found here.

Latest versions of the social occupational classification at the time of data collection (SOC90 in the NCDS and SOC00 in the BCS70) and employment type (i.e. employer, self-employed, employee) were used to allocate gender-specific Cambridge scale codes to the participants and their partners (Prandy and Lambert, 2003). The highest household Cambridge score was used. Participants' own gender-specific Cambridge score was included and for those married or cohabiting the score of their spouse or partner was also considered if higher, on the premise that married or cohabiting couples are commonly exposed to each other's friendship choices (Sacker et al., 2001). In the NCDS, this resulted in a continuous scale ranging from 10 (least advantaged) to 99 (most advantaged). In the BCS70, highest household Cambridge scale ranged from 20 (least advantaged) to 99 (most advantaged).

5.3 Overview of the statistical analytical approach

In this section, an overview of the statistical analytical methods used to achieve each research objective is provided along with a rationale for their selection. A more detailed description of the operationalisation of each analytical method is provided in the methods sections of chapters 6, 7 and 8, corresponding to research objectives 1, 2 and 3 respectively.

5.3.1 Objective 1: A cross-cohort comparison of HRB clustering

5.3.1.1 *Latent Profile Analysis*

In order to test the hypothesis that distinct HRB cluster patterns will be present (see section 4.3), Latent Profile Analysis (LPA) (Collins and Lanza, 2010) was chosen to identify HRB clustering in the two cohort studies. LPA incorporates continuous variables, but is otherwise identical to latent class analysis (Collins and Lanza, 2010). LPA models have two parameters: the probability of a participant being in a certain class (or in this case HRB cluster) and the probability of their response to the observed variables given their class (or HRB cluster) membership (Colman et al., 2007).

[10] Participants lived in the cities and surrounding areas of Cambridge, Peterborough, Glasgow, Leeds and Leicester.

The rationale for selecting this method was four-fold: 1) it has a number of statistical procedures to guide the selection of clusters (Wang and Wang, 2012); 2) it is more flexible when working with large sample sizes and non-standardised variables (Hofstetter et al., 2014); 3) LPA was also considered advantageous in comparison to cluster analysis due to its theoretical underpinnings (i.e. probability modelling[11]) (Hofstetter et al., 2014); 4) unlike cluster analysis, LPA identifies an unobserved categorical latent variable which explains associations between observed variables (Collins and Lanza, 2010). Using latent rather than observed variables is advantageous because it takes into account variance in the observed variables and consequently minimises measurement error (Hoyle, 2012). Despite the advantages of such data reduction techniques, they have been criticised for being subjective and therefore difficult to replicate across different population samples (Martinez et al., 1998; McAloney et al., 2013). The replication of this analysis within two cohort studies will serve to reduce this element of subjectivity.

In order to explicitly test the hypothesis of differences in HRB cluster patterns according to cohort and gender groups (see section 4.3), measurement invariance analysis was conducted. Measurement invariance assesses 'whether the nature of latent classes differs across known subgroups in the populations' (Finch, 2015: 192).

5.3.2 Objective 2: Pre-adolescent SEP predicting HRB cluster membership in mid-adulthood

In order to test the hypotheses for research objective 2, that pre-adolescent SEP will predict HRB cluster membership in mid-adulthood both directly and indirectly via mid-adulthood SEP, two types of analysis were used. The first approach was used to derive latent variables representing dimensions of SEP in pre-adolescence in mid-adulthood and the second tested the paths between these latent variables and HRB cluster membership.

5.3.2.1 *Exploratory and confirmatory factor analysis*

As with objective 1, two data reduction techniques were chosen to capture latent dimensions of SEP in pre-adolescence and mid-adulthood. These latent variables incorporate measurement error inherent in observed variables (Hoyle, 2012). These were exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) (Wang and Wang, 2012). However, in contrast to LPA, EFA and CFA are considered to be variable-centred (Hofstetter et al., 2014). Each dimension of SEP is conceptualised as being a single continuum on which participants in each cohort are placed (Graham, 2007).

[11] Unlike cluster analysis, LPA takes into account uncertainty in cluster assignment and considers the size of the cluster when assigning participants.

EFA is used in the first instance to explore the shared component of observed indicator variables (Tabachnick et al., 2001). This is followed by CFA which allows measurement errors of the observed variables to be correlated and tests for correlations between the latent variables (Wang and Wang, 2012).

5.3.2.2 *Path analysis*

Path analysis was undertaken to ascertain the extent to which latent variables representing dimensions of SEP in pre-adolescence predict HRB cluster membership in mid-adulthood and whether this association is mediated by dimensions of SEP in mid-adulthood. Path models which can test relations between latent continuous or categorical variables are a type of structural equation model (SEM) (Hoyle, 2012). SEM refers to a wider family of statistical models which bring together narrower statistical models into one framework (Hoyle, 2012).

All SEMs, including the path models used in this thesis, consist of measurement models, representing data reduction methods used to derive the latent variables, which are incorporated into a structural part which is built to conceptually to illustrate relationships between latent and/or observed variables (Hayes, 2013). Section 7.2.3.6 in chapter 7 provides a detailed description along with diagrams of these path models.

Path analysis is considered to be a more powerful tool to elucidate mediation effects when compared to simple regression analysis (Hayes, 2013). This approach allows all conceptualised relationships between variables to be estimated simultaneously. These path models provide evidence of a statistical relationship between variables and thus cannot infer causality (Krieger and Smith, 2016; Russo and Williamson, 2007). Nevertheless, these path models take us a step forward in detecting causal effects due to the temporal ordering of the independent and the dependent variable (Hoyle, 2012), the distance in time between the independent and mediator variable (i.e. reducing the likelihood of variance in the mediator variable being accounted for by the independent variable) (Hoyle and Robinson, 2003) and the use of a latent mediator variable (i.e. measurement error in the mediator variable can be problematic in estimating indirect effects) (Howe et al., 2016).

5.3.3 Objective 3: Transitions in HRB cluster membership over time and the role of mid-adulthood SEP

5.3.3.1 *Latent Transition Analysis*

To test the first hypothesis relating to objective 3 (see section 4.3), that HRB cluster membership will remain relatively stable between ages 33 and 42 although some change is likely to occur, Latent Transition Analysis (LTA) was used.

LTA is a longitudinal extension of LPA (Collins and Lanza, 2010) and is another type of SEM (Hoyle, 2012). LTA is used in this thesis to methodologically build upon the LPA models undertaken to achieve objective 1. The measurement part of the LTA models is the derivation of the HRB clusters at each age (i.e. 33 and 42) and the structural part of the LTA models identifies transitions between HRB clusters at ages 33 and 42 in the NCDS.

LTA models consist of three parameters (Collins and Lanza, 2010). The first is the probability of being in a particular class (or, in this case, HRB cluster) at each time point. The second is the probability of a participants' response to the observed variables given their class (or HRB cluster) membership at each time point, thus assessing the degree of error in each observed indicator in capturing the latent variable. The third is the probability of transitioning to a class (or cluster) at the second time point (i.e. age 42), given class (or HRB cluster) membership at the first time point (i.e. age 33).

5.3.3.2 *Latent Transition Analysis with a covariate*

To test the second hypothesis (see section 4.3), that more advantaged mid-adulthood SEP will be associated with positive transitions in HRB cluster membership, LTA models incorporating mid-adulthood SEP as a covariate were used.

LTA models that incorporate a covariate are also considered a type of SEM (Hoyle, 2012). These 'LTA with a covariate' models build upon the LTA models mentioned above by including additional structural paths testing the extent to which dimensions of SEP in mid-adulthood predict HRB cluster membership at age 33 and their relationship with transitions between HRB clusters during mid-life. Section 8.2.3.4 in chapter 8 provides a detailed description along with diagrams of these 'LTA with a covariate' models.

6 Chapter 6: A cross-cohort comparison of HRB clustering (objective 1)

6.1 Introduction

This chapter addresses the first of the three research objectives set out in chapter 4 of this thesis. In brief, the chapter outlines work undertaken to elucidate the distinct clustering of four HRBs – smoking, alcohol consumption, diet and physical activity – in two British birth cohort studies in order to test for cohort and gender differences.

It is hypothesised that distinct patterns of HRB clusters will be identified some will consist of multiple positive HRBs (i.e. a cluster of non-smokers who will have healthier diets and lower levels of alcohol consumption), others will be characterised by multiple negative HRBs (i.e. a cluster consisting of heavy smokers and drinkers who will have poorer diets but higher levels of physical activity).

Moreover, it is hypothesised that cohort and gender differences in HRB cluster membership and patterns will be detected. More women and those in the later-born cohort will belong to clusters characterised by multiple positive HRBs and more men and those born earlier will belong to clusters characterised by multiple negative HRBs. Gender convergence for smoking and alcohol consumption over time will mean differences in HRB cluster patterns will be less apparent in the later-born cohort.

Latent profile analysis (LPA), an advanced person-centred data reduction technique (Hofstetter et al., 2014; McAloney et al., 2013), is used to elucidate the clustered patterns of HRBs (see section 5.3.1.1). These clustered patterns are perceived as representing underlying health lifestyles that are shared by population subgroups (see chapter 3 for a fuller description of the conceptual framework).

Similarities in HRB clustering across two cohort studies can evidence a generalisation to mid-age adults within Britain today. Cohort differences can increase our understanding of how the experiences unique to a particular cohort may influence their health lifestyles and can serve to generate new hypotheses in relation to the health lifestyles of later-born cohorts.

This chapter begins with an outline of the methods used to achieve objective 1. This is followed by the results and a brief summary of the findings, and it concludes by discussing the strengths and limitations of the research.

6.2 Methods

6.2.1 Analytical sample

Data were taken from the National Child Development Study (NCDS) and the British Cohort Study (BCS70). A full description of the two cohort studies is provided in chapter 5 (see section 5.1.1). In order to achieve this first objective, data were taken from the NCDS when participants were aged 33 (collected in 1991) and from the BCS70 when participants were aged 34 (collected in 2004). As mentioned in chapter 5 (see section 5.2.1), data for BCS70 participants at age 30 (collected in 2000) supplemented information about their diet which was unavailable at age 34.

In the NCDS, data were available for 11,407 participants (response rate 73%). Participants with complete information on at least one HRB were analysed, yielding a sample of 11,373 (99.7%), 5,586 men and 5,787 women. In the BCS70, a total of 9,665 participants were included at age 34 (response rate 75%). Participants with information on at least one HRB, yielded a sample of 9,646 (99.8%), 4,613 men and 5,033 women.

6.2.2 Measures

The four HRBs of interest: Smoking, alcohol, diet and physical activity were measured using six variables: cigarette smoking, alcohol unit consumption, the frequency of fruit and vegetable consumption, fried food consumption and sweet food consumption and leisure-time physical activity frequency. A full description on the derivation of these measures is given in chapter 5 (see section 5.2.1).

6.2.3 Statistical analysis

<u>Descriptive analysis</u>

Descriptive analysis was undertaken on the six variables capturing the four HRBs: alcohol, smoking, diet and physical activity for the total sample and according to cohort and gender.

<u>Latent Profile Analysis (LPA)</u>

Latent Profile Analysis (LPA) (Collins and Lanza, 2010) was conducted using Mplus Version 7 (Muthén, 2014), to identify HRB clustering (see chapter 5, section 5.3, for an overview of this statistical method). LPA models assume that observed variables are conditionally independent, and that associations between them are explained by the latent (i.e. unobserved) variable (Collins and Lanza, 2010). However, in this analysis the assumption of conditional independence was relaxed for the three diet variables which were free to correlate within each cluster. This decision was theoretically driven: it was considered plausible that the frequency consumption of these three food groups were not independent from one another and therefore could not be fully explained by the latent variable.

HRB variables were continuous or ordered, rather than binary, to retain more information on individual differences in the data. Preliminary analysis found smoking to be rare in the largest cluster, this variable also had a long right-tailed distribution. To aid model convergence, the mean and variance of smoking was fixed at zero in the largest cluster and the distribution was condensed by dividing the variable by ten. Preliminary analysis also identified little difference across all of the clusters in terms of the variance for the three continuous diet variables. Additionally, variance for smoking was also similar across the remaining clusters (after fixing variance for smoking in the largest cluster at zero). Therefore, for the sake of model parsimony and stability (Vermunt and Magidson, 2002), the default option available in Mplus Version 7 (Muthén, 2014) was employed which constrains the variance for these continuous HRB variables to be the same across the HRB clusters. In all models 4,000 different starting values were used to identify the maximum likelihood solution (Collins and Lanza, 2010) and avoid the local maxima.

Selecting an appropriate number of clusters and defining them can be challenging (McAloney et al., 2013) and subjective (Martinez et al., 1998). The replication of this analysis within two cohort studies serves to reduce subjectivity. Furthermore, an advantage of LPA is the formal statistical procedure that can be employed to guide cluster selection (Wang and Wang, 2012). To determine an optimal number of clusters, several LPA models were estimated, adding another cluster (k) to each consecutive model and comparing fit indices to the previous model (k-1). Fit indices included the likelihood ratio chi-squared test; entropy; adjusted Bayesian Information Criterion (aBIC); and the Lo-Mendell Rubin likelihood ratio test (LMR) (Collins and Lanza, 2010; Nylund et al., 2007). Emphasis was placed upon the aBIC which balances model fit and parsimony (Collins and Lanza, 2010) adjusts for sample size (Finch, 2015) and has been found to perform well in large samples (Dziak et al., 2012; Nylund et al., 2007). As recommended (Wang and Wang, 2012), alongside these fit statistics, the cluster size (i.e. is there enough statistical power for further analysis?) was taken in account and minimum cluster size criterion were established. Cluster interpretability (i.e. are they theoretically meaningful?) was also taken into consideration (Wang and Wang, 2012).

In order to establish differences according to cohort and gender, multi-group LPA, which identifies structural differences in cluster patterns (Collins and Lanza, 2010), was employed. As recommended (Collins and Lanza, 2010), prior to conducting multi-group LPA, models were run separately for each subgroup (NCDS men, BCS70 men, NCDS women, BCS70 women) to establish whether the same number of clusters emerged. This was followed by multi-group LPA models, run separately for men and women, stratifying the sample according to cohort. Wald chi-square tests were performed to detect differences in HRB means and response probabilities within and across each cohort, for men and women. Wald chi-square tests were used to detect cohort differences in the proportion of participants in each cluster.

Measurement invariance analysis

Measurement invariance analysis was conducted to assess cluster equivalence (i.e. whether the nature of the HRB clusters are the same) across the cohorts (Finch, 2015) and consisted of five stages. Stage 1 established whether, within each subgroup (separated according to gender and cohort), the same optimal number of clusters could be identified. In stage 2, chi-square difference tests using the log-likelihood from multi-group LPA models were conducted: model one allowed values of the observed indicator variables to be free across the cohorts within each gender, model two fixed the observed indicator values to be the same across the cohorts within each gender.

However, scholars argue that such hypothesis testing can be difficult in LPA because models can become very sparse (containing cells with few participants) meaning that the difference in the chi-square between models 1 & 2 is not adequately approximated by the log-likelihood chi-square ratio distribution (Collins et al., 1994). Additionally, with large sample sizes (N>2000) the log-likelihood chi-square difference test can detect statistical differences indicating non-invariance when there are only small substantive differences across groups (Meade and Lautenschlager, 2004).

In consequence, stage three was a cross-validation analysis (Collins and Lanza, 2010). This approach estimated a LPA model based on group 1 (training dataset) and applied these parameters to group 2 (validation dataset) and vice versa. This is done to determine whether the model calibrated in the training dataset has an acceptable model fit in the validation dataset. Assessing whether individuals remain within the same cluster in the validated and calibrated models is also an indication of membership stability to a particular cluster (i.e. classification certainty) (Collins and Lanza, 2010). Models that cross-validate well indicate that the nature of the latent clusters are similar across the groups and that measurement invariance holds (Collins et al., 1994).

For stage 4, indicator variable specific entropy estimates produced by Mplus Version 7 (Muthén, 2014), were examined to determine the extent to which each observed variable contributes to the unobserved (latent) variable. An entropy of below <0.2 suggests that the observed HRB variable contributes little in defining the latent HRB clusters (Asparouhov and Muthén, 2014).

Alongside these assessments of measurement invariance, scholars suggest that researchers consider if the nature of the clusters differ according to subgroups by examining differences in cluster membership (i.e. prevalence) and cluster patterns (i.e. item means and probabilities) (Collins and Lanza, 2010). Therefore, for stage 5 of the measurement invariance analysis, chi-square Wald tests were conducted in order to identify whether levels of the HRBs differed significantly ($p<0.05$) across the clusters within each cohort and within each cluster across cohorts.

Sensitivity analysis for missing data

As mentioned in section 6.2.1 of this chapter, which describes the analytical sample, all participants had information pertaining to at least one HRB. Missing data in relation to the remaining three HRBs was managed using the Full Information Maximum Likelihood (FIML) function in Mplus Version 7 (Muthén, 2014). This approach utilises all available information in the data under a missing at random (MAR) assumption (Enders, 2010). This assumption was deemed adequate as it was considered plausible that missing data for one HRB could be explained by information relating to other HRBs in the model. In order to determine the influence of missing data in the LPA models, a sensitivity analysis was undertaken comparing the HRB item means and response probabilities from complete case models with those from FIML models.

6.3 Results

6.3.1 Descriptive analysis

Table 6.1, summarises participants' responses to all six HRB variables in the total analytical sample and separated according to gender and cohort. For both men and women, behaviours tended to be healthier (e.g. smoking fewer cigarettes per day, higher frequency of fruit and vegetable consumption, higher frequency of physical activity) in the BCS70 compared to the NCDS. An exception was alcohol consumption where a higher proportion of participants in the BCS70 drank above recommended limits compared to the NCDS, particularly amongst women.

For smoking, overall the proportion of men smoking ≥1 cigarette per day was slightly higher than women (29% men; 28% women; p=0.01), the difference is more pronounced when comparing cohort differences within each gender group, for both genders the NCDS is significantly higher than the BCS70 (NCDS=32% men, 28% women; BCS70=26% men, 23% women; p<0.001). The mean number of cigarettes smoked was significantly higher in men compared to women overall (mean=18 men; mean=15 women; p<0.001) and was higher in the NCDS compared to the BCS70 for each gender group (NCDS mean=19 men, mean=16 women; BCS70 mean=16 men, mean=14 women; p<0.001).

Overall men and women ate fruit and vegetables 3–6 days per week. Men in the NCDS ate fruit and vegetables less frequently than men in the BCS70 (NCDS men mean=4.5; BCS70 men mean=4.9; p<0.001.), as did women (NCDS women mean=5.6; BCS70 women mean=5.8; p<0.001.). Men and women ate fried food 1–2 days per week on average; men in the BCS70 ate these foods less than men in the NCDS (NCDS men mean=3.7; BCS70 men mean=3.0; p<0.001), as did women (NCDS women mean=2.9; BCS70 women mean=2.5; p<0.001). In both cohorts, men and women ate sweet foods on average 3–6 days per week.

Focusing on the lowest frequency category of leisure-time physical activity (≤3 times a month), showed that the proportion was slightly lower in women compared to men (31% men; 29% women; p<0.001), indicating a slightly higher frequency of physical activity in women. In women, the proportion was lower in the BCS70 compared to the NCDS (31% NCDS; 27% BCS70; p<0.001), indicating that the frequency of physical activity was higher amongst in women in the BCS70. These differences according to cohort were not found amongst men (p>0.05).

For alcohol consumption, a higher proportion of participants in the BCS70 drank above recommended limits compared to the NCDS, particularly amongst women. There were clear gender differences in the proportion drinking above the UK recommended guidelines, almost three times higher amongst men (29% men; 11% women; p<0.001). For women, the proportion drinking above the UK recommended guidelines was higher in the BCS70 (8% NCDS; 15% BCS70; p<0.001). For men, there was a small increase in the proportion drinking above the UK recommended guidelines in the BCS70 compared to the NCDS (28% NCDS; 31% BCS70; p<0.001), indicating a convergence in drinking behaviour for men and women in the BCS70.

Table 6.1 Health-related behaviour (HRB) characteristics of the analytical sample: Total pooled and stratified by cohort and gender

HRB variables	**Total Pooled N=21,019 (100%)**	**Men NCDS n=5,586 (100%)**	**Men BCS70 n=4,613 (100%)**	**Women NCDS n=5,787 (100%)**	**Women BCS70 n=5,033 (100%)**
	Mean (sd)	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**
Number of cigarettes smoked per day [a]	16.4 (8.5)	18.5 (9.5)	16.0 (7.9)	16.2 (8.2)	13.7 (6.7)
Frequency of fruit and vegetable consumption [b]	5.2 (2.1)	4.5 (1.9)	4.9 (2.1)	5.6 (1.9)	5.8 (2.2)
Frequency of fried food consumption [b]	3.02 (1.3)	3.7 (1.5)	3.0 (1.2)	2.9 (1.3)	2.5 (1.0)
Frequency of sweet food consumption [b]	4.7 (2.2)	4.7 (2.1)	4.6 (2.3)	4.8 (2.2)	4.6 (2.2)
	n (%)	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Diet *Missing*	696 (3.31%)	9 (0.2%)	475 (10.3%)	12 (0.2%)	392 (7.8%)
Proportion smoking cigarettes daily					
0	15,022 (71.5%)	3,797 (68.0%)	3,404 (73.8%)	3,964 (68.5%)	3,857 (76.6%)
1–10	1,934 (9.2%)	458 (8.2%)	385 (8.4%)	573 (9.9%)	518 (10.3%)
11–20	3,159 (15.0%)	912 (16.33%)	680 (14.7%)	984 (17.0%)	583 (11.6%)
21+	842 (4.0%)	393 (7.0%)	135 (2.9%)	249 (4.3%)	65 (1.3%)
Missing	62 (0.3%)	26 (0.5%)	9 (0.2%)	17 (0.3%)	10 (0.2%)
Frequency of leisure-time physical activity					
≤3 times a month	6,300 (30.0%)	1,773 (31.7%)	1,391 (30.2%)	1,775 (30.7%)	1,361 (27.0%)
Once a week	4,102 (19.5%)	1,166 (20.9%)	825 (17.9%)	1,314 (22.7%)	797 (15.8%)
2–3 days a week	4,932 (23.5%)	1,292 (23.1%)	1,237 (26.8%)	1,110 (19.2%)	1,293 (25.7%)
4–7 days a week	5,611 (26.7%)	1,330 (23.8%)	1,156 (25.1%)	1,551 (26.8%)	1,574 (31.3%)
Missing	74 (0.4%)	25 (0.5%)	4 (0.09%)	37 (0.6%)	8 (0.2%)
Alcohol units consumed in the previous week [c]					
No units	4,292 (20.4%)	754 (13.5%)	569 (12.3%)	1,670 (28.9%)	1,299 (25.8%)
Within limits (≤14 units women, ≤21 units men)	12,484 (59.4%)	3,280 (58.7%)	2,578 (55.9%)	3,640 (62.9%)	2,986 (59.3%)
Above limits (≥15 units women, ≥22 units men)	4,212 (20.0%)	1,549 (27.7%)	1,450 (31.4%)	474 (8.2%)	739 (14.7%)
Missing	31 (0.2%)	3 (0.05%)	16 (0.4%)	3 (0.05%)	9 (0.2%)

a. Range 1–80.

b. A Higher score indicates a higher consumption frequency. Range 0–10. Diet score equivalent (rounded to zero decimal places): 'never' [0] 'occasionally /less than 1 day a week' [1–2] '1–2 days a week' [3–4] '3–6 days a week' [5–6] once a day' [7–8] 'more than once a day' [9–10].

c. 'No units' category includes never drinkers and non-frequent drinkers who report 0 units in the previous week. Frequent drinkers who report 0 units in the previous week have been placed in category 'within limits'.

6.3.2 Latent Profile Analysis (LPA)

6.3.2.1 *Model selection*

The aBIC for LPA models run separately for each cohort within each gender (see Table 6.2), suggested four clusters were preferred over three clusters. However, for all groups the smallest cluster in the 4-cluster models fell below the minimum cluster size criterion. Moreover, the HRB patterns in the very small fourth cluster added little to model interpretability. For men in both cohorts and women in the NCDS, the additional fourth cluster was distinguished according to heavy smoking but was otherwise similar to the smallest cluster in the 3-cluster model characterised by heavy smokers. For women in the BCS70, the additional fourth cluster was distinguished according to fried food consumption, but was otherwise similar to the second largest cluster in the 3-cluster model characterised by moderate smoking.

The estimates of the 4-cluster multi-group LPA models can be found in Appendix 6.1 and Appendix 6.2. Model fit indices and minimum cluster size criterion for these models are shown in Table 6.2. The purpose of the minimum cluster size criterion was to ensure adequate statistical power for further analysis. Adequate cluster size was determined by detecting a ‘small’ difference using Cohen’s effect size of 20% (Cohen, 1992) between two independent proportions (0.1, 0.3), with 80% power and significance level of 0.05. Sample size calculations were conducted in Stata version 14 (StataCorp, 2014) using the ‘power two proportions’ command.

In sum, on the basis of model fit statistics, cluster size and interpretability, the 3-cluster multi-group LPA model was chosen for both genders.

Table 6.2 Goodness of fit indices for Latent Profile Analysis (LPA) models, stratified by cohort for each gender

NCDS Men	Log-likelihood	aBIC	LMR	Entropy	Smallest Cluster Size (n)	MMC (n=124)[a]
2 cluster	39034.135	78220.878	<0.001	0.990	1768 (31.7%)	Yes
3 cluster	38917.438	78058.340	0.03	0.978	82 (1.5%)	No
4 cluster	38850.141	77994.599	0.06	0.917	80 (1.4%)	No
BCS70 Men	**Log-likelihood**	**aBIC**	**LMR**	**Entropy**	**Smallest Cluster Size (n)**	**MMC (n=124)[a]**
2 cluster	27440.492	55028.237	<0.001	0.993	1200 (26.0%)	Yes
3 cluster	27316.210	54848.039	0.04	0.975	79 (1.7%)	No
4 cluster	27214.809	54713.604	0.003	0.981	24 (0.5%)	No
NCDS Women	**Log-likelihood**	**aBIC**	**LMR**	**Entropy**	**Smallest Cluster Size (n)**	**MMC (n=124)[a]**
2 cluster	39075.405	78304.409	<0.001	0.991	1808 (31.2%)	Yes
3 cluster	38915.657	78056.225	<0.001	0.899	515 (8.9%)	Yes
4 cluster	38849.292	77994.811	0.01	0.905	59 (1.0%)	No
BS70 Women	**Log-likelihood**	**aBIC**	**LMR**	**Entropy**	**Smallest Cluster Size (n)**	**MMC (n=124)[a]**
2 cluster	28922.370	57994.432	<0.001	0.993	1166 (23.2%)	Yes
3 cluster	28844.586	57908.362	0.02	0.940	183 (3.6%)	Yes
4 cluster	28789.143	57866.976	0.09	0.951	33 (0.7%)	No

Note: aBIC=adjusted Bayesian Information Criteron; LMR= Lo-Mendell-Rubin Likelihood Ratio Test *p* value. MMC=Meets Minimum Cluster Size Criterion. Superscript a= Adequate cluster size determined by detecting a 20% difference in two independent proportions (0.1, 0.3), with 80% power and significance level of 0.05.

6.3.2.2 *Measurement invariance analysis*

Measurement invariance analysis was conducted, consisting of five stages, in order to determine cluster equivalence across the cohorts.

The above results indicated that the 3-cluster solution was optimal for all subgroups, thus completing stage 1 of the measurement invariance analysis.

Table 6.3 and Table 6.4, present the results of stage 2, chi-square difference tests comparing the log-likelihood from multi-group LPA models which allow the observed indicator variables to be free and then fixed across the cohorts. Significant p values (p<0.001) were found in all instances, indicating that the fit of the model with fixed parameters was substantially worse. As the measurement models for the two cohorts differed significantly, full measurement invariance did not hold (Finch, 2015).

Table 6.3 Estimates from multi-group models with and without cluster patterns and membership constrained to be equal (Men)

Men FIML multi-group 3 cluster model	Log-likelihood	Scaling correction factor	Number of parameters	Chi-square difference (p value)
Cluster patterns free	-73256.573	1.2612	83	533.69 (<0.001)
Cluster patterns fixed	-73841.447	1.0215	66	
Cluster membership free	-73256.573	1.2612	83	32.65 (<0.001)
Cluster membership fixed	-73276.564	1.2621	81	

Table 6.4 Estimates from multi-group models with and without cluster patterns and membership constrained to be equal (Women)

Women FIML multi-group 3 cluster model	Log-likelihood	Scaling correction factor	Number of parameters	Chi-square difference (p value)
Cluster patterns free	-75233.802	1.1511	83	870.54 (<0.001)
Cluster patterns fixed	-75896.731	1.0553	66	
Cluster membership free	-75233.802	1.1511	83	43.80 (<0.001)
Cluster membership fixed	-75279.763	1.1277	81	

Table 6.5 and Table 6.6, show the results of stage 3, a cross-validation analysis, to determine whether the model calibrated in the training dataset has an acceptable model fit in the validation dataset. Based on model fit indices and entropy, we found that models calibrated in each subgroup (separated by cohort and gender) cross-validated reasonably well when applied to data from the same gender in the other cohort. This was taken as evidence of cluster equivalence across cohorts (Collins et al., 1994).

Table 6.5 Estimates from cross-validation analysis (Men)

Men 3 cluster FIML models	Log-likelihood	Entropy	Adjusted BIC	Lo-Mendell-Rubin LRT *p* value[a]
NCDS Men calibrated	-38559.333	0.981	77341.679	0.04
NCDS Men validated	-40265.503	0.950	80531.007	<0.001
BCS70 Men calibrated	-26838.585	0.980	53888.898	<0.001
BCS70 Men validated	-27698.580	0.992	55397.160	<0.001

Note: Calibrated models = model parameters estimated in this gender and cohort. Validated models = gender and cohort in which the model is validated using saved model parameter estimates from the model calibrated in the same gender but opposite cohort data. Superscript a = Tests the null hypotheses that the addition of a fourth cluster does not improve model fit.

Table 6.6 Estimates from cross-validation analysis (Women)

Women 3 cluster FIML models	Log-likelihood	Entropy	Adjusted BIC	Lo-Mendell-Rubin LRT *p* value[a]
NCDS Women calibrated	-38558.954	0.901	77342.313	<0.001
NCDS Women validated	-40604.876	0.907	81209.751	<0.001
BCS70 Women calibrated	-28595.963	0.943	57408.650	0.02
BCS70 Women validated	-29787.017	0.933	59574.034	>0.05

Note: Calibrated models = model parameters estimated in this gender and cohort. Validated models = gender and cohort in which the model is validated using saved model parameter estimates from the model calibrated in the same gender but opposite cohort data. Superscript a = Tests the null hypotheses that the addition of a fourth cluster does not improve model fit.

The stability of cluster membership in the validated and calibrated models was investigated and are shown in Table 6.7 to Table 6.10. Cluster stability was deemed to be excellent for the largest two clusters (clusters 2 and 3), with ≥95% of individuals being assigned to the same cluster in the calibrated and validated models. Cluster classification was also deemed to be good for the smallest cluster (cluster 1) amongst women given that ≥77% remained in the same cluster in the calibrated and validated models. However, there appeared to be classification uncertainty for cluster 1 amongst men. This could, in part, be due to the particularly small number of male participants assigned to this cluster in the calibrated models.

Table 6.7 Cluster stability in the validated and calibrated models (NCDS Men)

NCDS Men	Cluster 1 validated n (%)	Cluster 2 validated n (%)	Cluster 3 validated n (%)	Total
Cluster 1 calibrated	26 (8.1)	56 (3.9)	0	82
Cluster 2 calibrated	288 (89.2)	1,398 (96.2)	0	1,686
Cluster 3 calibrated	9 (2.8)	0	3,809 (100.0)	3,818
Total	323 (100.0)	1,454 (100.0)	3,809 (100.0)	5,586

Table 6.8 Cluster stability in the validated and calibrated models (BCS70 Men)

BCS70 Men	Cluster 1 validated n (%)	Cluster 2 validated n (%)	Cluster 3 validated n (%)	Total
Cluster 1 calibrated	11 (44.0)	65 (5.5)	3 (0.1)	79
Cluster 2 calibrated	14 (56.0)	1,110 (94.5)	0	1,124
Cluster 3 calibrated	0	0	3,410 (99.9)	3,410
Total	25 (100.0)	1,175 (100.0)	3,413 (100.0)	4,613

Table 6.9 Cluster stability in the validated and calibrated models (NCDS Women)

NCDS Women	Cluster 1 validated n (%)	Cluster 2 validated n (%)	Cluster 3 validated n (%)	Total
Cluster 1 calibrated	444 (78.6)	70 (5.6)	1 (0.1)	515
Cluster 2 calibrated	121 (21.4)	1,171 (94.4)	0	1,292
Cluster 3 calibrated	0	0	3,980 (99.9)	3,980
Total	565 (100.0)	1,241 (100.0)	3,981 (100.0)	5,787

Table 6.10 Cluster stability in the validated and calibrated models (BCS70 Women)

BCS70 Women	Cluster 1 validated n (%)	Cluster 2 validated n (%)	Cluster 3 validated n (%)	Total
Cluster 1 calibrated	136 (76.8)	46 (4.7)	1 (0.1)	183
Cluster 2 calibrated	41 (23.2)	943 (95.4)	0	984
Cluster 3 calibrated	0	0	3,866 (99.9)	3,866
Total	177 (100.0)	989 (100.0)	3,867 (100.0)	5,033

Table 6.11 provides the indicator variable specific entropy estimates, used to achieve stage 4 of the measurement invariance analysis. This demonstrates that the six HRB variables all had entropy >0.2, suggesting that each variable contributed to the latent HRB variable (Asparouhov and Muthén, 2014). Moreover, the entropy values for each HRB variable were similar across the cohorts, although these differed for women more than men. These results were taken as further evidence of some cluster equivalence across the cohorts, although the greater contribution of alcohol to the formation HRB clusters amongst BCS70 women compared to NCDS women, again suggests partial measurement invariance for alcohol consumption across the cohorts.

Table 6.11 HRB variable specific entropy estimates according to cohort and gender

Indicator variable	NCDS Men	BCS70 Men	NCDS Women	BCS70 Women
Smoking	0.97	0.93	0.84	0.91
Alcohol	0.38	0.43	0.26	0.41
Fruit and vegetables	0.39	0.42	0.32	0.42
Fried food	0.39	0.47	0.30	0.43
Sweet food	0.38	0.41	0.27	0.41
Physical activity	0.38	0.42	0.28	0.41

Finally, Chi-square Wald tests, identifying whether HRBs differed significantly both across the clusters within each cohort and within each cluster across cohorts, were undertaken in order to determine whether the nature of the HRB clusters differed substantively. These cluster patterns and the results of the chi-square Wald test are presented in the following section of this chapter which presents the results of the optimal LPA solution (see Table 6.12 and Table 6.13).

These tests implied that the cluster patterns across the cohorts were similar, except for alcohol consumption, particularly amongst women and suggested partial measurement invariance across cohorts and genders. Attempts were made to run LPA models which pooled the subgroups together (increasing statistical power) and allowing for one LPA model with partial measurement invariance for alcohol (i.e. accounting for alcohol differences). However, due to increased complexity and data sparseness (given the size of the smallest cluster) the partial measurement invariant LPA model would not converge. Therefore, the three cluster multi-group LPA models, comparing cohorts and run separately for each gender, were considered the optimal solution.

In summary, on the basis of the above analysis, it was concluded that there was some cluster equivalence for each gender group across the two cohorts. However, the results do suggest that the clusters may not be equivalent for alcohol consumption amongst women and that the stability of membership of the smallest cluster (i.e. cluster 1) amongst men in the two cohorts is questionable.

6.3.2.3 *The 3-cluster LPA model*

The same cluster labels were assigned across the three cluster multi-group LPA models, aiding interpretability. This was based on the measurement invariance analysis (outlined in the preceding section), which found evidence for cluster equivalence across the cohorts, within each gender group, except in relation to alcohol consumption amongst women.

Cluster 1, labelled ‘Risky’, had patterns riskier than the others (i.e. heavy smoking). Cluster 2, was labelled ‘Moderate Smokers’, because smoking behaviour notably distinguished this cluster from the others, although levels of smoking were lower than the ‘Risky’ cluster. Cluster 3, labelled ‘Mainstream’, was the largest cluster, representing the most prevalent HRB patterns in the data, described below.

Cluster patterns

For both genders, cluster patterns were similar across the cohorts for smoking, fruit and vegetable consumption, fried food consumption and physical activity. Patterns diverged slightly for sweet food and alcohol consumption amongst men but notably for alcohol consumption amongst women.

Wald chi-square tests found the estimated mean number of cigarettes smoked per day was higher for members in the 'Risky' and 'Moderate Smokers' clusters ('Risky' NCDS men=41 cigarettes, NCDS women=21 cigarettes; 'Moderate Smokers' NCDS men=17 cigarettes, NCDS women=14 cigarettes), compared to those in the 'Mainstream' cluster ($p<0.01$), which was fixed at zero in line with our methodological approach.

Members of the 'Risky' cluster had lower frequencies of fruit and vegetable consumption (mean NCDS men=2.61; BCS70 men=3.75; NCDS women=3.39; BCS70 women=3.67) and higher frequencies of fried food consumption (mean NCDS men=4.73; BCS70 men=6.73; NCDS women=4.02; BCS70 women=3.37) compared to members of the 'Moderate Smokers' and 'Mainstream' clusters ($p<0.01$).

The frequency of leisure-time physical activity was highest for the members of the 'Mainstream' cluster (≥once per week NCDS men=72%; BCS70 men=73%; NCDS women=73%; BCS70 women=76%), followed respectively by the members of the 'Moderate Smokers' and 'Risky' clusters ($p<0.01$).

Sweet food consumption frequency was generally highest in the 'Mainstream' cluster and lowest in the 'Risky' cluster ($p<0.01$). The exception was BCS70 men whose sweet food consumption frequency was high in the 'Mainstream' cluster (mean=4.59) but highest in the 'Risky' cluster (mean=5.23, $p<0.01$). In substantive terms, this difference across the two clusters was deemed negligible, given that after rounding sweet food consumption to one decimal place, the mean sweet food consumption was 5 in both the 'Risky' and 'Mainstream' clusters, equating to the same frequency category '3–6 days per week'. In women, sweet food consumption frequency in the 'Mainstream' cluster was significantly lower amongst BCS70 members (mean=4.60) compared to NCDS members (mean=4.85, $p<0.01$). Again, whilst different, after rounding to one decimal place, the mean sweet food consumption was 5, thus equating to the same frequency category '3–6 days per week' in both cohorts.

For both genders, alcohol consumption was lowest for members of the 'Mainstream' cluster across cohorts (p<0.01). For men, NCDS members of the 'Risky' cluster had the highest proportion drinking alcohol above recommended limits (51%) compared to the 'Moderate Smokers' cluster (36%), whereas proportions were similar for BCS70 members ('Risky'=42%; 'Moderate Smokers'=43%). For women, the proportion of BCS70 members drinking above recommended limits was almost double that of NCDS members, across all three clusters ('Risky' NCDS=19%, BCS70=32%; 'Moderate Smokers' NCDS=9%, BCS70=18%; 'Mainstream' NCDS=7%, BCS70=19%, p<0.01).

Cluster membership

For both genders, Wald chi-square tests indicated a significant difference in cluster membership across the cohorts. For men and women, a significantly higher proportion (p<0.01) of BCS70 participants (Men=73.8%; Women=76.7%) were members of 'Mainstream' cluster compared to NCDS participants (Men=68.2%; Women=68.6%).

The cluster prevalence based on the estimated models is presented in Table 6.12 and Table 6.13 Cluster prevalence based on assigning each participant to their most likely HRB cluster, known as 'modal assignment' (Heron et al., 2015), as well as the probabilities of cluster assignment are shown in Appendix 6.3.

Sensitivity analysis for missing data

Estimates from models using FIML, presented in Table 6.12 and Table 6.13 , were very similar to estimates using complete cases, which can be found in Table 6.14 and Table 6.15 .

Table 6.12 Estimated means and item response probabilities FIML of 3-cluster multi-group Latent Profile Analysis (LPA) model for men

	NCDS Men n=5,586 (100%)			BCS70 Men n=4,613 (100%)		
	Cluster 1 'Risky' n=96 (1.7%)≠	**Cluster 2 'Moderate Smokers' n=1679 (30.1%)≠**	**Cluster 3 'Mainstream' n=3811 (68.2%)≠**	**Cluster 1 'Risky' n=93 (2.0%)≠**	**Cluster 2 'Moderate Smokers' n=1117 (24.2%)≠**	**Cluster 3 'Mainstream' n=3403 (73.8%)≠**
	Mean (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	40.84 (3.67)*†	17.22 (0.31)*†	0	19.82 (4.46)†	15.57 (0.35)†	0
Frequency of fruit and vegetable consumption	2.61 (0.37)*†	3.95 (0.05)*†	4.64 (0.03)*†	3.75 (0.28)*†	4.29 (0.07)*†	5.10 (0.04)*†
Frequency of fried food consumption	4.73 (0.45)*†	3.99 (0.05)*†	3.36 (0.02)*†	6.74 (0.29)*†	3.02 (0.04)*†	2.86 (0.02)*†
Frequency of sweet food consumption	3.58 (0.45)*	4.18 (0.06)*	4.71 (0.04)*	5.23 (0.53)*	4.34 (0.08)*	4.59 (0.04)*
	IRP (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity						
≤3 times a month	0.61 (0.07)*	0.39 (0.01)*	0.28 (0.01)*†	0.49 (0.07)*	0.41 (0.02)*	0.26 (0.01)*†
Once a week	0.12 (0.05)	0.21 (0.01)	0.21 (0.01)	0.09 (0.04)	0.18 (0.01)	0.18 (0.01)
2–3 days a week	0.14 (0.04)	0.19 (0.01)	0.25 (0.01)	0.20 (0.05)	0.20 (0.01)	0.29 (0.01)
4–7 days a week	0.13 (0.04)	0.21 (0.01)	0.26 (0.01)	0.23 (0.06)	0.22 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week						
No units	0.26 (0.08)*	0.14 (0.01)*†	0.13 (0.01)*†	0.26 (0.06)*	0.13 (0.01)*†	0.12 (0.01)*†
Within limits (≤14 units women, ≤21 units men)	0.23 (0.06)	0.50 (0.01)	0.63 (0.01)	0.31 (0.07)	0.44 (0.02)	0.61 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.51 (0.08)	0.36 (0.01)	0.24 (0.01)	0.42 (0.09)	0.43 (0.02)	0.28 (0.01)

Note: Cluster prevalence based on estimated model. *=cluster means and response probabilities are significantly different (p ≤ 0.05) across the three clusters within each cohort. †=cluster means and response probabilities are significantly different (p ≤ 0.01) across the cohorts. ≠ = cluster membership is significantly different (p ≤ 0.01) across the cohorts. Estimated using the Wald chi-square test.

Table 6.13 Estimated means and item response probabilities of FIML 3 cluster multi-group Latent Profile Analysis (LPA) model for women

	NCDS Women Total N=5,787 (100%)			BCS70 Women Total N=5,033 (100%)		
	Cluster 1 'Risky' n=561 (9.7%)≠	**Cluster 2 'Moderate Smokers' n=1254 (21.7%)≠**	**Cluster 3 'Mainstream' n=3972 (68.6%)≠**	**Cluster 1 'Risky' n=227 (4.5%)≠**	**Cluster 2 'Moderate Smokers' n=944 (18.8%)≠**	**Cluster 3 'Mainstream' n=3862 (76.7%)≠**
	Mean (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	20.96 (1.00)*†	14.07 (0.31)*	0	19.18 (1.88)†	12.30 (0.39)	0
Frequency of fruit and vegetable consumption	3.39 (0.15)*†	5.57 (0.14)*	5.79 (0.03)*†	3.67 (0.16)*†	5.41 (0.20)*	5.97 (0.04)*†
Frequency of fried food consumption	4.02 (0.15)*†	2.69 (0.07)*†	2.55 (0.02)*†	3.37 (0.30)*†	2.32 (0.07)*†	2.36 (0.02)*†
Frequency of sweet food consumption	3.76 (0.24)*	4.40 (0.10)*†	4.85 (0.04)*†	3.68 (0.27)*	4.50 (0.12)*†	4.60 (0.04)*†
	IRP (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity						
≤3 times a month	0.62 (0.03)*†	0.29 (0.03)*	0.27 (0.01)*†	0.55 (0.06)*†	0.31 (0.03)*	0.25 (0.01)*†
Once a week	0.16 (0.02)	0.21 (0.01)	0.24 (0.01)	0.08 (0.03)	0.15 (0.01)	0.17 (0.01)
2–3 days a week	0.07 (0.02)	0.20 (0.01)	0.21 (0.01)	0.07 (0.03)	0.23 (0.02)	0.28 (0.01)
4–7 days a week	0.15 (0.02)	0.31 (0.01)	0.28 (0.01)	0.30 (0.05)	0.32 (0.02)	0.31 (0.01)
Alcohol units consumed in the previous week						
No units	0.27 (0.03)*	0.30 (0.01)*†	0.29 (0.01)*†	0.40 (0.05)*	0.27 (0.02)*†	0.24 (0.01)*†
Within limits (≤14 units women, ≤21 units men)	0.54 (0.03)	0.61 (0.01)	0.65 (0.01)	0.28 (0.08)	0.54 (0.02)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.19 (0.03)	0.09 (0.01)	0.07 (0.01)	0.32 (0.08)	0.18 (0.02)	0.13 (0.01)

Note: Cluster prevalence based on estimated model. *=cluster means and response probabilities are significantly different (p ≤ 0.05) across the three clusters within each cohort. †=cluster means and response probabilities are significantly different (p ≤ 0.01) across the cohorts. ≠ = cluster membership is significantly different (p ≤ 0.01) across the cohorts. Estimated using the Wald chi-square test.

Table 6.14 Estimated means and item response probabilities of 3-cluster multiple-group Latent Profile Analysis (LPA) model, using complete cases, for men

	NCDS Men Total N=5,525 (100%)			BCS70 Men Total N=4,195 (100%)		
	Cluster 1 'Risky' n=91 (1.6%)≠	**Cluster 2 'Moderate Smokers' n=1677 (30.2%)≠**	**Cluster 3 'Mainstream' n=3767 (68.2%)≠**	**Cluster 1 'Risky' n=83 (2.0%)≠**	**Cluster 2 'Moderate Smokers' n=1001 (23.9%)≠**	**Cluster 3 'Mainstream' n=3111 (74.2%)≠**
	Mean (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	41.24 (3.88)*†	17.23 (0.32)*†	0*	18.14 (2.15)†	15.60 (0.26)†	0
Frequency of fruit and vegetable consumption	2.62 (0.42)*	3.95 (0.05)*†	4.64 (0.03)*†	3.77 (0.27)*	4.29 (0.07)*†	5.10 (0.04)*†
Frequency of fried food consumption	4.69 (0.51)*†	3.99 (0.05)*†	3.36 (0.02)*†	6.50 (0.27)*†	3.01 (0.04)*†	2.86 (0.02)*†
Frequency of sweet food consumption	3.56 (0.49)*†	4.18 (0.06)*	4.71 (0.04)*	5.29 (0.49)*†	4.34 (0.08)*	4.59 (0.04)*
	IRP (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity						
≤3 times a month	0.60 (0.07)*	0.39 (0.01)*	0.28 (0.01)*†	0.49 (0.07)*	0.40 (0.02)*	0.27 (0.01)*†
Once a week	0.13 (0.06)	0.21 (0.01)	0.21 (0.01)	0.09 (0.04)	0.18 (0.01)	0.18 (0.01)
2–3 days a week	0.14 (0.05)	0.19 (0.01)	0.25 (0.01)	0.21 (0.05)	0.20 (0.01)	0.30 (0.01)
4–7 days a week	0.13 (0.05)	0.21 (0.01)	0.26 (0.01)	0.21 (0.05)	0.22 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week						
No units	0.25 (0.08)*	0.14 (0.01)*†	0.13 (0.01)*†	0.27 (0.06)*	0.14 (0.01)*†	0.12 (0.01)*†
Within limits (≤14 units women, ≤21 units men)	0.23 (0.06)	0.50 (0.01)	0.63 (0.01)	0.33 (0.06)	0.44 (0.02)	0.61 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.52 (0.09)	0.36 (0.01)	0.24 (0.01)	0.40 (0.08)	0.43 (0.02)	0.27 (0.01)

Note: Cluster prevalence based on estimated model. *=cluster means and response probabilities are significantly different (p ≤ 0.05) across the three clusters within each cohort. †=cluster means and response probabilities are significantly different (p ≤ 0.01) across the cohorts. ≠ = cluster membership is significantly different (p ≤ 0.01) across the cohorts. Estimated using the Wald chi-square test.

Table 6.15 Estimated means and item response probabilities of 3-cluster multiple-group Latent Profile Analysis (LPA) model, using complete cases, for women

	NCDS Women Total N=5,716 (100%)			**BCS70 Women Total N=4,739 (100%)**		
	Cluster 1 'Risky' n=549 (9.6%)≠	**Cluster 2 'Moderate Smokers' n=1251 (21.9%)≠**	**Cluster 3 'Mainstream' n=3916 (68.5%)≠**	**Cluster 1 'Risky' n=215 (4.5%)≠**	**Cluster 2 'Moderate Smokers' n=873 (18.4%)≠**	**Cluster 3 'Mainstream' n=3652 (77.1%)≠**
	Mean (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	21.01 (1.01)*	14.09 (0.31)*	0	19.27 (1.86)	12.29 (0.39)	0
Frequency of fruit and vegetable consumption	3.38 (0.15)*†	5.56 (0.14)*†	5.79 (0.03)*†	3.66 (0.16)*†	5.42 (0.20)*†	5.97 (0.04)*†
Frequency of fried food consumption	4.03 (0.16)*†	2.70 (0.08)*	2.55 (0.02)*†	3.35 (0.29)*†	2.32 (0.07)*	2.36 (0.02)*†
Frequency of sweet food consumption	3.75 (0.25)*†	4.41 (0.10)*	4.85 (0.04)*†	3.68 (0.27)*†	4.51 (0.12)*	4.60 (0.04)*†
	IRP (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity						
≤3 times a month	0.62 (0.03)*	0.29 (0.03)*†	0.27 (0.01)*†	0.55 (0.06)*	0.30 (0.03)*†	0.25 (0.01)*†
Once a week	0.16 (0.02)	0.21 (0.01)	0.24 (0.01)	0.08 (0.03)	0.15 (0.01)	0.17 (0.01)
2–3 days a week	0.07 (0.02)	0.20 (0.01)	0.21 (0.01)	0.07 (0.03)	0.23 (0.02)	0.28 (0.01)
4–7 days a week	0.15 (0.02)	0.31 (0.01)	0.28 (0.01)	0.30 (0.05)	0.32 (0.02)	0.31 (0.01)
Alcohol units consumed in the previous week						
No units	0.27 (0.03)*†	0.30 (0.01)*†	0.29 (0.01)*†	0.40 (0.05)*†	0.28 (0.02)*†	0.24 (0.01)*†
Within limits (≤14 units women, ≤21 units men)	0.54 (0.03)	0.61 (0.01)	0.65 (0.01)	0.28 (0.08)	0.53 (0.02)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.19 (0.03)	0.09 (0.01)	0.07 (0.01)	0.32 (0.08)	0.18 (0.02)	0.13 (0.01)

Note: Cluster prevalence based on estimated model. *=cluster means and response probabilities are significantly different ($p \leq 0.05$) across the three clusters within each cohort. †=cluster means and response probabilities are significantly different ($p \leq 0.01$) across the cohorts. ≠ = cluster membership is significantly different ($p \leq 0.01$) across the cohorts. Estimated using the Wald chi-square test.

6.4 Summary of findings

The hypotheses stated at the beginning of this chapter (see section 6.1) were partially confirmed. Three distinct clusters of HRBs were identified and subsequently labelled ‘Risky’, ‘Moderate Smokers’ and ‘Mainstream’. The ‘Mainstream’ cluster largely consisted of participants with multiple positive HRBs (e.g. non-smokers, moderate drinkers, frequent fruit and vegetable consumers), the ‘Risky’ cluster was largely characterised by individuals with multiple negative HRBs (e.g. heavy smokers and drinkers) and the ‘Moderate Smokers’ cluster was a mixture of both.

Cohort differences in HRB cluster membership were found, those in the later-born cohort belonging to clusters characterised by multiple positive HRBs. However, there was little evidence of differences in cluster membership according to gender. Whilst HRB cluster patterns were similar according to cohort and gender with respect to some HRBs, other HRBs were found to differ and there was evidence of convergence in alcohol consumption for men and women in the later-born cohort.

For both genders, cluster patterns were similar across the two cohorts in relation to smoking, fruit and vegetable consumption, fried food consumption and physical activity. The HRBs of members in the ‘Mainstream’ cluster tended to be more beneficial to health than the other two clusters (i.e. not smoking, eating fruit and vegetables more frequently, chips and fried food less frequently and being more active), based upon evidence linking these four HRBs to mortality (Kvaavik et al., 2010; Loef and Walach, 2012). However, the frequency of sweet food consumption was generally higher in the ‘Mainstream’ cluster. The exception was for men in the BCS70 whereby sweet food consumption was found, in substantive terms, to be the same in the ‘Risky’ and ‘Mainstream’ cluster. In the later-born cohort there was a significant shift in membership towards the ‘Mainstream’ cluster. Moreover, the hypothesis that heavy smokers and drinkers would also be more physically activity was not supported with members of the ‘Risky’ cluster having lower levels of physical activity in comparison to the other clusters.

The distribution of alcohol consumption across the three clusters differed by cohort in both genders but was particularly apparent for women. The proportion of BCS70 women drinking above recommended guidelines across the three clusters was almost double that of NCDS women, with a sizeable proportion (>10%) of women in the BCS70 drinking above recommended levels in the 'Mainstream' cluster. At the same time, BCS70 women in the 'Mainstream' cluster consumed sweet foods less frequently than NCDS women.

A higher proportion of 'Risky' cluster members (NCDS men, BCS70 men and women) reported drinking 'no units' in the previous week. Although members of the 'Risky' cluster drink differently (i.e. not drinking or drinking excessively) they are assigned to the same cluster by sharing other behaviours, particularly smoking. Research investigating these four HRBs suggests smoking to be the most persistent (Paavola et al., 2004) and is strongly associated with heavy alcohol consumption (Bien and Burge, 1990; Chiolero et al., 2006; Room, 2004; Zacny, 1989) and alcohol abstainers who have previously drunk alcohol (De Leon et al., 2007).

Membership of the 'Mainstream' cluster was found to be higher in the BCS70, compared to the NCDS. The shift to the 'Mainstream' cluster is considered to be beneficial for health in some respects, especially cigarette smoking. Moreover, BCS70 members had higher frequencies of fruit and vegetable consumption, lower frequencies of fried food consumption and were more physically active in this cluster, compared to NCDS members. However, in comparison to NCDS participants, a higher proportion of BCS70 men and women in the 'Mainstream' cluster were drinking alcohol above the recommended guidelines (DOH, 1995). Additionally, the frequency of sweet food consumption tended to be higher in the 'Mainstream' cluster compared to the other two clusters. The exception was men in the BCS70 who consumed sweet foods on average '3–6 days per week' in both the 'Risky' and 'Mainstream' cluster. Amongst women, sweet food consumption frequency was lower in the 'Mainstream' cluster for BCS70 compared to NCDS members, although substantively they equated to the same frequency category of '3–6 days per week'.

Despite some cohort differences, the measurement invariance analyses indicated largely consistent cluster patterns of smoking, fruit and vegetable consumption, fried food consumption and physical activity across the two cohorts for both genders, implying that these clusters could be generalised to individuals in mid-adulthood in Britain today.

6.5 Strengths and limitations

The research undertaken in this chapter has a number of strengths. This work maximised the efficiencies of data reduction techniques by treating variables in the study model as continuous or ordered, identifying clusters that may have been missed if variables were dichotomised. LPA allows for the investigation of multiple rather than individual HRBs

(McAloney et al., 2013) and provided new insights with existing data by detecting a previously unobserved mixture of three clusters. The study detected cohort differences in HRB clustering according to gender, made possible by the large sample size. Furthermore, the replication of this analysis within two cohort studies reduced the possibility of subjectivity, which is a criticism of data reduction techniques (Martinez et al., 1998; McAloney et al., 2013).

The purpose of this first chapter was to empirically derive clustered patterns of HRBs in each cohort. The complex web of HRB patterns identified here suggests distinct typologies are practised by different types of people providing a person-centred understanding. However, the size of the 'Mainstream' cluster suggests that they are likely to be a heterogeneous group which should be taken into consideration when interpreting these results. Consequently, consideration of additional covariates would allow a greater understanding of the types of people who may share these patterns of HRBs, improving the relevance of these results to policy and public health practice. Furthermore, additional covariates would strengthen the assertions made here regarding the equivalence of the clusters across the two cohorts. The addition of covariates pertaining to social circumstances is addressed in the second objective of this doctoral work which is fully outlined in chapter 7.

When interpreting these results, other limitations should also be considered. To aid LPA model convergence the mean and variance of smoking in the 'Mainstream' cluster was set at 0 because whilst smoking in this cluster did exist it was considered rare (cigarettes per day: NCDS mean=0.5; BCS70 mean=0.3). Sensitivity analysis indicated that this decision did not affect the LPA model estimates (results not shown), with only a small proportion of smokers (NCDS=7.6%, BCS70=5.2%) in this cluster.

It was necessary to use data from ages 30 and 34 in the BCS70, because dietary information was not available at age 34. The dietary habits of participants were considered to have remained relatively stable during this period based upon empirical evidence (Parsons et al., 2006). However, this approach left 663 individuals (7%) with incomplete data. Men had more missing data on diet at age 30 than women (p<0.001). Despite this limitation, similar estimates were found in sensitivity analysis comparing models using FIML and those using complete cases suggesting little impact of this caveat on the overall findings.

During the twelve years separating the two cohorts the average serving size of spirits and wine has increased (Stead et al., 2013), potentially underestimating alcohol consumption among BCS70 participants. On the other hand, a higher proportion of participants in the BCS70 drinking above recommended limits were found when compared to the NCDS. Therefore, correcting for a potential underestimation of alcohol consumption in the BCS70 would not change the direction of these findings.

This study relies on self-reported measures of HRB which can be biased (Conry et al., 2011; Heroux et al., 2012; Schneider et al., 2009; Worsley et al., 2012). However, both cohorts collected data on HRB variables using well-structured questionnaires and in the BCS70 all interviews were assisted with a computer, reducing interviewer error. Although HRB measures were not identical across the cohorts, using data from two purposefully similar birth cohort studies allowed for a valid comparison (Ekinsmyth et al., 1992) and reduced the likelihood of cohort and gender differences due to differential measurement. Furthermore, differential measurement would suggest a uniformed bias, instead the results suggested larger cohort differences amongst women, indicating other contextual factors are likely to be at play.

The inclusion of additional measures of the four HRBs would have been insightful and may have uncovered some of the heterogeneity that is likely to exist within the large 'Mainstream' cluster. For example, the intensity of leisure-time physical activity using metabolic equivalents (Conry et al., 2011; van Nieuwenhuijzen et al., 2009) as well as consideration of other types of physical activity, such as that undertaken as part of occupation and commuting, may have been beneficial (Poortinga, 2007; van Nieuwenhuijzen et al., 2009). At the same time, leisure-time physical activity is considered to be more salient for health than occupational physical activity, with only the former found to be associated with aerobic fitness (Mundwiler et al., 2017). Leisure-time physical activity is also strongly associated with commuting physical activity and can therefore be a marker of exercise associated with travel (Becker and Zimmermann-Stenzel, 2008). The use of food diaries may have improved the measurement of dietary intake (Conry et al., 2011; Patterson et al., 1994), although these are not without their limitations (Thompson and Subar, 2008). As recommended (Nugawela et al., 2016), information pertaining to alcohol consumption frequency and episodic heavy alcohol use (i.e. binge drinking) may also have been useful in elucidating additional cluster patterns.

Social desirability bias, i.e. a tendency to respond in a way that reflects social norms (Crowne and Marlowe, 1960), is a possibility potentially leading to an underestimation of health-damaging HRBs and overestimating of health-promoting HRBs. This may explain, to some extent, the large size of the 'Mainstream' cluster. However, there are similarities between the prevalence of current smoking in both cohorts (NCDS 1991=32%, BCS70 2004=24%) and that reported by the Health and Social Care Information Centre (HSCIC) in 2014, for persons aged 35–49 during the same time periods (1990=34%, 2004=29%) (HSCIC, 2014b).

Also, interpretation of gender and cohort differences in cluster membership requires caution due to differences in cluster patterns – e.g. alcohol consumption in the ‘Risky’ cluster for NCDS women resembled that of the ‘Moderate Smokers’ cluster for BCS70 women and membership of the ‘Risky’ cluster was higher amongst women compared to men in both cohorts. This indicates ‘partial’ measurement invariance (Collins and Lanza, 2010) (i.e. the ‘Risky’ clusters cannot be interpreted the same way across subgroups). This demonstrates the importance of investigating measurement invariance in HRB cluster patterns according to cohort and gender, rather than treating these demographic variables as covariates that only predict HRB cluster membership. As mentioned in section 2.2.3.3, the finding of gender differences in the nature of HRB clusters has been identified elsewhere (Bondy and Rehm, 1998).

However, the measurement invariance analysis suggested equivalence of the ‘Moderate Smokers’ and ‘Mainstream’ clusters across the two cohorts. This work identified consistent cluster patterns for smoking, fruit and vegetable consumption, fried food consumption and physical activity. Therefore, it is considered likely that the ‘Moderate Smokers’ and ‘Mainstream’ cluster patterns, are similar among individuals in mid-adulthood in Britain more generally.

7 Chapter 7: Pre-adolescent SEP predicting HRB cluster membership in mid-adulthood (objective 2)

7.1 Introduction

This chapter addresses the second research objective which is fully outlined in chapter 4. To summarise, it investigates the extent to which socio-economic position (SEP) in pre-adolescence and mid-adulthood predict membership of three distinct clustered patterns of HRBs: ‘Risky’, ‘Moderate Smokers’ and ‘Mainstream’, derived in the previous chapter (see chapter 6).

This chapter seeks to extend existing research (outlined fully in the literature review chapter of this thesis, see section 2.2), by considering the role of childhood SEP in predicting mid-adulthood lifestyles. There is a clear need for further research on the effects of childhood SEP on HRB clustering within the British context. Only one other study has been identified that examines the effects of childhood and adulthood SEP on HRB clustering in a sample of Swedish working-age adults (Falkstedt et al., 2016), finding disadvantaged SEP in childhood and adulthood was predictive of membership in clusters characterised by three or four health-damaging behaviours.

A direct effect of pre-adolescent SEP on mid-adulthood lifestyles is hypothesised (see section 4.3). This direct effect is conceived to occur through behaviours embedded in pre-adolescence (i.e. behavioural pathway). More disadvantaged pre-adolescent SEP will be associated with membership of HRB clusters characterised by multiple negative HRBs. Additionally, an indirect effect of pre-adolescent SEP on mid-adulthood lifestyles is hypothesised. This indirect effect is conceived to occur through mid-adulthood SEP. The accumulation of resources from pre-adolescence to mid-adulthood will dictate the social circumstances in which people live their lives in mid-adulthood and consequently their lifestyles (i.e. social pathway). More disadvantaged pre-adolescent SEP will be associated with more disadvantaged mid-adulthood SEP which will be associated with membership of HRB clusters characterised by multiple negative HRBs. Examining the role of pre-adolescent SEP in relation to adult HRB clustering could show how social circumstances at an early stage of the lifecourse might link to a particular HRB cluster.

7.2 Methods

7.2.1 Analytical sample

Data were taken from both the National Child Development Study (NCDS) and the British Birth Cohort Study (BCS70). The analytical sample was the same as the one used in the previous chapter (see chapter 6), including participants who had information on at least one out of the four HRBs (smoking, alcohol, diet, physical activity) in early mid-adulthood. This yielded an analytical sample of 11,373 participants at age 33 in the NCDS and 9,464 participants at age 34 in the BCS70. All of the participants in the existing analytical samples used in chapter 6 were retained in this analysis because they had information on at least one SEP indicator from either pre-adolescence or mid-adulthood. The data used here were ethically collected (see section 5.1.3).

7.2.2 Measures

7.2.2.1 *Outcome: HRB cluster membership*

The dependent variable in the analysis is HRB cluster membership, representing three distinct clustered patterns of HRBs: 'Risky', 'Moderate Smokers' and 'Mainstream' and derived separately for each cohort and gender group (see section 6.3.2.3).

HRB cluster membership was treated as observed, rather than latent. Treating latent variables as observed variables is known as the 'three-step approach' (Vermunt, 2010). This approach aids model convergence because it is less computationally intensive and retains the nature of the HRB clusters when incorporating covariates in further analyses (Heron et al., 2015; Vermunt, 2010). This was operationalised by assigning each participant to their most likely HRB cluster, known as 'modal assignment' (Heron et al., 2015). However, it should be acknowledged that, unlike the latent variable, the observed dependent variable did not consider any classification error in HRB cluster assignment. The absence of classification error is important because it can lead to under-estimated standard errors of regression coefficients in path models (Clark and Muthén, 2009; Heron et al., 2015).

The HRB cluster assignment classification error in the original latent variable was considered to be low. This was indicated by an entropy of 0.9 in the original measurement model used to derive the latent HRB variable (see section 6.3.2.1), which is above the 0.8 cut-off point suggested (Clark et al., 2013). Therefore, bias introduced by 'modal assignment' in the observed dependent variable used in this analysis was likely to be minimal. However, further steps were taken to consider any possible bias. These steps are fully outlined below in section 7.2.3.7.

7.2.2.2 *Predictor: Pre-adolescent SEP*

As mentioned in chapters 3, social circumstances in pre-adolescence are conceived to shape HRBs via material and cultural dimensions of SEP. Differentials in material resources will shape HRBs by determining physical access to health-promoting HRBs. Exposure to particular HRBs in pre-adolescence will also depend on cultural resources that shape social group habitus.

The indicators used to capture the material and cultural dimensions of SEP in pre-adolescence are outlined in chapter 5 (see section 5.2.2.1).

7.2.2.3 *Mediator: Mid-adulthood SEP*

As mentioned in chapters 3, social circumstances in mid-adulthood are conceived to shape HRBs via material, cultural and occupational dimensions of SEP. Material resources determine physical access to health-promoting HRBs, material and occupational resources relate to differentials in psychosocial stress which influence the uptake of health-damaging HRBs and social group habitus is shaped by cultural resources.

The indicators used to capture the material, occupational and cultural dimensions of SEP in mid-adulthood are presented in chapter 5 (see section 5.2.2.2).

7.2.3 Statistical analysis

An overview of the statistical analysis used in this chapter is provided in chapter 5 (see section 5.3). This section provides further details on how these statistical methods were operationalised in order to address research objective 2 of this thesis.

All statistical analyses were stratified according to cohort and gender, given the outcome variable, i.e. HRB cluster membership, was derived separately for each cohort and gender group (see chapter 6).

7.2.3.1 *Descriptive and bivariate analysis*

Descriptive and bivariate analyses were performed in Stata version 14 (StataCorp, 2014), to determine the distribution of the SEP indicator variables and ascertain the statistical significance and direction of the relationship between each SEP indicator and the nominal HRB cluster dependent variable. Bivariate analyses used multinomial logistic regression models with the ‘Mainstream’ cluster as the baseline category.

7.2.3.2 *Exploratory and Confirmatory Factor Analysis to derive material, occupational and material dimensions of SEP*

The samples were randomly split in half. Exploratory Factor Analysis (EFA) was undertaken on one half of the data in order to explore the shared component of the SEP indicator variables for each dimension of SEP. Following this exploratory work, Confirmatory Factor Analysis (CFA) was undertaken on the other half of the data in order to confirm that the SEP indicators adequately captured the dimensions of SEP in pre-adolescence (material and cultural) or mid-adulthood (material, occupational and cultural) that they were hypothesised to measure. Both types of analyses were stratified according to cohort and gender groups.

Both EFA and CFA models were run in Mplus Version 7 (Muthén, 2014) and estimated using Weighted Least Squares (WLS) with robust standard errors (MV) which is less computationally demanding than robust maximum likelihood (MLR). WLSMV is the Mplus default estimator when using a combination of continuous and categorical indicator variables (Wang and Wang, 2012). A sensitivity analysis comparing models estimated using WLSMV and MLR found them to be similar (results not shown). Slow computation and non-convergence was more likely to occur for EFA and CFA models estimated using MLR, reinforcing the decision to use WLSMV.

A reflective approach to modelling the latent dimensions of SEP was employed (Hagger-Johnson et al., 2011). This approach assumed that the latent dimensions of SEP explained observed associations between the SEP indicator variables. The reflective approach was chosen, as opposed to the formative approach, on the basis of previous work suggesting that reflective models are easier to compare across studies (Hagger-Johnson et al., 2011) and that model identification is less problematic when using the reflective than the formative approach (von Stumm et al., 2013).

When running EFA models with a WLSMV estimator, a polychoric matrix is used to handle categorical and continuous indicator variables (Muthén, 2012). Oblique rotation was chosen due to the high likelihood of significant correlations between the latent dimensions of SEP at each age. Eigenvalues >1, factor loadings >0.32 and factor communality (Tabachnick et al., 2001) were assessed to determine whether the SEP indicators contributed to their respective dimension of SEP.

In the CFA models, standardised factor loadings, standard errors and p values of the SEP indicator variables were examined. Indicator variables with factor loadings >0.32 and p values <0.05 were considered to contribute moderately to the dimension of SEP that they were hypothesised to measure (Tabachnick et al., 2001). Indicators with weaker loadings (<0.32) were also retained if they were significant for at least one gender group ($p<0.05$). Model fit was assessed using the Comparative Fit Index (CFI) (Bentler, 1990) as well as the Root Mean Square Error of Approximation (RMSEA) and the RMSEA close fit test (Steiger, 1990). Adequate model fit was determined by a CFI of >0.9, the RMSEA being <0.05 and a RMSEA close fit test p value >0.05 (Wang and Wang, 2012).

7.2.3.3 *Handling multicollinearity between dimensions of SEP in pre-adolescence and mid-adulthood*

The result of the EFA and CFA analyses were five continuous latent variables representing each dimension of SEP at each age: material and cultural in pre-adolescence and material, occupational and cultural in mid-adulthood. A higher score was indicative of more social disadvantage. The SEP indicator variables were found to significantly contribute to their respective dimension of SEP, with most factors loadings being >0.32 and p values <0.05. The results of the EFA and CFA are presented in Appendix 7.1 to Appendix 7.4.

Unlike EFA, CFA allows for causal correlations between the latent variables (Wang and Wang, 2012). The CFA identified high correlations (ranging from 0.6 to 0.9, $p<0.001$) between the dimensions of SEP at each age. These correlations were considered to pose a high risk of multicollinearity in future models incorporating HRB cluster membership as a dependent variable (Farrar and Glauber, 1967).

Higher order CFA models were considered to circumvent this problem (Thompson, 2004). However, convergence of these higher order models was problematic for both pre-adolescent and mid-adulthood SEP indicators, partly due to the high number of categorical variables being used. Whilst it is possible to construct a continuous latent variable using categorical indicator variables (Muthén, 2012), the difficulties in model identification are likely due to the limited ability of categorical indicator variables to fully capture the tails of the continuous factors.

Additionally, the CFA modification indices in models where all five of the latent variables were estimated together suggested a strong correlation between two mid-adulthood SEP indicators. These were the NS-SEC, capturing the occupational dimension of SEP and the highest household Cambridge scale, capturing the cultural dimension of SEP. This correlation was considered plausible given their similarity in measurement, although these two indicators are conceived to be theoretically distinct. The lack of discrimination of these measures in the CFA models was likely due to the NS-SEC being reduced to three ordinal categories.[12] The high

[12]Nominal variables could not be included in CFA models in version 7 of Mplus when using the WLSMV estimator. Models using the MLR estimator to estimate these dimensions would not converge.

correlation of these measures in the CFA models indicated that that the occupational and cultural dimensions of SEP could not be considered distinct from one another.

In consequence, the five dimensions of SEP (material and cultural in pre-adolescence and material, occupational and cultural in mid-adulthood) could not be distinguished from one another for further analyses. Instead a single SEP construct was derived at each age (i.e. one in pre-adolescence and one in mid-adulthood). These constructs were conceptualised as uni-dimensional, parsimoniously capturing the interrelationship between the dimensions of SEP at each age.

7.2.3.4 *Testing paths between dimensions of SEP in pre-adolescence and mid-adulthood*

Although the dimensions of SEP could not be modelled separately for further analyses that included HRB cluster membership as a dependent variable, each dimension of SEP in mid-adulthood was regressed onto each dimension of SEP in pre-adolescence in order to examine the hypothesised paths between the dimensions of SEP in pre-adolescence and mid-adulthood as described in the conceptual model (see chapter 3), by identifying the direction and significance of the relationship between the dimensions of SEP at each age.

This was operationalised using linear regression in Mplus Version 7 (Muthén, 2014). By default, Mplus Version 7 (Muthén, 2014) assumes the latent factor to be normally distributed, fixing the mean at zero. The TECH4 information in the Mplus output (Muthén, 2012) showed the latent factors to have a mean of zero accompanied by an estimated variance for each latent variable (Appendix 7.5) which suggested the distribution of the latent factors for each dimension of SEP were adequate for the purposes of this analysis.

The results of this analysis are presented in Appendix 7.6 and confirm statistically significant ($p<0.001$) positive associations between dimensions of SEP in pre-adolescence and mid-adulthood (i.e. being more socially disadvantaged in each dimension of SEP in pre-adolescence predicts being more socially disadvantaged in each dimension of SEP in mid-adulthood).

7.2.3.5 *Exploratory and Confirmatory Factor Analysis to derive a uni-dimensional construct of SEP*

Through preliminary EFA and CFA analyses it was established that material, occupational and cultural dimensions of SEP could not be distinguished from one another for further analyses (see 7.2.3.2). Consequently, two uni-dimensional constructs, representing SEP at each age, were derived. EFA models were run in one half of the data in order to explore the shared component of the SEP indicator variables for each uni-dimensional SEP construct. CFA models were run on the other half of the data in order to confirm that the SEP indicators adequately captured the uni-dimensional construct of SEP construct that they were hypothesised to measure.

The same thresholds used in the previous EFA and CFA models (see section 7.2.3.2) were applied to determine the contribution of each indicator variable to the uni-dimensional construct of SEP that they were hypothesised to measure (i.e. factor loadings >0.32 and p values <0.05 were considered to contribute moderately (Tabachnick et al., 2001)). The same model fit criteria were also applied to the CFA models (i.e. CFI of >0.9, the RMSEA being <0.05, and a RMSEA close fit test p value >0.05 (Wang and Wang, 2012)).

The CFA models allowed measurement errors of the SEP indicators to be correlated (Wang and Wang, 2012). To improve model fit, the assumption that the association between SEP indicators is fully accounted for by the latent variable was relaxed and the measurement errors among certain SEP indicators were free to correlate. These correlations were guided by model modification indices and theoretical plausibility, such as variables which were similar in wording and measurement (Wang and Wang, 2012).

7.2.3.6 *Path analysis*

Path analysis that incorporates latent variables is a type of SEM (Hoyle, 2012). As mentioned in section 5.3.2.2, the path model consisted of a measurement model which was incorporated into a structural path model built conceptually to illustrate relationships between uni-dimensional SEP constructs in pre-adolescence and mid-adulthood and HRB cluster membership in mid-adulthood.

The path model simultaneously estimated the direct relationship between the SEP construct in pre-adolescence and HRB clustering in mid-adulthood as well as the indirect effect via the SEP construct in mid-adulthood. These models were run separately according to cohort and gender. The measurement part of the model was the CFA model used to derive uni-dimensional constructs representing SEP in pre-adolescent and mid-adulthood (see section 7.2.3.5). The structural element of each model simultaneously regresses the observed dependent categorical variable containing three HRB clusters: ‘Risky’, ‘Moderate Smokers’ and ‘Mainstream’, onto the SEP constructs for pre-adolescent (the independent variable) and mid-adulthood (the mediator variable).

Mediation analysis was employed in order to decompose direct and indirect effects (Fairchild and McQuillin, 2010). This analysis separated the influence of pre-adolescent SEP on mid-adulthood HRB into two components 1) the direct effect of pre-adolescent SEP on mid-adulthood HRB 2) the indirect effect of pre-adolescent SEP on mid-adulthood HRB through its influence on mid-adulthood SEP.

Figure 7.1 is a diagram of the CFA model (i.e. the measurement model part of the path model) and describes how observed indicator variables, representing the material, occupational and cultural dimensions of SEP, are conceived to be related to the SEP construct at each age.

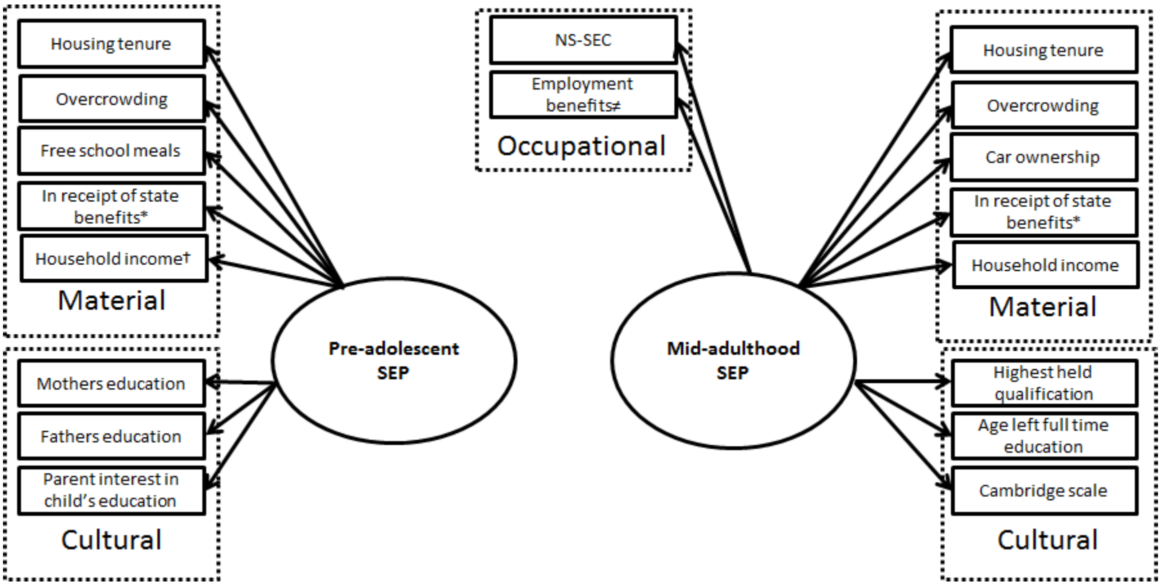
Figure 7.2 provides a visual representation of the structural part of the path model illustrating the tested paths between pre-adolescent and mid-adulthood SEP constructs and mid-adulthood HRB cluster membership. The path labelled ‘c*’ represents the direct relationship between pre-adolescent SEP and mid-adulthood HRB cluster membership. The path labelled ‘a’ represents the relationship between pre-adolescent SEP and mid-adulthood SEP and the path labelled ‘b’ represents the relationship between mid-adulthood SEP and mid-adulthood HRB cluster membership. Together, paths ‘a’ and ‘b’ represent the indirect relationship between pre-adolescent SEP and mid-adulthood HRB cluster membership (i.e. ‘a’ x ‘b’). The total effect of pre-adolescent SEP on mid-adulthood HRB cluster membership is represented by ‘c’ which is equal to the sum of the direct and indirect effect (i.e. $c = c^* + ab$) (Hayes, 2013).

This analysis was operationalised using conventional mediation (Howe et al., 2016), i.e. simultaneously regressing HRB cluster membership onto pre-adolescent SEP with and without mid-adulthood SEP. These models were estimated using the ‘Model Indirect’ command in Mplus Version 7 (Muthén, 2012). A WLSMV estimator was used which is recommended when modelling categorical outcomes (Wang and Wang, 2012).

When modelling a binary outcome, probit logistic regressions are estimated in order to identify the relationship between the dependent, independent and mediator variables (Muthén, 2012). The goodness of fit statistics, previously used in the CFA models (i.e. CFI, RMSEA), were used again to determine the fit of the path models. As recommended by others (Hayes, 2013), bias-corrected bootstrapped confidence intervals, based on 10,000 iterations, were estimated in order to account for asymmetric confidence intervals of the indirect effect.

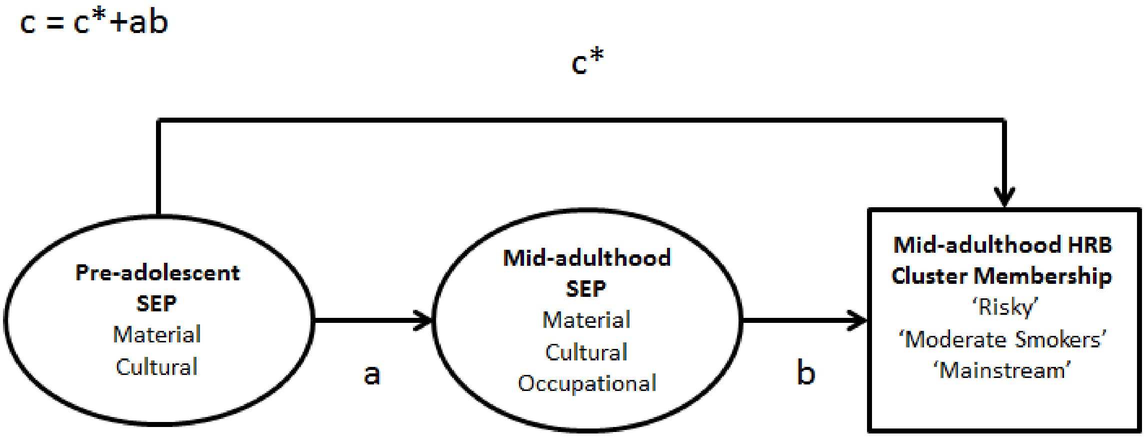
Currently the ‘Model Indirect’ command cannot be applied to nominal outcomes, therefore two binary HRB variables were created to analyse HRB cluster membership in Mplus Version 7 (Muthén, 2014). The ‘Mainstream’ cluster was included as the reference category for both variables given that it represents the most prevalent HRB pattern. The first binary variable compared membership of the ‘Risky’ cluster (coded as 1) with ‘Mainstream’ cluster membership (coded as 0), the second compared membership of the ‘Moderate Smokers’ cluster (coded as 1) with ‘Mainstream’ cluster membership (coded as 0).

Figure 7.1 The measurement model part of path models demonstrating the relationship between material, occupational and cultural dimension indicator variables and the SEP construct in pre-adolescence and mid-adulthood in the NCDS and BCS70



Note: Models run separately according to gender and cohort (NCDS men N=5,586; NCDS women N=5,787; BCS70 men N=4,613; BCS70 women N=5,033).
SEP = socio-economic position.
* = In pre-adolescence: unemployment benefit, income support and family income supplement. In mid-adulthood: income support, unemployment benefit and housing benefit.
† = Household income in pre-adolescence only available in BCS70.
≠ = In the NCDS employment benefits are: offered a pension scheme, has chance to buy company shares, access to a company car, given private medical insurance. In the BCS70 employment benefits are: offered a pension scheme, part of pension scheme.
Pre-adolescence = age 11 in the NCDS, age 10 in the BCS70.
Mid-adulthood = age 33 in the NCDS, age 34 in the BCS70.
Ovals represent the SEP construct (latent variable). Rectangles represent the dimension indicators (observed variables).

Figure 7.2 The structural part of path models estimating the effect of pre-adolescent SEP on mid-adulthood HRB cluster membership in the NCDS and BCS70



Note: Models run separately according to gender and cohort (NCDS men N=5,586; NCDS women N=5,787; BCS70 men N=4,613; BCS70 women N=5,033).
SEP = socio-economic position.
Pre-adolescence = age 11 in the NCDS, age 10 in the BCS70.
Mid-adulthood = age 33 in the NCDS, age 34 in the BCS70.
Oval represents latent variable estimated in measurement model part of path models. Rectangle represents observed variable (based on estimates from a previous measurement model).
Path a x b = indirect path between pre-adolescent SEP and mid-adulthood HRB cluster membership.
Path c* = direct path between pre-adolescent SEP and HRB cluster membership.
Total effect of pre-adolescent SEP on HRB cluster membership denoted as 'c' (=c*+ab).

7.2.3.7 *Sensitivity analysis using multinomial logistic regression models*

Sensitivity analysis was conducted by running multinomial logistic regression models. The rationale for these models was two-fold. Firstly, these models enabled the dependent variable to be modelled as a nominal rather than binary variable. Secondly, the multinomial regression models enabled an investigation of the effect of assigning participants to their most likely HRB cluster, known as 'modal assignment' (Heron et al., 2015).

7.2.3.7.1 HRB cluster membership as a nominal outcome

The multinomial regression models, allowed all three HRB clusters to be included in the analysis. Whilst these models are not directly comparable with the path models, given that mid-adulthood SEP was not formally modelled as a mediator in the multinomial regressions, they do elucidate direct effects of pre-adolescent SEP on HRB cluster membership, after accounting for mid-adulthood SEP. Thus, the extent to which estimates in the path analysis were impacted by the specification of the dependent variable as two binary rather than one nominal variable could be ascertained by comparing the direct effects identified in the multinomial regression models with the direct effects identified in the path analysis. Similar results would imply that differential treatment of the dependent variable does not alter the overall conclusions made on the basis of the path models.

In order to analyse the HRB cluster membership as a nominal rather than binary dependent variable, the 'Risky' cluster was coded 1, the 'Moderate Smokers' cluster was coded 2 and the 'Mainstream' cluster was coded 3. The latter cluster was the reference category as it represented the most prevalent HRB cluster patterns in the data.

7.2.3.7.2 Modal assignment and classification error

As mentioned above (section 7.2.2.1), the dependent variable was treated as observed in the path models therefore participants were assigned to their most likely HRB cluster, known as 'modal assignment' (Heron et al., 2015). Whilst this approach aids model convergence and prevents previously defined latent variables being influenced by covariates in the model, it does not consider classification error and can lead to an underestimation of standard errors in regression models (Heron et al., 2015).

As recommended (Clark and Muthén, 2009), the precautionary step of imposing a higher threshold to determine if associations between the dependent variable and other variables were statistically significant was employed for the path models. The alpha threshold was increased from $p<0.05$ to $p<0.01$.

Classification error was assumed to be low in HRB cluster assignment due to the entropy of the original measurement model used to derive the latent HRB variable (see section 6.3.2.1) being above the 0.8 cut-off point suggested as indicating low classification error (Clark and Muthén, 2009). However, recently scholars have suggested that using entropy as a global measure of classification error in the model is not sufficient to justify most likely cluster membership assignment and that the classification probabilities for each cluster should also be taken into considered (Heron et al., 2015).

Therefore, comparisons of estimates from multinomial regression models with and without classification error in the dependent variable indicated the extent to which the absence of this classification error in the path models may influence the results. These models regressed a latent dependent variable, representing HRB cluster membership, onto observed independent and mediator variables, representing pre-adolescent and mid-adulthood SEP constructs. One of the models incorporated the classification error associated with the assignment of participants to their most likely HRB cluster (i.e. using starting values from the original 3-cluster measurement models) whereas the other model did not incorporate this classification error.

This sensitivity analysis was operationalised using a SEM framework (Hoyle, 2012) by embedding a LPA measurement model to derive the HRB clusters into a structural model testing the relationship between the latent HRB cluster variable and the pre-adolescent and mid-adulthood uni-dimensional SEP constructs. The SEMs were estimated using the 'mixture' type command and the MLR estimator in Mplus Version 7 (Muthén, 2014). This approach allowed for the derivation of the latent HRB cluster variable but could not be used to derive the latent SEP constructs, which were estimated using the WLSMV estimator (see section 7.2.3.2). Consequently, the SEMs incorporated factor scores, which are observed rather than latent variables, representing the SEP constructs in pre-adolescence and mid-adulthood. Factor scores for the SEP constructs were saved using the regression method (DiStefano et al., 2009). This method of factor score estimation provides a high correlation between the observed estimated factor score and the latent construct (DiStefano et al., 2009).

7.2.3.8 *Sensitivity analysis for handling missing data*

It was deemed possible that the attrition of participants in these cohort studies between pre-adolescence and mid-adulthood could lead to bias in estimating SEP differentials in behavioural outcomes (Howe et al., 2013b). In consequence, attempts were made to manage missing data in these cohorts. Missing data patterns on SEP indicators in pre-adolescence and mid-adulthood were assessed descriptively and are presented in Appendix 7.7 and Appendix 7.8.

Missing data were handled in the EFA, CFA and SEM using the WLSMV estimator function in Mplus Version 7 (Muthén, 2014). In the absence of covariates in the model, this technique considers all available data for each pair of variables (pairwise deletion) and assumes missingness can be explained by other SEP indicators in the model. This pairwise approach is less restrictive than a missing completely at random assumption but can lead to a loss of information and is slightly more restrictive than FIML (only available with an MLR estimator) which uses all available information under a missing at random assumption (Wang and Wang, 2012).

The inclusion of auxiliary variables that can predict missingness is recommended (Graham, 2003). Therefore, to validate the assumptions made using the pairwise approach (i.e. that missing values can be adequately explained by pairs of variables in the model) a sensitivity analysis was conducted whereby the variance of variables shown to predict missing data in pre-adolescence and mid-adulthood in the NCDS (Atherton et al., 2008; Hawkes and Plewis, 2006) and the BCS70 (Mostafa and Wiggins, 2014) was incorporated into the path model and free to correlate with the SEP constructs in pre-adolescence and mid-adulthood in the path.[13] Subsequent changes to the model estimates and fit statistics were assessed. Appendix 7.9 provides a list of the covariates included in the path models for both men and women in each cohort.

7.3 Results

7.3.1 Descriptive and bivariate analyses

Table 7.1 and Table 7.2, present the descriptive statistics of the analytical sample in each cohort study. There were both similarities and differences in the distribution of pre-adolescence and mid-adulthood SEP across the two cohorts.

In pre-adolescence, participants in the BCS70 appeared to be less materially disadvantaged compared to those in the NCDS. For example, the prevalence of overcrowding (<1 person per room NCDS=32.1%, BCS70=42.1%) and living in council housing (NCDS=34.3%, BCS70=24.6%) was statistically lower in the later-born cohort ($p<0.001$). In terms of the cultural dimension, a higher proportion of mothers and fathers stayed at school past minimum leaving age in the

[13] The Mplus 'AUXILIARY (m)' command could not be used because the outcome was not continuous.

BCS70 compared to the NCDS (NCDS mothers=23.1%, BCS70 mothers=25.9%, p<0.001; NCDS fathers=19.7%, BCS70 fathers=24.1%, p<0.001).

For mid-adulthood material dimension indicators, overall BCS70 participants tended to be more advantaged than NCDS participants. For example, mean income was higher for mid-age participants in the BCS70 than in the NCDS (NCDS mean=195.0 (sd 1090.7), BCS70 mean=335.4 (sd 784), p<0.001), after accounting for inflation.[14] In terms of occupation dimension indicators, the NCDS had a higher proportion of working-class participants compared to the BCS70 (NCDS=39.3%, BCS70=29.3%, p<0.001). Moreover, there was greater access to employer pension schemes in the BCS70 compared to the NCDS (NCDS=42.2%, BCS70=58.1%, p<0.001), particularly amongst women (NCDS=33.0%, BCS70=54.2%, p<0.001). For the cultural dimension, the distribution of qualifications in each cohort indicated increased homogeneity in the BCS70 and the age left full-time education was higher in the BCS70 compared to the NCDS (NCDS mean=17.2 (sd 2.1), BCS70 mean=18.7 (sd 3.9), p<0.001).

[14] NCDS mean income in 1991 was £195, BCS70 mean income in 2004 was £335. After accounting for inflation, £195 in 1991 was equivalent to £249 in 2004 (calculated via http://www.whatsthecost.com/cpi.aspx).

Table 7.1 Descriptive statistics for pre-adolescent SEP indicator variables

Pre-adolescent SEP indicator variables	**Total NCDS N=11,373 (100%)**	**Men NCDS n=5,586 (100%)**	**Women NCDS n=5,787 (100%)**	**Total BCS70 n=9,464 (100%)**	**Men BCS70 n=4,613 (100%)**	**Women BCS70 n=5,033 (100%)**
Housing tenure						
Owner occupied/private rent/tied to occupation/other	5,746 (50.5%)	2,849 (51.0%)	2,897 (50.1%)	6,030 (62.5%)	2,890 (62.7%)	3,140 (62.4%)
Council rented	3,903 (34.3%)	1,893 (33.9%)	2,010 (34.7%)	2,368 (24.6%)	1,111 (24.1%)	1,257 (25.0%)
Missing	1,724 (15.2%)	844 (15.1%)	880 (15.2%)	1,248 (12.9%)	612 (13.3%)	636 (12.6%)
Overcrowding						
<1 person per room	3,649 (32.1%)	1,820 (32.6%)	1,829 (31.6%)	3,870 (40.1%)	1,810 (39.2%)	2,060 (40.9%)
1 person per room	2,423 (21.3%)	1,195 (21.4%)	1,228 (21.2%)	2,288 (23.7%)	1,093 (23.7%)	1,195 (23.7%)
>1 to 1.5 people per room	2,553 (22.5%)	1,222 (21.9%)	1,331 (23.0%)	1,785 (18.5%)	874 (19.0%)	911 (18.1%)
>1.5 to 2 people per room	826 (7.3%)	400 (7.2%)	426 (7.4%)	314 (3.3%)	156 (3.4%)	158 (13.1%)
>2 people per room	197 (1.7%)	102 (1.8%)	95 (1.6%)	109 (1.1%)	46 (1.0%)	63 (1.3%)
Missing	1,726 (15.2%)	847 (15.2%)	878 (15.2%)	1,280 (13.5%)	634 (13.7%)	646 (12.8%)
Free school meals						
No	8,658 (76.1%)	4,275 (76.5%)	4,383 (75.7%)	7,373 (76.4%)	3,551 (77.0%)	3,822 (75.9%)
Yes	876 (7.7%)	410 (7.3%)	466 (8.1%)	1,042 (10.8%)	458 (9.9%)	584 (11.6%)
Missing	1,839 (16.2%)	901 (16.1%)	938 (16.2%)	1,231 (12.8%)	604 (13.1%)	627 (12.5%)
Benefits received						
No benefits	7,245 (63.7%)	3,539 (63.4%)	3,706 (64.0%)	6,053 (62.8%)	2,914 (63.2%)	3,139 (62.4%)
=>1 benefits	725 (6.4%)	348 (6.2%)	377 (6.5%)	818 (8.5%)	353 (7.7%)	465 (9.2%)
Missing	3,403 (29.9%)	1,699 (30.4%)	1,704 (29.4%)	2,775 (28.8%)	1,346 (29.2%)	1,429 (28.4%)
Weekly gross household income (BCS70 only)						
<£35 per week	N/A	N/A	N/A	120 (1.2%)	59 (1.3%)	61 (1.2%)
£35 – £49 per week				307 (3.2%)	137 (3.0%)	170 (3.4%)
£50 – £99 per week				2,200 (22.8%)	1,025 (22.2%)	1,175 (23.4%)
£100 – £149 per week				2,753 (28.5%)	1,322 (28.7%)	1,431 (28.4%)
£150 – £199 per week				1,344 (13.9%)	659 (14.3%)	685 (13.6%)
£200 – £249 per week				539 (5.6%)	250 (5.4%)	289 (5.7%)
>£249 per week				497 (5.2%)	246 (5.3%)	251 (5.0%)
Missing				1,886 (19.6%)	915 (19.8%)	971 (19.3%)
Mothers education						
Stayed past minimum school-leaving age	2,622 (23.1%)	1,288 (24.8%)	1,334 (23.1%)	2,494 (25.9%)	1,177 (25.5%)	1,317 (26.2%)
Did not stay past minimum school-leaving age	7,486 (65.8%)	3,683 (71.0%)	3,803 (65.7%)	5,055 (52.4%)	2,419 (52.4%)	2,636 (52.4%)
Missing	1,265 (11.1%)	615 (11.0%)	650 (11.2%)	2,097 (21.7%)	1,017 (22.1%)	1,080 (21.5%)
Fathers education						
Stayed past minimum school-leaving age	2,237 (19.7%)	1,085 (19.4%)	1,152 (19.9%)	2,325 (24.1%)	1,103 (23.9%)	1,222 (24.3%)
Did not stay past minimum school-leaving age	6,913 (60.8%)	3,412 (61.1%)	3,501 (60.5%)	4,899 (50.8%)	2,358 (51.1%)	2,541 (50.5%)
Missing	2,223 (19.5%)	1,089 (19.5%)	1,134 (19.5%)	2,422 (25.1%)	1,152 (25.0%)	1,270 (25.2%)
Parental interest in education						
Over concerned (NCDS only)	209 (1.8%)	107 (1.9%)	102 (1.8%)	N/A	N/A	N/A
Very interested	3,282 (28.9%)	1,558 (27.9%)	1,724 (29.8%)	3,710 (38.5%)	1,736 (37.6%)	1,974 (39.2%)
Some interest (NCDS) / Moderately interested (BCS70)	3,631 (31.9%)	1,806 (32.3%)	1,825 (31.5%)	2,589 (26.8%)	1,248 (27.1%)	1,341 (26.6%)
Little (NCDS) / Very little interest (BCS70)	1,708 (15.0%)	890 (15.9%)	818 (14.1%)	387 (4.0%)	198 (4.3%)	189 (3.8%)
Uninterested (BCS70)	N/A	N/A	N/A	220 (2.3%)	117 (2.5%)	103 (2.1%)
Missing	2,543 (22.4%)	1,225 (21.9%)	1,318 (22.8%)	2,740 (28.4%)	1,314 (28.5%)	1,426 (28.3%)

Table 7.2 Descriptive statistics for mid-adulthood SEP indicator variables

Mid-adulthood SEP indicator variables	**Total NCDS N=11,373 (100%)**	**Men NCDS n=5,586 (100%)**	**Women NCDS n=5,787 (100%)**	**Total BCS70 n=9,464 (100%)**	**Men BCS70 n=4,613 (100%)**	**Women BCS70 n=5,033 (100%)**
Housing tenure						
Owner occupied/private rent/other	8,849 (77.8%)	4,325 (77.4%)	4,524 (78.2%)	8,551 (88.7%)	4,194 (90.9%)	4,357 (86.6%)
Council rented	1,588 (14.0%)	669 (12.0%)	919 (15.9%)	1,049 (10.9%)	397 (8.6%)	652 (13.0%)
Missing	936 (8.2%)	592 (10.6%)	344 (5.9%)	46 (0.5%)	22 (0.5%)	24 (0.5%)
Overcrowding						
<1 person per room	7,486 (65.8%)	3,714 (66.5%)	3,772 (65.2%)	7,279 (75.5%)	3,534 (76.6%)	3,745 (74.4%)
1 person per room	2,366 (20.8%)	1,153 (20.6%)	1,213 (21.0%)	1,579 (16.4%)	730 (15.8%)	849 (16.9%)
>1 to 1.5 people per room	1,142 (10.0%)	512 (9.2%)	630 (10.9%)	634 (6.6%)	282 (6.11%)	352 (7.0%)
>1.5 people per room	149 (1.31%)	69 (1.2%)	80 (1.4%)	92 (1.0%)	33 (0.8%)	59 (1.2%)
Missing	230 (2.0%)	138 (2.5%)	92 (1.6%)	62 (0.6%)	34 (0.7%)	28 (0.6%)
Car ownership						
Yes	9,658 (84.9%)	4,759 (85.2%)	4,899 (84.7%)	6,934 (71.9%)	3,353 (72.7%)	3,581 (71.2%)
No	1,604 (14.1%)	779 (14.0%)	825 (14.3%)	2,040 (21.2%)	864 (18.7%)	1,176 (23.4%)
Missing	111 (1.0%)	48 (0.9%)	63 (1.09%)	672 (7.0%)	396 (8.6%)	276 (5.5%)
Benefits received						
No benefits						
=>1 benefits	10,160 (89.3%)	5,096 (91.2%)	5,064 (87.5%)	8,929 (92.6%)	4,353 (94.4%)	4,576 (90.9%)
Missing	1,165 (10.2%)	467 (8.4%)	698 (12.1%)	703 (7.3%)	253 (5.5%)	450 (8.9%)
	48 (0.4%)	23 (0.4%)	25 (0.4%)	14 (0.2%)	7 (0.2%)	7 (0.1%)
NS-SEC[a]						
Higher managerial, administrative and professional	3,558 (31.3%)	1,998 (35.8%)	1,560 (27.0%)	4,136 (42.9%)	2,132 (46.2%)	2,004 (39.8%)
Intermediate	2,677 (23.5%)	1,043 (18.7%)	1,634 (28.2%)	1,949 (20.2%)	830 (18.0%)	1,119 (22.2%)
Routine and manual	4,473 (39.3%)	2,328 (41.7%)	2,145 (37.1%)	2,823 (29.3%)	1,515 (32.8%)	1,308 (26.0%)
Missing	665 (5.9%)	217 (3.9%)	448 (7.7%)	738 (7.7%)	136 (3.0%)	602 (12.0%)
Access to employer pension scheme						
Yes	4,797 (42.2%)	2,887 (51.7%)	1,910 (33.0%)	5,604 (58.1%)	2,874 (62.3%)	2,730 (54.2%)
No	2,605 (22.9%)	997 (17.9%)	1,608 (27.8%)	1,377 (14.3%)	715 (15.5%)	662 (13.2%)
Missing	3,971 (34.9%)	1,702 (30.5%)	2,269 (39.2%)	2,665 (27.6%)	1,024 (22.2%)	1,641 (32.6%)
Has joined employer pension scheme (BCS70 only)						
Yes	N/A	N/A	N/A	5,461 (56.6%)	2,920 (63.3%)	2,541 (50.5%)
No				1,554 (16.1%)	680 (14.7%)	874 (17.4%)
Missing				2,631 (27.3%)	1,013 (22.0%)	1,618 (32.2%)
Chance to buy shares (NCDS only)						
Yes	1,841 (16.2%)	1,184 (21.2%)	657 (11.4%)	N/A	N/A	N/A
No	5,561 (48.9%)	2,700 (48.3%)	2,861 (49.4%)			
Missing	3,971 (34.9%)	1,702 (30.5%)	2,269 (39.2%)			

Note: [a] Using social occupational classification 1990 in the NCDS and social occupational classification 2000 in the BCS70.

Table 7.2 Descriptive statistics for mid-adulthood SEP indicator variables

Mid-adulthood SEP indicator variables	**Total NCDS N=11,373 (100%)**	**Men NCDS n=5,586 (100%)**	**Women NCDS n=5,787 (100%)**	**Total BCS70 n=9,464 (100%)**	**Men BCS70 n=4,613 (100%)**	**Women BCS70 n=5,033 (100%)**
Access to company car (NCDS only)						
Yes	1,295 (11.4%)	1,058 (18.9%)	237 (4.1%)	N/A	N/A	N/A
No	6,107 (53.7%)	2,826 (50.6%)	3,281 (56.7%)			
Missing	3,971 (34.9%)	1,702 (30.5%)	2,269 (39.2%)			
Private medical insurance (NCDS only)						
Yes	1,366 (12.0%)	964 (17.3%)	402 (7.0%)	N/A	N/A	N/A
No	6,036 (53.1%)	2,920 (52.3%)	3,116 (53.8%)			
Missing	3,971 (34.9%)	1,702 (30.5%)	2,269 (39.2%)			
Highest held qualification						
No qualifications	1,402 (12.3%)	619 (11.1%)	783 (13.5%)	899 (9.3%)	486 (10.5%)	413 (8.2%)
CSE 2–5/NVQ1	1,386 (12.2%)	607 (10.9%)	779 (13.5%)	1,457 (15.1%)	703 (15.2%)	754 (15.0%)
O Level/ NVQ2	3,803 (33.4%)	1,669 (29.9%)	2,134 (36.9%)	3,173 (32.9%)	1,539 (33.4%)	1,634 (32.5%)
A Level/NVQ3	1,567 (13.8%)	1,000 (17.9%)	567 (9.8%)	884 (9.2%)	417 (9.0%)	467 (9.3%)
Diploma or higher qualification below degree/NVQ4	1,577 (13.9%)	785 (14.1%)	792 (13.7%)	2,605 (27.0%)	1,167 (25.3%)	1,438 (28.6%)
Degree or higher/NVQ5 or 6	1,401 (12.3%)	770 (13.8%)	631 (10.9%)	606 (6.3%)	291 (6.3%)	315 (6.3%)
Missing	237 (2.1%)	136 (2.4%)	101 (1.8%)	22 (0.2%)	10 (0.2%)	12 (0.2%)
	Mean(sd)	**Mean(sd)**	**Mean(sd)**	**Mean(sd)**	**Mean(sd)**	**Mean(sd)**
Age left full-time education (range 14 to 33/34)	17.2 (2.1)	17.2 (2.3)	17.2 (2.1)	18.7 (3.9)	18.6 (3.9)	18.7 (3.9)
Missing	454 (4.0%)	259 (4.6%)	195 (3.4%)	22 (0.2%)	12 (0.3%)	10 (0.2%)
Highest household Cambridge scale [a] (range 10 to 99)	55.3 (14.8)	54.0 (14.6)	56.7 (14.9)	58.5 (13.9)	57.1 (13.9)	59.9 (13.7)
Missing	647 (5.7%)	205 (3.7%)	442 (7.6%)	749 (7.8%)	142 (3.08%)	607 (12.1%)
Weekly net household income adjusted for household size (range £0 to £90,000)	195.0 (1090.7)	209.8 (1516.0)	181.7 (419.1)	335.4 (784.0)	353.0 (946.2)	319.4 (600.0)
Missing	1,683 (14.8%)	937 (16.8%)	746 (12.9%)	1,493 (15.5%)	736 (16.0%)	757 (15.0%)

Note: [a] Using social occupational classification 1990 in the NCDS and social occupational classification 2000 in the BCS70.

Table 7.3 to Table 7.6 present the results of the bivariate analysis using multinomial regression models with the 'Mainstream' cluster as the baseline category. The results show the crude association between each indicator variable and the nominal HRB cluster dependent variable. These models were run to ascertain the statistical significance and direction of the relationship between each SEP indicator and HRB cluster membership.

In both cohorts, nearly all of the associations were statistically significant ($p<0.001$). However, for NCDS women two of the three employee benefits (having a company car and the opportunity to buy shares) were not significantly associated with HRB cluster membership ($p>0.05$).

Table 7.3 Bivariate analyses of pre-adolescent SEP indicator variables and HRB cluster membership in the NCDS

Pre-adolescent SEP indicator variables	**NCDS Men sample N=5,586, RRR (CI)**			**NCDS Women sample N=5,787, RRR (CI)**		
	Mainstream (n=3,818)	**Risky (n=82)**	**Moderate Smokers (n=1,686)**	**Mainstream (n=3,980)**	**Risky (n=515)**	**Moderate Smokers (n=1,292)**
Housing tenure	Ref			Ref		
Owner occupied/private rent/other		1.00	1.00		1.00	1.00
Council rented		2.16 (1.32, 3.55)	1.27 (1.12, 1.44)		2.78 (2.27, 3.41)	1.63 (1.42, 1.87)
Overcrowding	Ref			Ref		
<1 person per room		1.00	1.00		1.00	1.00
1 person per room		1.37 (1.11, 1.70)	1.19 (1.12, 1.26)		1.64 (1.50, 1.80)	1.27 (1.20, 1.36)
>1 to 1.5 people per room						
>1.5 to 2 people per room						
>2 people per room						
Free school meals	Ref			Ref		
No		1.00	1.00		1.00	1.00
Yes		3.36 (1.80, 6.25)	2.00 (1.62, 2.46)		3.55 (2.71, 4.64)	2.18 (1.75, 2.71)
Benefits received	Ref			Ref		
No benefits		1.00	1.00		1.00	1.00
=>1 benefits		2.86 (1.46, 5.60)	1.63 (1.29, 2.05)		3.32 (2.47, 4.47)	1.79 (1.40, 2.29)
Mothers education	Ref			Ref		
Stayed past minimum school-leaving age		1.00	1.00		1.00	1.00
Did not stay past minimum school-leaving age		0.99 (0.59, 1.64)	1.38 (1.20, 1.60)		2.49 (1.89, 3.28)	1.32 (1.13, 1.55)
Fathers education	Ref			Ref		
Stayed past minimum school-leaving age		1.00	1.00		1.00	1.00
Did not stay past minimum school-leaving age		1.39 (0.73, 2.63)	1.34 (1.15, 1.57)		3.42 (2.48, 4.74)	1.55 (1.31, 1.84)
Parental interest in education	Ref			Ref		
Over concerned/Very interested		1.00	1.00		1.00	1.00
Some interest		1.88 (1.35, 2.60)	1.42 (1.31, 1.55)		2.52 (2.18, 2.92)	1.49 (1.35, 1.64)
Little interest						

Table 7.4 Bivariate analyses of mid-adulthood SEP indicator variables and HRB cluster membership in the NCDS

Mid-adulthood SEP indicator variables	NCDS Men sample N=5,586, RRR (CI)			NCDS Women sample N=5,787, RRR (CI)		
	Mainstream (n=3,818)	**Risky (n=82)**	**Moderate Smokers (n=1,686)**	**Mainstream (n=3,980)**	**Risky (n=515)**	**Moderate Smokers (n=1,292)**
Housing tenure Owner occupied/private rent/other Council rented	Ref	 1.00 11.12 (6.55, 18.87)	 1.00 3.50 (2.95, 4.15)	Ref	 1.00 7.98 (6.48, 9.82)	 1.00 2.98 (2.52, 3.52)
Overcrowding <1 person per room 1 person per room >1 to 1.5 people per room >1.5 people per room	Ref	 1.00 1.70 (1.31, 2.22)	 1.00 1.41 (1.30, 1.52)	Ref	 1.00 1.66 (1.49, 1.85)	 1.00 1.21 (1.11, 1.31)
Car ownership Yes No	Ref	 1.00 3.8 (2.36, 6.10)	 1.00 1.66 (1.42, 1.94)	Ref	 1.00 3.78 (3.05, 4.68)	 1.00 2.09 (1.76, 2.48)
Benefits received No benefits =>1 benefits	Ref	 1.00 3.95 (2.22, 7.04)	 1.00 2.97 (2.44, 3.61)	Ref	 1.00 4.55 (3.65, 5.69)	 1.00 2.37 (1.98, 2.85)
NS-SEC Higher managerial, administrative and professional Intermediate Routine and manual	Ref	 1.00 2.46 (1.80, 3.34)	 1.00 1.70 (1.59, 1.82)	Ref	 1.00 2.21 (1.92, 2.53)	 1.00 1.39 (1.28, 1.51)
Access to employer pension scheme Yes No	Ref	 1.00 0.38 (0.21, 0.68)	 1.00 0.54 (0.46, 0.63)	Ref	 1.00 0.40 (0.31, 0.52)	 1.00 0.65 (0.55, 0.76)
Chance to buy shares Yes No	Ref	 1.00 1.12 (0.61, 2.05)	 1.00 0.82 (0.70, 0.96)	Ref	 1.00 0.83 (0.59, 1.16)	 1.00 0.88 (0.71, 1.09)
Access to company car Yes No	Ref	 1.00 0.50 (0.23, 1.07)	 1.00 0.71 (0.60, 0.84)	Ref	 1.00 1.01 (0.61, 1.68)	 1.00 1.15 (0.84, 1.58)
Private medical insurance Yes No	Ref	 1.00 0.40 (0.17, 0.95)	 1.00 0.71 (0.60, 0.84)	Ref	 1.00 0.69 (0.45, 1.07)	 1.00 0.74 (0.57, 0.97)
Highest held qualification No qualifications CSE 2–5/NVQ1 O Level/ NVQ2 A Level/NVQ3 Diploma or higher qualification below degree/NVQ4 Degree or higher/NVQ5 or 6	Ref	 1.00 0.52 (0.44, 0.62)	 1.00 0.70 (0.67, 0.73)	Ref	 1.00 0.49 (0.45, 0.53)	 1.00 0.74 (0.71, 0.77)
Age left full-time education	Ref	0.68 (0.53, 0.86)	0.85 (0.82, 0.88)	Ref	0.48 (0.42, 0.55)	0.83 (0.79, 0.86)
Highest household Cambridge scale	Ref	0.94 (0.92, 0.95)	0.97 (0.96, 0.97)	Ref	0.94 (0.93, 0.95)	0.97 (0.96, 0.97)
Household income (log transformed)	Ref	0.56 (0.41, 0.76)	0.71 (0.65, 0.78)	Ref	0.57 (0.51, 0.63)	0.75 (0.69, 0.81)

Table 7.5 Bivariate analyses of mid-adulthood SEP indicator variables and HRB cluster membership in the BCS70

Pre-adolescent SEP indicator variables	**BCS70 Men sample N=4,613, RRR (CI)**			**BCS70 Women sample N=5,033, RRR (CI)**		
	Mainstream (n=3,410)	**Risky (n=79)**	**Moderate Smokers (n=1,124)**	**Mainstream (n=3,866)**	**Risky (n=183)**	**Moderate Smokers (n=984)**
Housing tenure	Ref			Ref		
Owner occupied/private rent/other		1.00	1.00		1.00	1.00
Council rented		2.52 (1.57, 4.06)	1.76 (1.51, 2.06)		4.39 (3.14, 6.14)	2.26 (1.93, 2.65)
Overcrowding	Ref			Ref		
<1 person per room		1.00	1.00		1.00	1.00
1 person per room		1.72 (1.39, 2.13)	1.18 (1.09, 1.27)		1.44 (1.24, 1.68)	1.21 (1.12, 1.30)
>1 to 1.5 people per room						
>1.5 to 2 people per room						
>2 people per room						
Free school meals	Ref			Ref		
No		1.00	1.00		1.00	1.00
Yes		3.93 (2.33, 6.65)	1.62 (1.31, 2.00)		3.11 (2.13, 4.53)	1.88 (1.54, 2.30)
Benefits received	Ref			Ref		
No benefits		1.00	1.00		1.00	1.00
=>1 benefits		2.80 (1.49, 5.27)	1.62 (1.27, 2.06)		3.18 (2.13, 4.75)	1.51 (1.20, 1.90)
Weekly gross household income (BCS70 only)	Ref			Ref		
<£35 per week		1.00	1.00		1.00	1.00
£35 – £49 per week		0.64 (0.51, 0.79)	0.90 (0.85, 0.96)		0.67 (0.58, 0.78)	0.85 (0.80, 0.91)
£50 – £99 per week						
£100 – £149 per week						
£150 – £199 per week						
£200 – £249 per week						
>£249 per week						
Mothers education	Ref			Ref		
Stayed past minimum school-leaving age		1.00	1.00		1.00	1.00
Did not stay past minimum school-leaving age		2.79 (1.41, 5.51)	1.43 (1.21, 1.69)		3.21 (1.99, 5.19)	1.73 (1.44, 2.07)
Fathers education	Ref			Ref		
Stayed past minimum school-leaving age		1.00	1.00		1.00	1.00
Did not stay past minimum school-leaving age		3.23 (1.52, 6.83)	1.32 (1.11, 1.57)		3.10 (1.87, 5.15)	1.71 (1.41, 2.07)
Parental interest in education	Ref			Ref		
Very/Moderately interested		1.00	1.00		1.00	1.00
Very little interest		2.36 (1.78, 3.13)	1.48 (1.34, 1.64)		2.09 (1.70, 2.56)	1.49 (1.34, 1.66)
Uninterested						

Table 7.6 Bivariate analyses of mid-adulthood SEP indicator variables and HRB cluster membership in the BCS70

Mid-adulthood SEP indicator variables	**BCS70 Men sample N=4,613, RRR (CI)**			**BCS70 Women sample N=5,033, RRR (CI)**		
	Mainstream (n=3,410)	**Risky (n=79)**	**Moderate Smokers (n=1,124)**	**Mainstream (n=3,866)**	**Risky (n=183)**	**Moderate Smokers (n=984)**
Housing tenure Owner occupied/private rent/other Council rented	Ref	 1.00 9.07 (5.54, 14.86)	 1.00 3.95 (3.18, 4.90)	Ref	 1.00 8.55 (6.26, 11.74)	 1.00 3.89 (3.24, 4.68)
Overcrowding <1 person per room 1 person per room >1 to 1.5 people per room >1.5 people per room	Ref	 1.00 2.57 (2.00, 3.29)	 1.00 1.50 (1.35, 1.66)	Ref	 1.00 1.94 (1.63, 2.32)	 1.00 1.59 (1.44, 1.75)
Unemployment Employed/sick or disabled/homemaker/student/other Unemployed	Ref	 1.00 3.98 (1.67, 9.45)	 1.00 1.97 (1.34, 2.89)	Ref	 1.00 2.65 (1.12, 6.27)	 1.00 1.53 (0.90, 2.62)
Car ownership Yes No	Ref	 1.00 2.09 (1.29, 3.40)	 1.00 1.46 (1.24, 1.73)	Ref	 1.00 2.73 (2.00, 3.71)	 1.00 2.02 (1.73, 2.37)
Benefits received No benefits =>1 benefits	Ref	 1.00 4.67 (2.46, 8.84)	 1.00 2.97 (2.29, 3.86)	Ref	 1.00 6.09 (4.29, 8.64)	 1.00 3.20 (2.58, 3.95)
NS-SEC Higher managerial, administrative and professional Intermediate Routine and manual	Ref	 1.00 2.82 (2.08, 3.82)	 1.00 1.76 (1.62, 1.90)	Ref	 1.00 2.56 (2.04, 3.21)	 1.00 1.65 (1.51, 1.81)
Access to employer pension scheme Yes No *Missing*	Ref	 1.00 2.86 (1.68, 4.86)	 1.00 1.58 (1.31, 1.90)	Ref	 1.00 2.78 (1.79, 4.29)	 1.00 1.51 (1.23, 1.85)
Has joined employer pension scheme (BCS70 only) Yes No *Missing*	Ref	 1.00 5.35 (3.19, 8.97)	 1.00 2.49 (2.07, 2.99)	Ref	 1.00 3.64 (2.39, 5.55)	 1.00 2.11 (1.75, 2.53)
Highest held qualification No qualifications CSE 2–5/NVQ1 O Level/ NVQ2 A Level/NVQ3 Diploma or higher qualification below degree/NVQ4 Degree or higher/NVQ5 or 6	Ref	 1.00 0.50 (0.41, 0.60)	 1.00 0.69 (0.66, 0.73)	Ref	 1.00 0.49 (0.43, 0.55)	 1.00 0.69 (0.65, 0.73)
Age left full-time education	Ref	0.76 (0.67, 0.86)	0.91 (0.89, 0.92)	Ref	0.80 (0.74, 0.86)	0.92 (0.90, 0.94)
Highest household Cambridge scale	Ref	0.93 (0.91, 0.94)	0.96 (0.66, 0.73)	Ref	0.93 (0.91, 0.94)	0.96 (0.95, 0.96)
Household income (log transformed)	Ref	0.55 (0.42, 0.73)	0.65 (0.59, 0.72)	Ref	0.57 (0.47, 0.70)	0.70 (0.64, 0.77)

7.3.2 Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) to derive uni-dimensional constructs of SEP in pre-adolescence and mid-adulthood

This section presents the results of the EFA and CFA models, used to derive two uni-dimensional constructs of SEP (i.e. one in pre-adolescence and one in mid-adulthood).

EFA model estimates suggested the indicators were adequate measures of each uni-dimensional construct of SEP (see Appendix 7.10 and Appendix 7.11). This was evidenced by statistically significant factor loadings ($p<0.05$) the majority being moderate in strength (>0.32) and eigenvalues >1.

For both cohorts, the majority of the indicator variables were deemed to make a moderate (factor loading >0.32) and significant ($p<0.05$) contribution to their respective uni-dimensional construct of SEP. Within each cohort, most indicators contributed at least moderately to the SEP construct (>0.32) although some indicators were found to be weaker (<0.32). However, all of the indicators were retained because they were statistically significant ($p<0.05$).

Table 7.7 presents the standardised factor loadings for indicator variables in the NCDS CFA models and Figure 7.3 and Figure 7.4 are diagrams of the NCDS CFA models. In the NCDS, modification indices suggested six correlations between pairs of indicator variables (see note 2 under Table 7.7), all of which were theoretically plausible and their addition improved the model fit substantially. The CFA model in the NCDS was found to be a good fit for men (CFI = 0.939, RMSEA = 0.046 (95% CI = 0.044, 0.047), RMSEA close fit test p value = 1.000) and women (CFI = 0.946, RMSEA = 0.046 (95% CI = 0.045, 0.048), RMSEA close fit test p value = 1.000). The Pearson r correlation between the pre-adolescent SEP and mid-adulthood SEP constructs was 0.73 ($p<0.001$) for men and 0.75 ($p<0.001$) for women. These correlations indicate that the constructs are distinct from one another given that they are below the 0.85 threshold which is commonly used to reject discriminant validity (Kline, 2011).

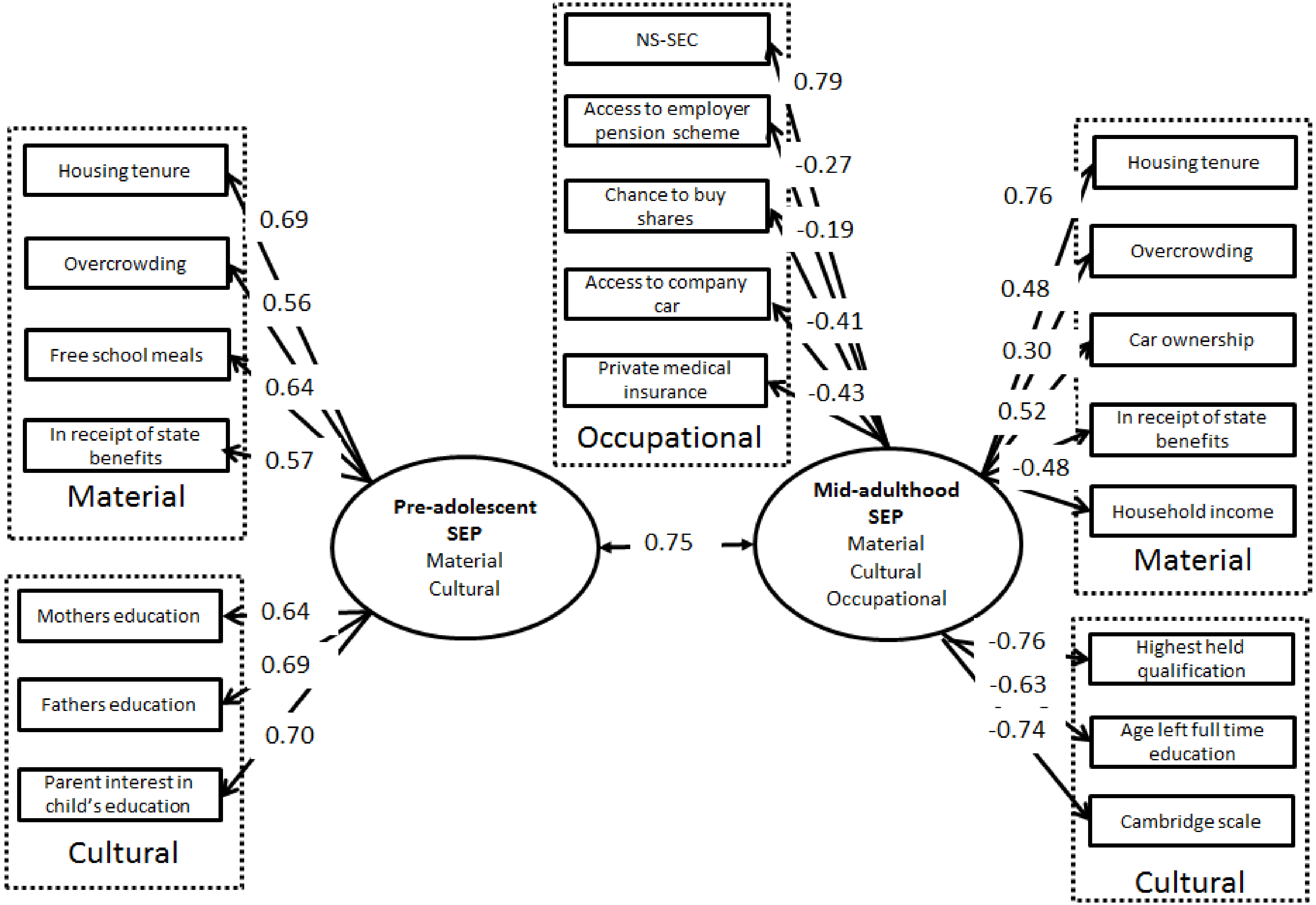
Table 7.7 Estimates from CFA models incorporating both pre-adolescent and mid-adulthood SEP indicator variables in the NCDS

Pre-adolescent Indicator Variables	**NCDS Men**		**NCDS Women**	
	Sample 2 Estimate(S.E)	**Total Sample Estimate(S.E)**	**Sample 2 Estimate(S.E)**	**Total Sample Estimate(S.E)**
Housing tenure	0.69 (0.02)	0.69 (0.02)	0.66 (0.02)	0.67 (0.02)
Overcrowding	0.57 (0.02)	0.56 (0.02)	0.57 (0.02)	0.58 (0.02)
Free school meals	0.62 (0.04)	0.64 (0.03)	0.58 (0.03)	0.57 (0.04)
Benefits received	0.57 (0.04)	0.57 (0.03)	0.61 (0.03)	0.54 (0.04)
Mothers education	0.67 (0.03)	0.64 (0.02)	0.69 (0.02)	0.65 (0.03)
Fathers education	0.67 (0.03)	0.69 (0.02)	0.70 (0.02)	0.71 (0.02)
Parental interest in education	0.70 (0.02)	0.70 (0.02)	0.69 (0.02)	0.69 (0.02)
Mid-adulthood Indicator Variables				
Housing tenure	0.80 (0.02)	0.76 (0.02)	0.80 (0.02)	0.78 (0.01)
Overcrowding	0.48 (0.04)	0.48 (0.03)	0.44 (0.03)	0.45 (0.02)
Car ownership	0.28 (0.03)	0.30 (0.03)	0.43 (0.03)	0.45 (0.02)
Benefits received	0.49 (0.03)	0.52 (0.03)	0.63 (0.03)	0.61 (0.01)
Household income	-0.46 (0.02)	-0.48 (0.01)	-0.54 (0.01)	-0.53 (0.01)
NS-SEC	0.79 (0.01)	0.79 (0.01)	0.74 (0.01)	0.75 (0.01)
Access to employer pension scheme	-0.26 (0.03)	-0.27 (0.03)	-0.52 (0.03)	-0.49 (0.02)
Chance to buy shares	-0.19 (0.03)	-0.19 (0.02)	0.02 (0.04)[†]	-0.01 (0.03)[†]
Access to company car	-0.43 (0.03)	-0.41 (0.03)	-0.25 (0.04)	-0.28 (0.03)
Private medical insurance	-0.45 (0.03)	-0.43 (0.03)	-0.21 (0.04)	-0.23 (0.03)
Highest held qualification	-0.75 (0.01)	-0.76 (0.01)	-0.81 (0.01)	-0.81 (0.01)
Age left full-time education	-0.61 (0.01)	-0.63 (0.01)	-0.68 (0.01)	-0.66 (0.01)
Highest household Cambridge scale	-0.72 (0.01)	-0.74 (0.01)	-0.71 (0.01)	-0.72 (0.01)

Note 1: Sample 2=sample split CFA models ran in one half (EFA models ran in sample 1). Standardised factor loadings. All loadings statistically significant ($p<0.001$), except superscripts †=$p>0.10$.

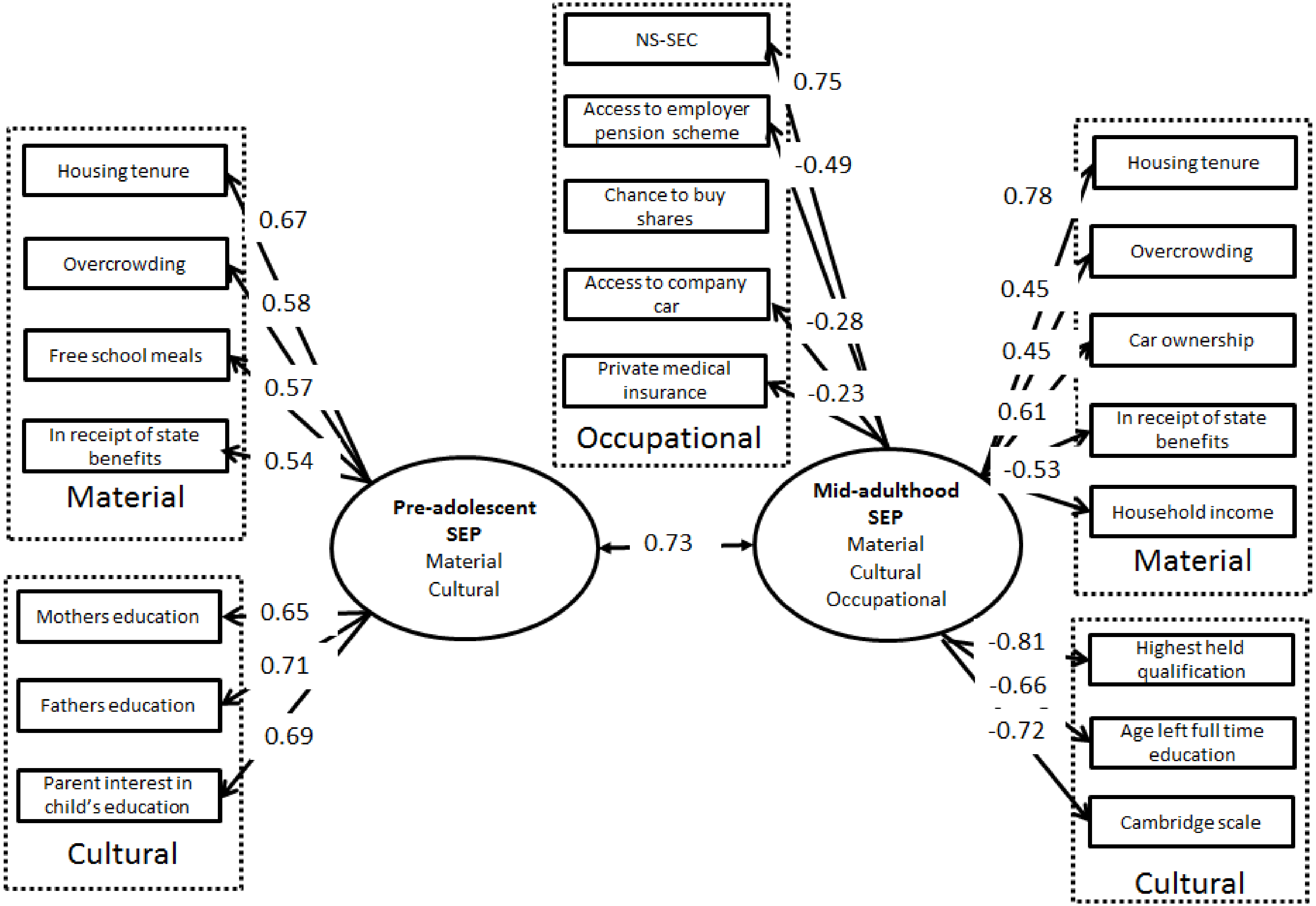
Note 2: Six correlations between indicator measurement errors included in the model based on modification indices 1) NS-SEC and highest household Cambridge scale; 2) Access to a company car and employer offering private medical insurance; 3) household income and employer offering private medical insurance; 4) the chance to buy shares in company and employer offering a pension scheme; 5) being in receipt of benefits and car ownership; 6) number of years in education and highest held qualification.

Figure 7.3 Estimates from CFA models incorporating both pre-adolescent and mid-adulthood SEP indicator variables in the NCDS (men)



Note: Five correlations between indicator measurement errors (listed in note 2 under table 9) not indicated for parsimony. One headed arrows between SEP latent constructs (oval) and observed indicator variables (rectangles) are statistically significant (p<0.001) standardised factor loadings. Two headed arrow between the SEP latent constructs is a Pearson r correlation (p<0.001).

Figure 7.4 Estimates from CFA models incorporating both pre-adolescent and mid-adulthood SEP indicator variables in the NCDS (women)



Note: Five correlations between indicator measurement errors (listed in note 2 under table 9) not indicated for parsimony. One headed arrows between SEP latent constructs (oval) and observed indicator variables (rectangles) are statistically significant (p<0.001) standardised factor loadings. Two headed arrow between the SEP latent constructs is a Pearson r correlation (p<0.001).

Table 7.8 presents the standardised factor loadings for indicator variables in the BCS70 CFA models and Figure 7.5 and Figure 7.6 are diagrams of the BCS70 CFA models. In the BCS70 CFA model, five correlations (see note 2 under Table 7.8) between pairs of indicator variables were included in the model, on the basis of the modification indices and theory. The addition of these correlations led to a weakened factor loading of mid-adulthood SEP on car ownership (0.1) amongst men, although it remained significant (<0.001) and was therefore retained. Moreover, adding these correlations improved model fit substantially for both men (CFI = 0.935, RMSEA = 0.049 (95% CI = 0.047, 0.051), RMSEA close fit test p value = 0.838) and women (CFI = 0.940, RMSEA = 0.048 (95% CI = 0.046, 0.050), RMSEA close fit test p value = 0.949). The Pearson r correlation between the pre-adolescent SEP and mid-adulthood SEP constructs was 0.67 (p<0.001) for men and 0.63 (p<0.001) for women, which is below the 0.85, indicating discriminant validity (Kline, 2011).

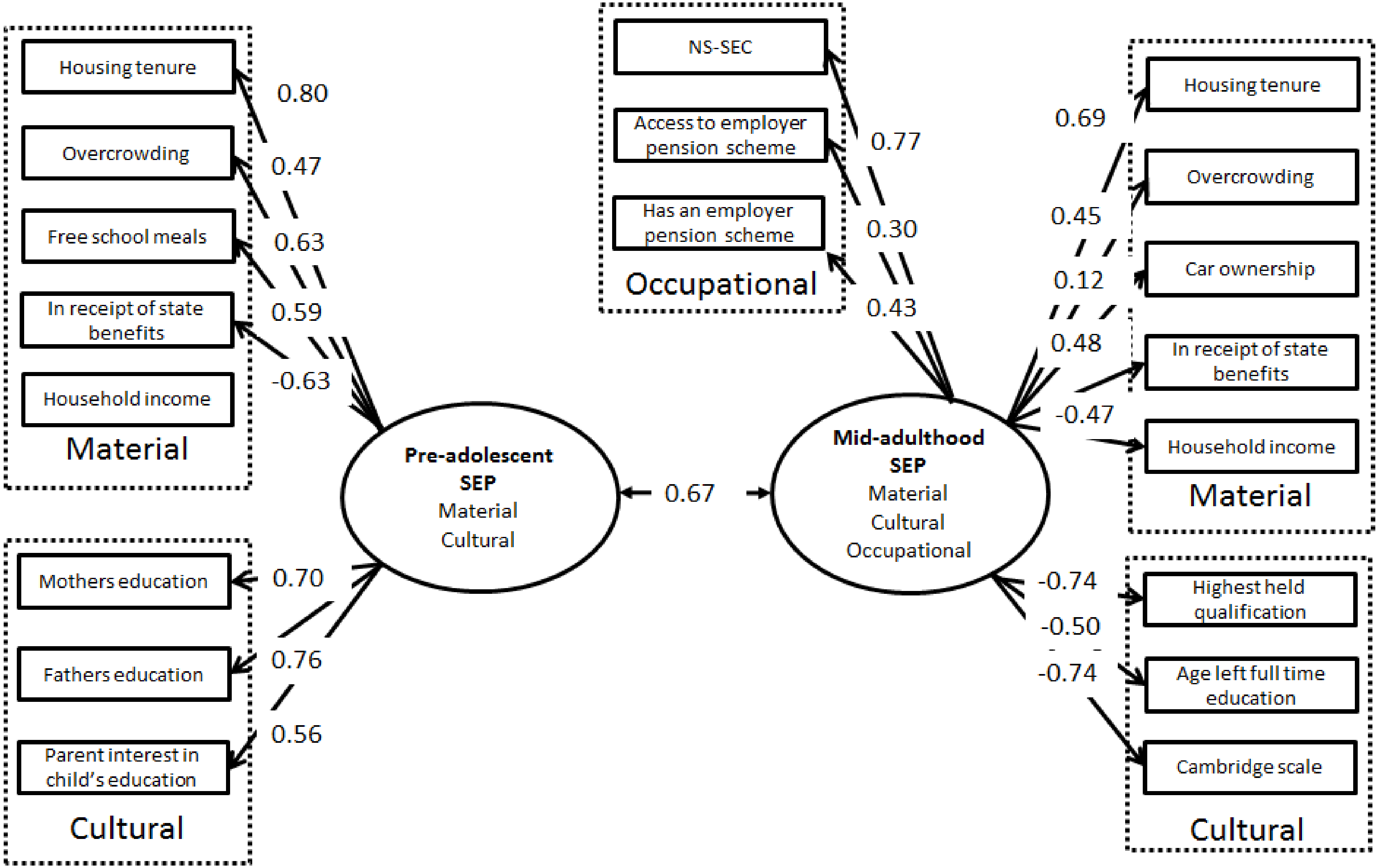
Table 7.8 Estimates from CFA models incorporating both pre-adolescent and mid-adulthood SEP indicator variables in the BCS70

Pre-adolescent Indicator Variables	**BCS70 Men**		**BCS70 Women**	
	Sample 2 Estimate(S.E)	**Total Sample Estimate(S.E)**	**Sample 2 Estimate(S.E)**	**Total Sample Estimate(S.E)**
Housing tenure	0.79 (0.02)	0.80 (0.02)	0.77 (0.02)	0.77 (0.01)
Overcrowding	0.48 (0.03)	0.47 (0.02)	0.48 (0.03)	0.46 (0.2)
Free school meals	0.69 (0.03)	0.63 (0.02)	0.69 (0.03)	0.68 (0.02)
Benefits received	0.58 (0.04)	0.59 (0.03)	0.63 (0.03)	0.65 (0.02)
Household income	-0.63 (0.02)	-0.63 (0.01)	-0.71 (0.02)	-0.70 (0.01)
Mothers education	0.70 (0.02)	0.70 (0.02)	0.67 (0.02)	0.65 (0.01)
Fathers education	0.76 (0.02)	0.76 (0.02)	0.71 (0.02)	0.71 (0.02)
Parental interest in education	0.57 (0.03)	0.56 (0.02)	0.53 (0.03)	0.56 (0.02)
Mid-adulthood Indicator Variables				
Housing tenure	0.69 (0.03)	0.69 (0.02)	0.71 (0.02)	0.73 (0.02)
Overcrowding	0.45 (0.04)	0.45 (0.03)	0.45 (0.04)	0.48 (0.02)
Car ownership	0.12 (0.04)†	0.12 (0.02)	0.29 (0.03)	0.32 (0.02)
Benefits received	0.42 (0.04)	0.48 (0.03)	0.60 (0.03)	0.61 (0.02)
Household income	-0.48 (0.01)	-0.47 (0.01)	-0.54 (0.02)	-0.53 (0.01)
NS-SEC	0.76 (0.02)	0.77 (0.01)	0.80 (0.02)	0.78 (0.01)
Access to employer pension scheme	0.34 (0.04)	0.30 (0.03)	0.37 (0.04)	0.36 (0.03)
Has an employer pension scheme	0.40 (0.03)	0.43 (0.02)	0.54 (0.03)	0.55 (0.02)
Highest held qualification	-0.72 (0.01)	-0.74 (0.01)	-0.73 (0.02)	-0.72 (0.01)
Age left full-time education	-0.53 (0.01)	-0.50 (0.01)	-0.40 (0.02)	-0.41 (0.01)
Highest household Cambridge scale	-0.71 (0.02)	-0.74 (0.01)	-0.72 (0.02)	-0.72 (0.01)

Note 1: Sample 2=sample split CFA models ran in one half (EFA models ran in sample 1).
Standardised factor loadings. All loadings statistically significant ($p<0.001$) except superscripts †=p=0.001.

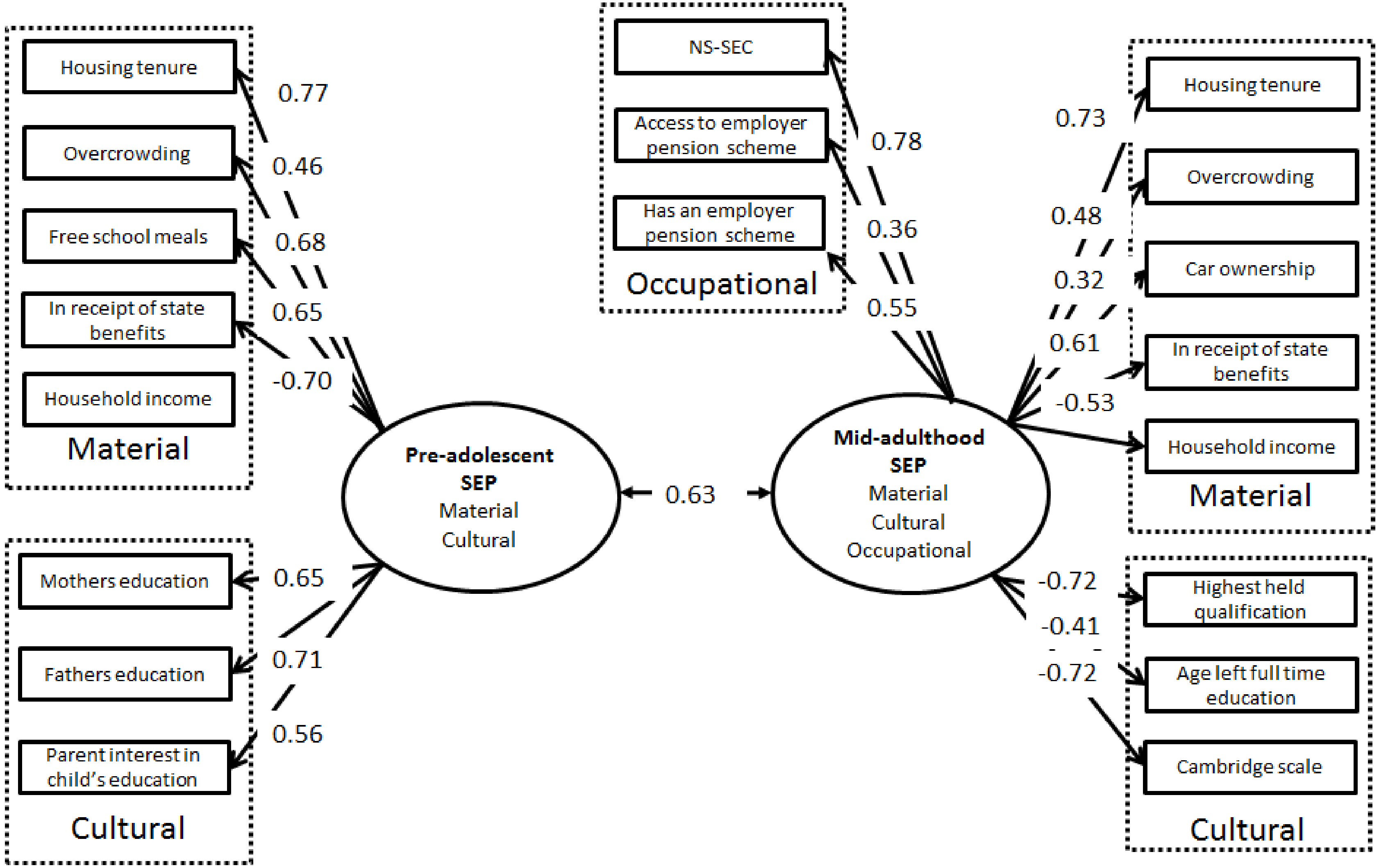
Note 2: Five correlations between indicator measurement errors included in the model based on modification indices 1) employer offering a pension scheme and cohort member joining a pension scheme; 2) receiving benefits in childhood and receiving free school meals; 3) living in social housing in adulthood and receiving benefits in adulthood; 4) NS-SEC and highest household Cambridge scale; 5) number of years in education and highest held qualification.

Figure 7.5 Estimates from CFA models incorporating both pre-adolescent and mid-adulthood SEP indicator variables in the BCS70 (men)



Note: Five correlations between indicator measurement errors (listed in note 2 under table 12) not indicated for parsimony. One headed arrows between SEP latent constructs (oval) and observed indicator variables (rectangles) are statistically significant (p<0.001) standardised factor loadings. Two headed arrow between the SEP latent constructs is a Pearson r correlation (p<0.001).

Figure 7.6 Estimates from CFA models incorporating both pre-adolescent and mid-adulthood SEP indicator variables in the BCS70 (women)



Note: Five correlations between indicator measurement errors (listed in note 2 under table 12) not indicated for parsimony. One headed arrows between SEP latent constructs (oval) and observed indicator variables (rectangles) are statistically significant (p<0.001) standardised factor loadings. Two headed arrow between the SEP latent constructs is a Pearson r correlation (p<0.001).

The product of each CFA model was two uni-dimensional latent constructs, derived separately by cohort and gender. These constructs were conceived to parsimoniously capture social circumstances in pre-adolescence and mid-adulthood, incorporating information pertaining to different dimensions of SEP. In both cohorts, a higher score on each construct indicated more disadvantaged SEP. By default, Mplus Version 7 (Muthén, 2014) fixes the factor mean at zero and estimates the variance. It is possible for the distributions of the estimated constructs (i.e. factor scores) to differ from the latent factors (Muthén, 2012) but in this instance these were also considered to be normal, following their examination in the TECH4 output of Mplus (Muthén, 2012) and in Stata version 14 (StataCorp, 2014) (see Appendix 7.12 and Appendix 7.13).

Appendix 7.14 presents the descriptive statistics for the relationship between individuals HRBs and SEP constructs in pre-adolescence and mid-adulthood.

7.3.3 Path analysis based on probit logistic regression models

Table 7.9 and Table 7.10 present standardised probit regression coefficients (with and without the predictive missing data covariates in the model) for the direct effect (path 'c*' in Figure 7.1), the indirect effect (path 'a'x'b' in Figure 7.2) and the total effect (denoted as 'c' in Figure 7.2) of pre-adolescent SEP on HRB cluster membership.

Each table refers to the results of a model with an alternative binary outcome. Table 15 refers to the model comparing 'Mainstream' and 'Risky' cluster membership whereas table 16 compares 'Mainstream' and 'Moderate Smokers' cluster membership. A lower alpha threshold (p<0.01 instead of p<0.05) was implemented to determine associations as being statistically significant, given treatment if the HRB cluster variable as observed with 'modal assignment' (Heron et al., 2015).

Model estimates of the path models changed very little following the inclusion of covariates shown to predict missingness in the NCDS (Atherton et al., 2008; Hawkes and Plewis, 2006) and the BCS70 (Mostafa and Wiggins, 2014), outlined in Appendix 7.9. This sensitivity analysis indicated that the pairwise approach employed using the WLSMV estimator in Mplus Version 7 (Muthén, 2014), under the assumption that missing values can be explained by pairs of variables in the model, is valid and that the inclusion of additional covariates to predict missingness did not influence the interpretation of the results.

Figure 7.7 to Figure 7.14 are the diagrams of the path models comparing either 'Risky' and 'Mainstream' cluster membership or 'Moderate Smokers' and 'Mainstream' cluster membership for each cohort and gender subgroup (without the predictive missing data covariates in the model).

For all subgroups, the total effect was significant (p<0.001) when comparing 'Risky' and 'Mainstream' cluster membership (NCDS men=0.25 (95% CI 0.16, 0.34), NCDS women=0.52 (95% CI 0.48, 0.57), BCS70 men=0.41 (95% CI 0.33, 0.50), BCS70 women=0.44 (95% CI 0.37, 0.50)). The total effect was also significant (p<0.001) when comparing 'Moderate Smokers' and 'Mainstream' cluster membership (NCDS men=0.21 (95% CI 0.18, 0.25), NCDS women=0.27 (95% CI 0.23, 0.30), BCS70 men=0.21 (95% CI 0.17, 0.25), BCS70 women=0.27 (95% CI 0.23, 0.31)). The coefficients suggested that a one unit increase in pre-adolescent SEP (=more disadvantaged) increased the probability of membership in the 'Risky' or 'Moderate Smokers' cluster compared to the 'Mainstream' cluster.

For men and women in both cohorts, a significant positive indirect effect (p<0.001) was identified when comparing 'Risky' and 'Mainstream' cluster membership (NCDS men=0.51 (95% CI 0.41, 0.62), NCDS women=0.39 (95% CI 0.33, 0.46), BCS70 men=0.39 (95% CI 0.30, 0.47), BCS70 women=0.36 (95% CI 0.31, 0.42)). The same significant effect (p<0.001) was found when comparing 'Moderate Smokers' and 'Mainstream' cluster membership (NCDS men=0.40 (95% CI 0.35, 0.45), NCDS women=0.25 (95% CI 0.19, 0.30), BCS70 men=0.37 (95% CI 0.33, 0.41), BCS70 women=0.28 (95% CI 0.24, 0.32)). A higher pre-adolescent SEP score (=more disadvantage) increased the probability of membership in the 'Risky' or 'Moderate Smokers' cluster by predicting a higher mid-adulthood SEP score (=more disadvantage). This indicates that the effect of pre-adolescent SEP on cluster membership in mid-adulthood can be at least partially explained through its effect on mid-adulthood SEP.

The mediating effect of mid-adulthood SEP largely explained the relationship between pre-adolescent SEP and HRB cluster membership amongst women. For women in the BCS70, the direct effect of pre-adolescent SEP on cluster membership was non-significant (p>0.01) when comparing membership of the 'Mainstream' cluster with 'Risky' (0.07 (95% CI -0.03, 0.17)) and 'Moderate Smokers' (-0.01 (95% CI -0.07, 0.05)) cluster membership. This indicates that more disadvantaged circumstances in pre-adolescence do not directly influence HRB cluster membership in mid-adulthood. The same non-significant (p>0.01) direct effect was found amongst women in the NCDS when comparing 'Moderate Smokers' and 'Mainstream' cluster membership (0.02 (95% CI -0.06, 0.10)). However, a marginally significant (p=0.02) positive direct effect was found for women in the NCDS when comparing 'Risky' and 'Mainstream' cluster membership (0.13 (95% CI 0.04, 0.22)), indicating that more disadvantaged circumstances in pre-adolescence increases the probability of membership of the ‘Risky’ cluster when compared to ‘Mainstream’ cluster membership and that this is not mediated by mid-adulthood SEP. However, this effect became non-significant (p>0.05) after including covariates shown to predict missingness (0.06 (95% CI -0.04, 0.15)).

Similarly, amongst men in both cohorts the mediating effect of mid-adulthood SEP explained a large proportion of the relationship between pre-adolescent SEP and HRB cluster membership. A non-significant direct effect (p>0.01) was found for BCS70 men when comparing ‘Risky’ and ‘Mainstream’ cluster membership (0.03 (95% CI (-0.10, 0.16)). However, there was a significant (p<0.001) negative direct effect of pre-adolescent SEP on ‘Moderate Smokers’ comparative to ‘Mainstream’ cluster membership (-0.16 (95% CI (-0.23, -0.10)). For NCDS men there was a significant (p<0.001) negative direct effect when comparing ‘Mainstream’ cluster membership with ‘Risky’ (-0.27 (95% CI -0.44, -0.09) and ‘Moderate Smokers’ (-0.19 (95% CI (-0.26, -0.12)) cluster membership. The negative direction of these direct effects indicates that more disadvantaged circumstances in pre-adolescence increases the probability of membership in the ‘Mainstream’ cluster when compared to ‘Moderate Smokers’ cluster membership (for men in both cohorts) and ‘Risky’ cluster membership (for men in the NCDS) and that this is not mediated by circumstances in mid-adulthood.

It is clear from these results that amongst men the signs of the total, indirect and direct effects are opposing. The direction of the direct effect is negative whilst the total and indirect effect is positive. Mediation models whereby the direct and indirect effects are different signs are known as ‘inconsistent mediation’ models (MacKinnon et al., 2000). In this case, the inclusion of pre-adolescent SEP clarifies the nature of the relationship between mid-adulthood SEP and HRB cluster membership. Further analyses presented in section 7.3.4 of this chapter shows that pre-adolescent SEP strengthens the effect of mid-adulthood SEP on HRB cluster membership.

Table 7.9 Probit regression coefficients for the total, direct and indirect effects of pre-adolescent SEP on HRB cluster membership ('Risky' vs 'Mainstream')

HRB cluster membership	Total effect	Total effect (includes missing covariates)	Direct effect	Direct effect (includes missing covariates)	Indirect effect	Indirect effect (includes missing covariates)	Model fit CFI; RMSEA (95% CI)	Model fit (includes missing covariates) CFI; RMSEA (95% CI)
NCDS Men	0.25 (0.16, 0.34)*	0.26 (0.17, 0.35)*	-0.27 (-0.44, -0.09)**	-0.33 (-0.50, -0.16)**	0.51 (0.41, 0.62)*	0.59 (0.48, 0.70)*	0.935; 0.048 (0.046, 0.049)	0.929; 0.045 (0.044, 0.046)
NCDS Women	0.52 (0.48, 0.57)*	0.52 (0.48, 0.57)*	0.13 (0.04, 0.22)†	0.06 (-0.04, 0.15)	0.39 (0.33, 0.46)*	0.47 (0.40, 0.53)*	0.942; 0.048 (0.047, 0.050)	0.934; 0.043 (0.042, 0.044)
BCS70 Men	0.41 (0.33, 0.50)*	0.41 (0.33, 0.50)*	0.03 (-0.10, 0.16)	0.02 (-0.11, 0.15)	0.39 (0.30, 0.47)*	0.39 (0.31, 0.47)*	0.932; 0.050 (0.048, 0.052)	0.929; 0.046 (0.044, 0.047)
BCS70 Women	0.44 (0.37, 0.50)*	0.44 (0.37, 0.50)*	0.07 (-0.03, 0.17)	0.07 (-0.02, 0.17)	0.36 (0.31, 0.42)*	0.37 (0.31, 0.43)*	0.939; 0.049 (0.047, 0.051)	0.933; 0.046 (0.044, 0.047)

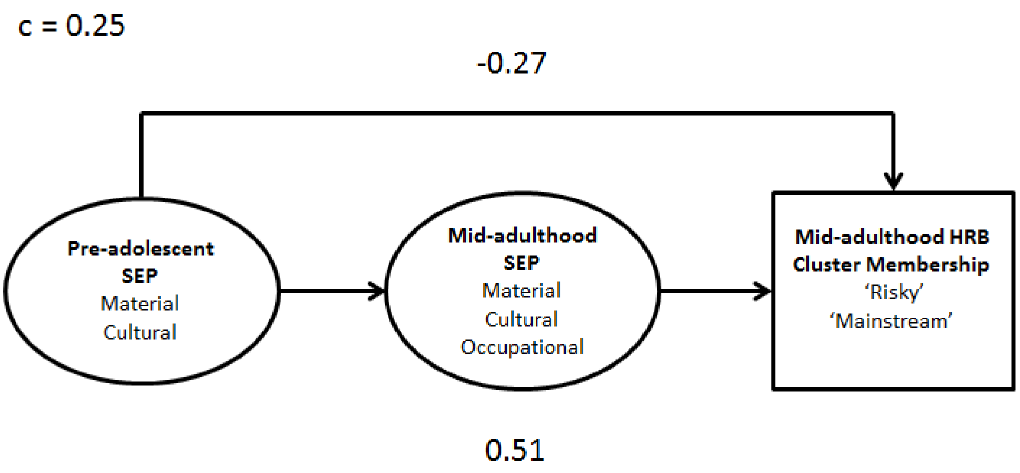
Note: 'Risky' coded 1, 'Mainstream' cluster membership coded 0. One unit increase in pre-adolescent and mid-adulthood SEP = more disadvantaged. Effect estimates are standardised probit regression coefficients and bias-corrected bootstrap 95% Cis (10,000 iterations), *p value ≤ 0.001, **p value ≤ 0.01, †p value=0.02. See figure 7.2 for a diagram of effects; direct effect = path 'a'x'b', indirect effect = path 'c*', total effect = 'c' (c* + ab).

Table 7.10 Probit regression coefficients for the total, direct and indirect effects of pre-adolescent SEP on HRB cluster membership ('Moderate Smokers' vs 'Mainstream')

HRB cluster membership	Total effect	Total effect (includes missing covariates)	Direct effect	Direct effect (includes missing covariates)	Indirect effect	Indirect effect (includes missing covariates)	Model fit CFI; RMSEA (95% CI)	Model fit (includes missing covariates) CFI; RMSEA (95% CI)
NCDS Men	0.21 (0.18, 0.25)*	0.23 (0.19, 0.27)*	-0.19 (-0.26, -0.12)*	-0.17 (-0.24, -0.10)*	0.40 (0.35, 0.45)*	0.40 (0.36, 0.45)*	0.934; 0.049 (0.047, 0.050)	0.927; 0.046 (0.044, 0.047)
NCDS Women	0.27 (0.23, 0.30)*	0.25 (0.21, 0.29)*	0.02 (-0.06, 0.10)	-0.06 (-0.14, 0.02)	0.25 (0.19, 0.30)*	0.31 (0.26, 0.37)*	0.941; 0.048 (0.046, 0.050)	0.934; 0.043 (0.041, 0.044)
BCS70 Men	0.21 (0.17, 0.25)*	0.21 (0.17, 0.25)*	-0.16 (-0.23, -0.10)*	-0.15 (-0.21, -0.09)*	0.37 (0.33, 0.41)*	0.37 (0.32, 0.41)*	0.931; 0.051 (0.049, 0.053)	0.929; 0.046 (0.044, 0.048)
BCS70 Women	0.27 (0.23, 0.31)*	0.28 (0.24, 0.32)*	-0.01 (-0.07, 0.05)	0.01 (-0.06, 0.06)	0.28 (0.24, 0.32)*	0.28 (0.24, 0.31)*	0.938; 0.050 (0.048, 0.051)	0.932; 0.046 (0.045, 0.048)

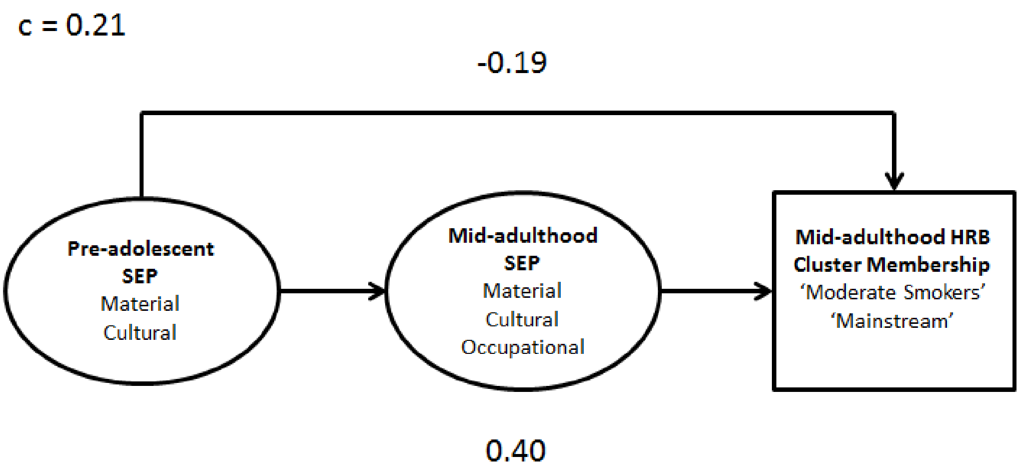
Note: 'Moderate Smokers' coded 1, 'Mainstream' cluster membership coded 0. One unit increase in pre-adolescent and mid-adulthood SEP = more disadvantaged. Effect estimates are standardised probit regression coefficients and bias-corrected bootstrap 95% Cis (10,000 iterations), *p value ≤ 0.001, **p value ≤ 0.01. See figure 7.2 for a diagram of effects: direct effect = path 'a'x'b', indirect effect = path 'c*', total effect = 'c' (c* + ab).

Figure 7.7 Estimates from path model comparing 'Risky' and 'Mainstream' cluster membership in the NCDS (men)



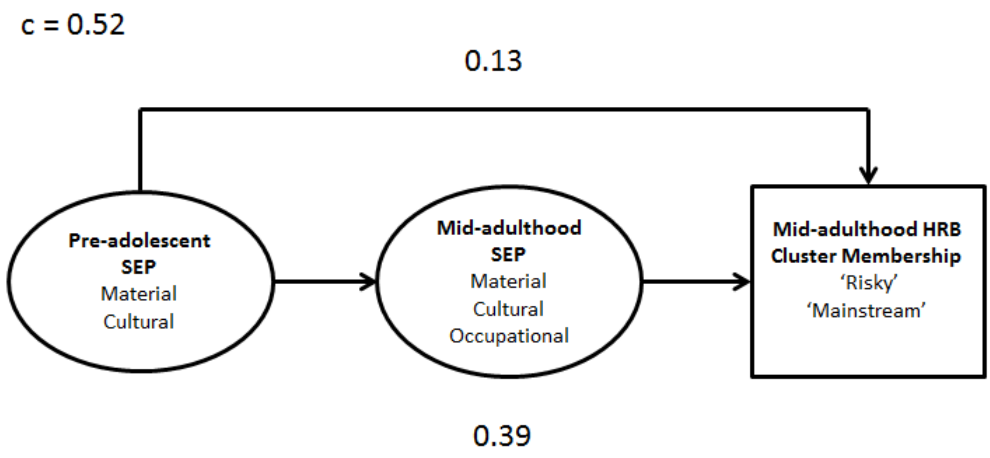
Note: Standardised probit regression coefficients. Solid bold arrows represent tested pathways. Bold arrows represent significant paths (p≤0.01). Dashed arrows represent non-significant paths (p>0.01).

Figure 7.8 Estimates from path model comparing 'Moderate Smokers' and 'Mainstream' cluster membership in the NCDS (men)



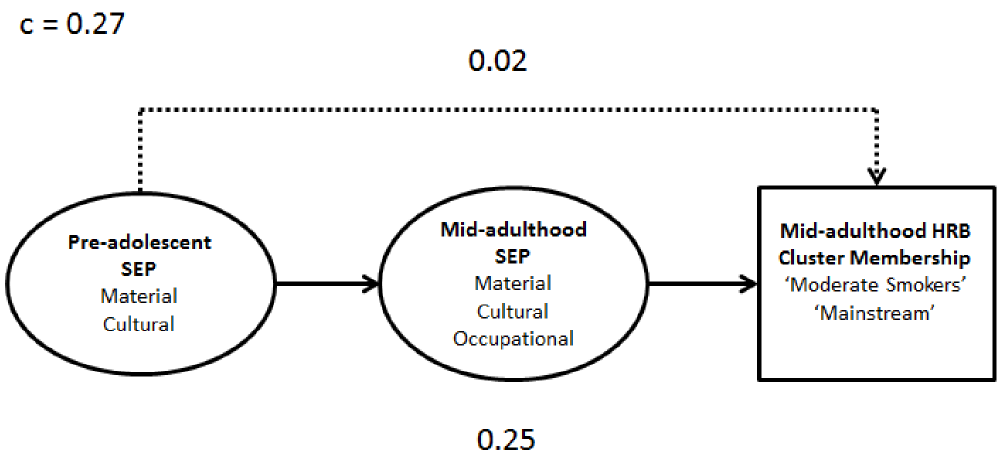
Note: Standardised probit regression coefficients. Solid bold arrows represent tested pathways. Bold arrows represent significant paths (p≤0.01). Dashed arrows represent non-significant paths (p>0.01).

Figure 7.9 Estimates from path model comparing 'Risky' and 'Mainstream' cluster membership in the NCDS (women)



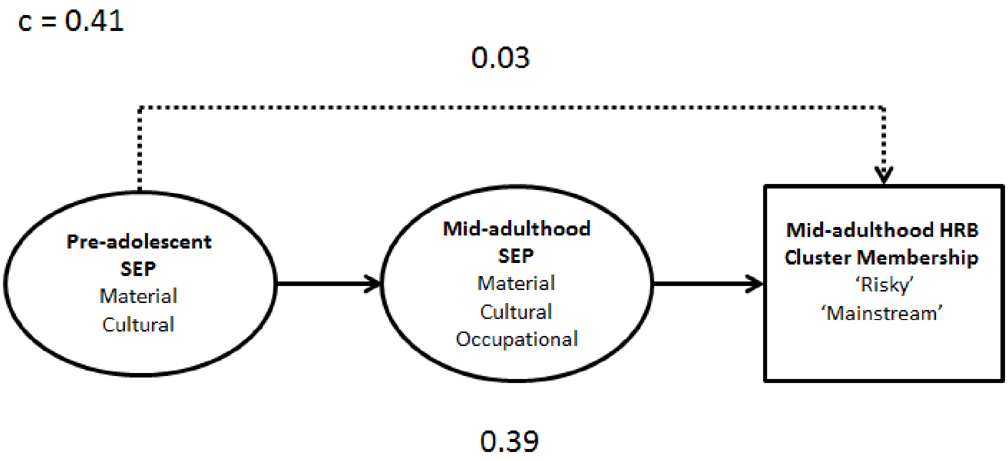
Note: Standardised probit regression coefficients. Solid bold arrows represent tested pathways. Bold arrows represent significant paths ($p \leq 0.01$). Dashed arrows represent non-significant paths ($p > 0.01$).

Figure 7.10 Estimates from path model comparing 'Moderate Smokers' and 'Mainstream' cluster membership in the NCDS (women)



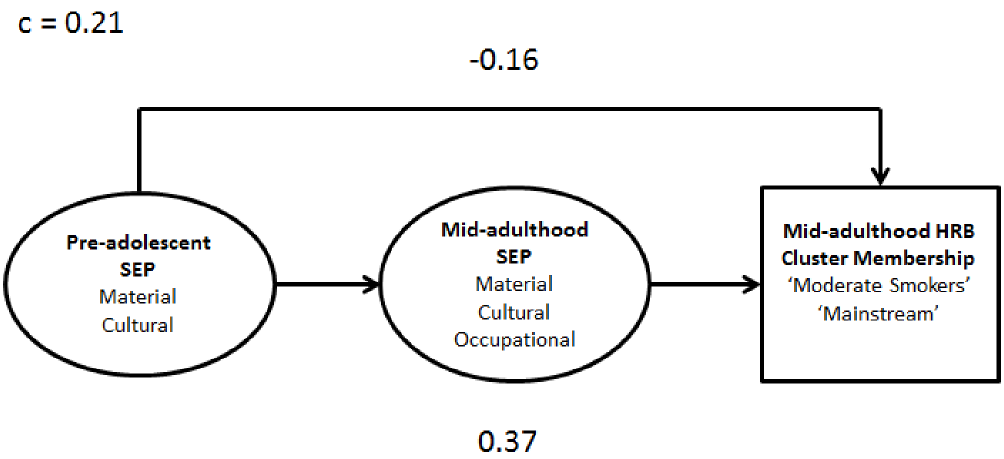
Note: Standardised probit regression coefficients. Solid bold arrows represent tested pathways. Bold arrows represent significant paths ($p \leq 0.01$). Dashed arrows represent non-significant paths ($p > 0.01$).

Figure 7.11 Estimates from path model comparing ‘Risky’ and ‘Mainstream’ cluster membership in the BCS70 (men)



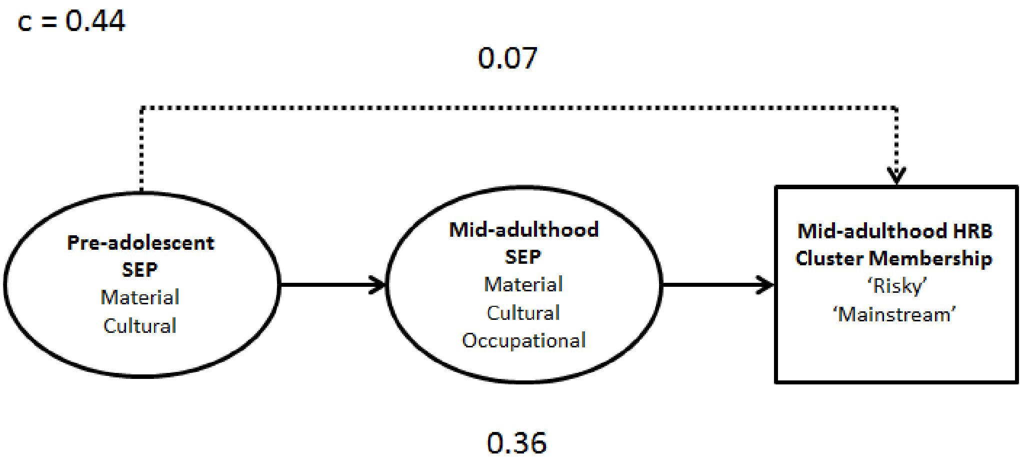
Note: Standardised probit regression coefficients. Solid bold arrows represent tested pathways. Bold arrows represent significant paths (p≤0.01). Dashed arrows represent non-significant paths (p>0.01).

Figure 7.12 Estimates from path model comparing ‘Moderate Smokers’ and ‘Mainstream’ cluster membership in the BCS70 (men)



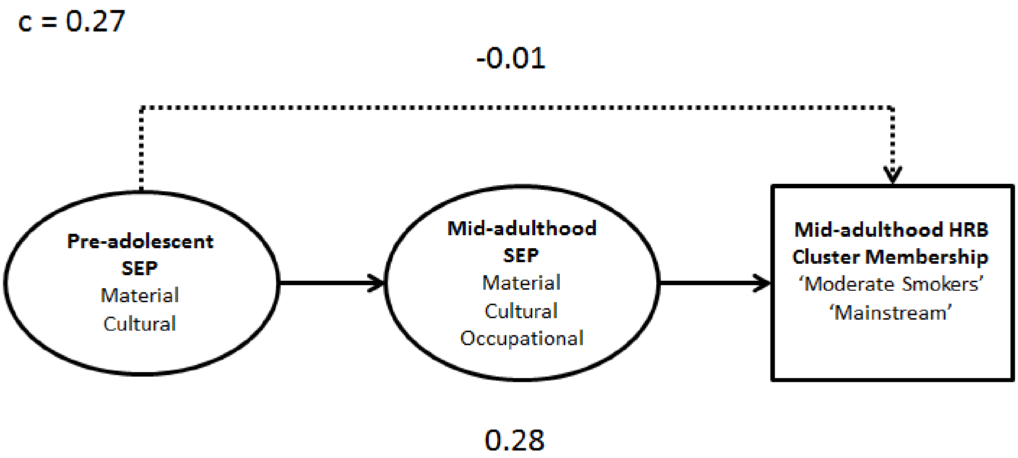
Note: Standardised probit regression coefficients. Solid bold arrows represent tested pathways. Bold arrows represent significant paths (p≤0.01). Dashed arrows represent non-significant paths (p>0.01).

Figure 7.13 Estimates from path model comparing 'Risky' and 'Mainstream' cluster membership in the BCS70 (women)



Note: Standardised probit regression coefficients. Solid bold arrows represent tested pathways. Bold arrows represent significant paths ($p \leq 0.01$). Dashed arrows represent non-significant paths ($p > 0.01$).

Figure 7.14 Estimates from path model comparing 'Moderate Smokers' and 'Mainstream' cluster membership in the BCS70 (women)



Note: Standardised probit regression coefficients. Solid bold arrows represent tested pathways. Bold arrows represent significant paths ($p \leq 0.01$). Dashed arrows represent non-significant paths ($p > 0.01$).

7.3.4 Pre-adolescent SEP strengthens the relationship between mid-adulthood SEP and HRB cluster membership.

In order to show how the inclusion of pre-adolescent SEP strengthens the relationship between mid-adulthood SEP and HRB cluster membership, estimates from three multinomial logistic regression models undertaken in Mplus Version 7 (Muthén, 2014) were compared. HRB cluster membership was included here as a three-category nominal variable. The first two models included each SEP construct (i.e. one in pre-adolescence and one in mid-adulthood) as the sole predictor of HRB cluster membership. The third model included both SEP constructs.

The estimates in Table 7.11 and Table 7.12 provide the results of these multinomial logistic regression models. The results demonstrate that amongst men in both cohorts the logit coefficient for mid-adulthood SEP as a predictor of 'Mainstream' comparative to 'Moderate Smokers' cluster membership increases following the inclusion of pre-adolescent SEP in the model (NCDS men before=1.73 (95% CI 1.58, 1.98), after=2.33 (95% CI 2.05, 2.62); BCS70 men before=1.95 (95% CI 1.76, 2.15), after=2.38 (95% CI 2.10, 2.67)). This suggests that the inclusion of pre-adolescent SEP in these models improves the contribution of mid-adulthood SEP on HRB cluster membership.

Moreover, the results demonstrate that the pre-adolescent SEP coefficient changes from positive to negative when adjusting for mid-adulthood SEP (NCDS men before=0.84 (95% CI 0.74, 0.93), after=-0.46 (95% CI -0.64, -0.27); BCS70 men before=0.61 (95% CI 0.52, 0.71), after=-0.28 (95% CI -0.42, -0.14)). The same effect was found comparing 'Risky' and 'Mainstream' cluster membership amongst men in the NCDS SEP (before=1.60 (95% CI 1.17, 2.03), after=-0.95 (95% CI -1.86, -0.04)).

In substantive terms, this effect implies that differentials in HRB patterns associated with social circumstances in mid-adulthood are strengthened through the accumulation of either advantaged or disadvantaged social circumstances from pre-adolescence.

Table 7.11 Estimates from multinomial logistic regression models with and without pre-adolescent and mid-adulthood SEP in the NCDS

SEP estimated constructs	**NCDS Men sample N=5,586 Logit coefficient (CI)**			**NCDS Women sample N=5,787 Logit coefficient (CI)**		
	Mainstream (n=3, 811)	**Risky (n=96)**	**Moderate Smokers (n=1,679)**	**Mainstream (n=3, 972)**	**Risk (n=561)**	**Moderate Smokers (n=1,254)**
Pre-adolescent SEP only	Ref	1.60 (1.17, 2.03)*	0.84 (0.74, 0.93)*	Ref	2.50 (2.27, 2.75)*	0.72 (0.60, 0.83)*
Mid-adulthood SEP only	Ref	3.67 (2.87, 4.48)*	1.73 (1.58, 1.89)*	Ref	4.56 (4.10, 5.03)*	1.23 (1.05, 1.40)*
Pre-adolescent SEP (adjusting for mid-adulthood SEP)	Ref	-0.95 (-1.86, -0.04)	-0.46 (-0.64, -0.27)*	Ref	0.73 (0.30, 1.16)**	0.12 (-0.10, 0.34)
Mid-adulthood SEP (adjusting for pre-adolescent SEP)	Ref	4.98 (3.49, 6.47)*	2.33 (2.05, 2.62)*	Ref	3.55 (2.77, 4.32)*	1.08 (0.74, 1.41)*

Note: adjustment for classification error in the model, *p value <0.001, **p value <0.05.

Table 7.12 Estimates from multinomial logistic regression models with and without pre-adolescent and mid-adulthood SEP in the BCS70

SEP estimated constructs	**BCS70 Men sample N=4,613 Logit coefficient (CI)**			**BCS70 Women sample N=5,033 Logit coefficient (CI)**		
	Mainstream (n=3,403)	**Risky (n=93)**	**Moderate Smokers (n=1,117)**	**Mainstream (n=3,862)**	**Risky (n=227)**	**Moderate Smokers (n=944)**
Pre-adolescent SEP only	Ref	1.79 (1.44, 2.13)*	0.61 (0.52, 0.71)*	Ref	1.97 (1.65, 2.30)*	0.69 (0.59, 0.80)*
Mid-adulthood SEP only	Ref	5.26 (4.25, 6.27)*	1.95 (1.76, 2.15)*	Ref	5.15 (4.35, 5.95)*	1.86 (1.65, 2.07)*
Pre-adolescent SEP (adjusting for mid-adulthood SEP)	Ref	0.25 (-0.28, 0.78)	-0.28 (-0.42, -0.14)*	Ref	0.50 (-0.02, 1.02)	-0.01 (-0.17, 0.17)
Mid-adulthood SEP (adjusting for pre-adolescent SEP)	Ref	4.79 (3.45, 6.16)*	2.38 (2.10, 2.67)*	Ref	4.38 (3.31, 5.45)*	1.87 (1.54, 2.19)*

Note: adjustment for classification error in the model, *p value <0.001, **p value <0.05.

7.3.5 Sensitivity analysis using multinomial logistic regression models

7.3.5.1 *HRB cluster membership as a nominal outcome*

Table 7.13 and Table 7.14 present the results of a multinomial logistic regression estimating the effect of pre-adolescent SEP and mid-adulthood SEP on HRB cluster membership, with the 'Mainstream' cluster as the reference category. Whilst the results of the multinomial logistic regression models are not directly comparable to the original path models, the results of the sensitivity analyses suggest that the estimated effects from models specifying the dependent variable as nominal are largely similar in direction and magnitude to those from the original path models specifying the dependent variable as two binary variables. Thus, differential treatment of the dependent variable is not considered to alter the overall conclusions made on the basis of the path models.

For example, in all cohort and gender subgroups there was a significant association ($p<0.001$) between mid-adulthood SEP and HRB cluster membership in the multinomial logistic regression models when comparing 'Risky' vs 'Mainstream' cluster membership (NCDS Men=4.54 (95% CI 3.36, 5.73); NCDS Women=2.77 (95% CI 2.26, 3.28); BCS70 Men=4.14 (95% CI 3.10, 5.19); BCS70 Women=3.54 (95% CI 2.87, 4.21)) and when comparing 'Moderate Smokers' and 'Mainstream' cluster membership (NCDS Men=2.35 (95% CI 2.07, 2.63); NCDS Women=1.27 (95% CI 0.97, 1.56); BCS70 Men=2.41 (95% CI 2.13, 2.70); BCS70 Women=2.02 (95% CI 1.72, 2.31)). This mimics path 'b' (see Figure 7.2) of the indirect effect between pre-adolescent SEP and HRB cluster membership which was found to be significant ($p<0.001$) for all subgroups in the original path models (see section 7.3.3).

Amongst men in both cohorts a one unit increase in pre-adolescent SEP (=more disadvantaged) was associated with an increased likelihood of membership of the 'Mainstream' cluster in comparison to the 'Moderate Smokers' clusters (NCDS Men=-0.46, (95% CI -0.64, -0.28); BCS70 MEN=-0.27, (95% CI -0.41, -0.13)). This is consistent with the significant ($p<0.001$) negative direct effect (path c*in Figure 7.2) identified in the original path models (see Figure 7.8 and Figure 7.12). A significant association, albeit weaker ($p<0.05$), was found comparing membership of the 'Mainstream' and 'Risky' clusters, amongst NCDS men (-0.90, 95% CI -1.62, -0.18)). This again, mimics the significant direct effect found in the original path models (see Figure 7.7). For BCS70 men, this association between pre-adolescent SEP and 'Mainstream' vs 'Risky' cluster membership was non-significant (0.17, 95% CI -0.28, 0.62), mirroring the non-significant indirect effect (p>0.01) found in the path analysis (see Figure 7.11).

For women in both cohorts, the non-significant (p>0.05) associations between pre-adolescent SEP and 'Moderate Smokers' cluster membership (NCDS Women=0.18 (95% CI -0.02, 0.37); BCS70 Women=0.03 (95% CI -0.11, 0.18)) mirror the non-significant direct effects (p>0.01) identified amongst women in the original path models (see Figure 7.10 and Figure 7.14).

Amongst NCDS women, more disadvantaged circumstances in pre-adolescence was significantly associated with an increased likelihood of membership of the 'Risky' cluster compared to the 'Mainstream' cluster (0.61 (95% CI 0.31, 0.91)), which is in contrast to the original path model (see Figure 7.9). For BCS70 women, there was no significant association between more disadvantaged circumstances in pre-adolescence when comparing 'Risky' and 'Mainstream' cluster membership (0.33 (95% CI 0.01, 0.65)), which is consistent with the estimates of the original path models (see Figure 7.13).

7.3.5.2 *Modal assignment and classification error*

The multinomial logistic regression results also suggested that modal assignment of participants to their most likely HRB cluster did not substantially alter the results. Comparing estimates from models that adjusted for HRB cluster classification error alongside estimates from models that did not adjust for classification error (see Table 7.13 and Table 7.14) demonstrated that there was little change to the logit coefficients although the 95% confidence intervals were wider when adjusting for HRB cluster classification error into the model.

Table 7.13 The effect of pre-adolescent and mid-adulthood SEP on HRB cluster membership in the NCDS using multinomial logistic regression

SEP estimated constructs	**NCDS Men sample N=5,586 Logit coefficient (CI)**					
	Mainstream≠ (n=3,811)	**Mainstream† (n=3,818)**	**Risky≠ (n=96)**	**Risky† (n=82)**	**Moderate Smokers≠ (n=1,679)**	**Moderate Smokers† (n=1,686)**
Pre-adolescent SEP	Ref	Ref	-0.95 (-1.86, -0.40)	-0.90 (-1.62, -0.18)**	-0.46 (-0.64, -0.27)*	-0.46 (-0.64, -0.28)*
Mid-adulthood SEP	Ref	Ref	4.98 (3.49, 6.47)*	4.54 (3.36, 5.73)*	2.33 (2.05, 2.62)*	2.35 (2.07, 2.63)*
	NCDS Women sample N=5,787 Logit coefficient (CI)					
SEP estimated constructs	**Mainstream≠ (n=3,972)**	**Mainstream† (n=3,980)**	**Risky≠ (n=561)**	**Risky† (n=515)**	**Moderate Smokers≠ (n=1,253)**	**Moderate Smokers† (n=1,292)**
Pre-adolescent SEP	Ref	Ref	0.73 (0.30, 1.16)**	0.61 (0.31, 0.91)*	0.12 (-0.10, 0.34)	0.18 (-0.02, 0.37)
Mid-adulthood SEP	Ref	Ref	3.55 (2.77, 4.32)*	2.77 (2.26, 3.28)*	1.08 (0.74, 1.41)*	1.27 (0.97, 1.56)*

Note: ≠=adjustment for classification error in the model, †=no adjustment for classification error in the model, *p value <0.001, **p value <0.05.

Table 7.14 The effect of pre-adolescent and mid-adulthood SEP on HRB cluster membership in the BCS70 using multinomial logistic regression

SEP estimated constructs	**BCS70 Men sample N=4,613 Logit coefficient (CI)**					
	Mainstream≠ (n=3,404)	**Mainstream† (n=3,410)**	**Risky≠ (n=94)**	**Risky† (n=79)**	**Moderate Smokers≠ (n=1,116)**	**Moderate Smokers† (n=1,292)**
Pre-adolescent SEP	Ref	Ref	0.25 (-0.28, 0.76)	0.17 (-0.28, 0.62)	-0.28 (-0.42, -0.14)*	-0.27 (-0.41, -0.13)**
Mid-adulthood SEP	Ref	Ref	4.79 (3.45, 6.13)*	4.14 (3.10, 5.19)*	2.38 (2.10, 2.67)*	2.41 (2.13, 2.70)*
	BCS70 Women sample N=5,033 Logit coefficient (CI)					
SEP estimated constructs	**Mainstream≠ (n=3,862)**	**Mainstream† (n=3,866)**	**Risky≠ (n=224)**	**Risky† (n=183)**	**Moderate Smokers≠ (n=947)**	**Moderate Smokers† (n=984)**
Pre-adolescent SEP	Ref	Ref	0.50 (-0.02, 1.02)	0.33 (0.01, 0.65)	-0.01 (-0.17, 0.16)	0.03 (-0.11, 0.18)
Mid-adulthood SEP	Ref	Ref	4.38 (3.31, 5.45)*	3.54 (2.87, 4.21)*	1.87 (1.54, 2.19)*	2.02 (1.72, 2.31)*

Note: ≠= adjustment for classification error in the model, †= no adjustment for classification error in the model, *p value <0.001, **p value <0.05.

7.4 Summary of findings

The results of this chapter provide support for some of the hypothesised relationships between pre-adolescence SEP and mid-adulthood HRB cluster membership. Using data from two British birth cohort studies and path analyses the findings showed how more disadvantaged social circumstances in pre-adolescence and mid-adulthood, captured by material, occupational and cultural dimensions of individual SEP, were together associated with membership of HRB clusters characterised by multiple negative HRBs in mid-adulthood.

A significant total effect of disadvantaged pre-adolescent SEP on mid-adulthood HRB cluster membership was found. For men and women in both cohorts more social disadvantage in pre-adolescence was associated with an increased probability of membership of the 'Risky' and 'Moderate Smokers' clusters compared to the 'Mainstream' cluster in mid-adulthood, the latter cluster characterised by HRB patterns that tend to be more beneficial for health (see chapter 6, section 6.4).

The results of this analysis supported the hypothesis of a social pathway through which pre-adolescent SEP influences mid-adulthood HRB cluster membership. For men and women in both cohorts, a substantial proportion of the total effect of pre-adolescent SEP on mid-adulthood HRB occurred via mid-adulthood SEP. Disadvantaged pre-adolescence SEP strongly predicted disadvantaged mid-adulthood SEP which was strongly associated with 'Risky' and 'Moderate Smokers' cluster membership in mid-adulthood.

There was relatively weak evidence to support the hypothesised behavioural pathway, i.e. a direct effect of pre-adolescent SEP on mid-adulthood lifestyles. There was a non-significant ($p \geq 0.01$) direct effect amongst women, although amongst NCDS women the direct effect was marginally significant ($p=0.02$) when comparing 'Risky' and 'Mainstream' cluster membership. Amongst men, the direct effect was found to be significant ($p \leq 0.01$) comparing 'Moderate Smokers' and 'Mainstream' cluster membership in both cohorts and was also significant in the NCDS when comparing 'Risky' and 'Mainstream' cluster membership. However, whilst being significant ($p \leq 0.01$) coefficients demonstrated these direct effects for men and women were half the size of the indirect effects. These results indicate that pre-adolescent SEP is less likely to shape HRB patterns in mid-adulthood in comparison to mid-adulthood SEP. This implies that SEP has a contemporaneous influence on HRB clustering in mid-adulthood, dictating the proximal environment in which people live their lives strongly influencing their lifestyles.

The opposing signs for the direct and indirect effect amongst men in the path models when comparing ‘Risky’ and ‘Moderate Smoker’ against ‘Mainstream’ cluster membership, indicated ‘inconsistent mediation’ (MacKinnon et al, 2000). The sensitivity analyses found that without the inclusion of mid-adulthood SEP in the model the coefficient for the effect of pre-adolescent SEP on HRB cluster membership was positive, instead of negative. The sensitivity analysis also demonstrated that the effect of mid-adulthood SEP on HRB cluster membership was strengthened by the inclusion of pre-adolescent SEP in the model. In substantive terms, this implies that differentials in HRB patterns associated with social circumstances in mid-adulthood are strengthened through the persistence of either advantaged or disadvantaged social circumstances since pre-adolescence.

This suggests that there could be an interaction between pre-adolescent SEP and mid-adulthood SEP. Attempting to disentangle and interpret the negative direct effect in isolation of the indirect effect in these models is therefore not appropriate, given the likely existence of an exposure-mediator interaction (Howe et al., 2016).

The aim of these analyses (outlined fully in chapter 4) was to test the hypothesis of a direct relationship between pre-adolescent SEP and mid-adulthood HRB cluster membership and did not include the exploration of an interaction between pre-adolescent and mid-adulthood SEP mediation effects. In consequence this line of enquiry was not pursued. However, it is acknowledged that future post-doctoral work could employ causally defined mediation methods, such as four-way decomposition analysis (Howe et al., 2016) which can test for interaction and mediation simultaneously. This would be a useful next step in order to isolate the direct and indirect effects, as well as the interaction between them, in these path models.

7.5 Strengths and limitations

A major strength of this work is the valuable contribution it makes to the scientific community, by employing a person-centred approach to investigate behavioural outcomes with a focus on HRB clustering. This takes into consideration how HRBs are dependent on one another rather than viewing them as non-related individual entities. Moreover, by using prospectively collected birth cohort data and adopting a social determinants and lifecourse framework, we have been able to elucidate a 'chain of risk' (Ben-Shlomo and Kuh, 2002: 287), through which social circumstances in pre-adolescence shape lifestyles in mid-age by determining social circumstances in mid-adulthood for men and women in two British cohorts.

This work circumvents some of the methodological limitations identified in the appraisal of existing research fully described in chapter 2 (see section 2.2.1) by using large samples with longitudinal prospectively collected data. The large sample size provides enough statistical power to separate analysis by subgroups, and to some extent, offsets the reduced precision associated with latent variable modelling (Ledgerwood and Shrout, 2011). However, it is acknowledged that the 'Risky' cluster is likely to lack statistical power due to its smaller prevalence (ranging from 2% to 9%), which may partly explain the non-significant ($p \geq 0.01$) direct effects between pre-adolescent SEP and 'Risky' compared to 'Mainstream' cluster membership for BCS70 men and women.

The data reduction techniques used here derived SEP constructs at two points in the lifecourse resulting in well-defined latent variables that were free from measurement error inherent in each SEP indicator (Hoyle, 2012). These constructs were conceived to parsimoniously capture the complex interrelationship between material, occupational and cultural dimensions of SEP.

This study is further strengthened by incorporating these latent variables into path models in order to test hypothesised paths within a conventional mediation framework. This approach is a more powerful tool in comparison to simple regression analysis (Hayes, 2013), by allowing all conceptualised relationships between the pre-adolescent and mid-adulthood SEP constructs and mid-adulthood HRB cluster membership of be estimated simultaneously. Accounting for measurement error by modelling a latent mediator variable is a particular strength of these path models, given that measurement error in the mediating variable can be problematic in estimating indirect effects (Howe et al., 2016; Hoyle and Robinson, 2003).

The prospective longitudinal data ensured temporality which minimises recall bias (Cohen et al., 2010) and improved causal inference between pre-adolescent and mid-adulthood SEP by establishing a temporal ordering of the independent and mediator variables (Hoyle and Robinson, 2003). Moreover, the distance in time between pre-adolescent and mid-adulthood SEP is advantageous by reducing the likelihood of variance in the mediation variable being accounted for by the independent variable, which would weaken the effect of mid-adulthood SEP on HRB cluster membership (Hoyle and Robinson, 2003). Mid-adulthood SEP was assumed to predict HRB cluster membership in these models, whilst this assumption is considered reasonable given the compelling evidence of this causal ordering (Pampel et al., 2010), this could not be explicitly tested in these models.

These path models were employed to test for mediation yet through doing so pre-adolescent SEP was identified as strengthening the relationship between mid-adulthood SEP and HRB cluster membership among men, implying a non-linear joint effect (Kline, 2011). Scholars argue that joint effects of childhood and adulthood SEP on behavioural outcomes are often missed because SEP measures are crude and unable to capture these complex processes (Singhammer and Mittelmark, 2010). Therefore, the detection of non-linear joint effects has been made possible here through the derivation of continuous SEP constructs that incorporate multiple observed indicators of SEP and account for measurement error.

At the same time, the detection of a non-linear joint effect highlights the limitations of using a conventional mediation framework in which to test paths between pre-adolescent SEP, mid-adulthood SEP and HRB clustering. This non-linear joint effect suggested an interaction between pre-adolescent SEP and mid-adulthood SEP on HRB clustering and by using this conventional framework it was not possible to disentangle and interpret the significant direct effect found amongst men. The limited ability of the conventional mediation framework to disentangle direct and indirect effects has been observed elsewhere (Howe et al., 2016).

Moreover, this analysis suggests a social gradient in HRB cluster membership and implies a severity continuum (Stapinski et al., 2016). The analysis found that individuals in the ‘Risky’ cluster are the most socially disadvantage, followed by the ‘Moderate Smokers’ and ‘Mainstream’ clusters respectively. As shown in chapter 6, these clusters are largely driven by smoking status which is strongly socially patterned (Pampel et al., 2010) and therefore it makes sense that overall the numbers of cigarettes smoked across the three HRB clusters largely reflects a social gradient in HRB cluster membership. However, it is acknowledged that non-linear effects on the relationship between SEP and HRB cluster membership, which have not been detected in this analysis, may exist. For example, given evidence of an association between higher levels of weekly alcohol consumption and more advantaged SEP (Bellis et al., 2016), it is possible that some socially advantaged men and women who drink heavily may have been assigned to the 'Risky' cluster.

The direct effect of pre-adolescent SEP on HRB clustering in adulthood was not consistently significant ($p>0.01$) for women across cohorts. It may be that this result was consistently significant only for men because some the SEP indicators, particularly those based on occupation (such as the NS-SEC and Cambridge scale), may capture men's social position better than women's whose ties to the labour market are weaker due to parental responsibilities (Bartley, 2016b).

Whilst the small significant direct effect ($p\leq0.01$) found amongst men is considered negligible in comparison to the indirect effect, the non-significant direct effects ($p\geq0.01$) found amongst women cannot be interpreted as mid-adulthood SEP fully mediating the association between pre-adolescent SEP and mid-adulthood HRB cluster membership. The derivation of latent variables means that they are free from measurement error (Hoyle, 2012). This improves accuracy but reduces precision impeding the ability to conclude that the direct effect is truly zero (Ledgerwood and Shrout, 2011). Moreover, these models focus solely on social influences in pre-adolescence and mid-adulthood on HRB clustering, it is possible that other mediators are involved that are not consider here (Hayes, 2013; Howe et al., 2016). Despite this uncertainty, there is no doubt that the total effect of pre-adolescent SEP on HRB cluster membership was significant ($p<0.001$) in all cohort and gender groups and that pre-adolescent SEP largely exerted its effect on HRB cluster membership via mid-adulthood SEP.

The inability to distinguish the three dimensions of SEP (material, occupational and cultural) from one another in these analyses is a limitation of this work. Methodologically speaking, strong correlations between the three dimensions of SEP were identified, posing a high risk of multicollinearity in subsequent path models (Farrar and Glauber, 1967). Steps to circumvent this problem via higher order CFA modelling (Thompson, 2004) were unsuccessful due to a heavy reliance on binary and ordinal SEP indicator variables in the measurement models which were limited in their ability to capture the continuous distribution of each dimension. Moreover, indicators of cultural and occupational dimensions of SEP in mid-adulthood were found to be highly related due to their similarities in measurement and could not be adequately distinguished from one another due to the inability of including nominal indicator variables in the Mplus CFA models (Muthén, 2014) and therefore collapsing the nominal occupational indicator into three ordinal categories.

Using theoretically driven and precise measures to capture each dimension of SEP, as opposed to proxy measures, is required in order to adequately separate these dimensions from one another and elucidate differences in the strength of associations with adulthood HRBs (Bartley et al., 1999; Stait and Calnan, 2016). It is acknowledged that the measures used to derive each SEP construct serve as proxies in capturing the paths conceptualised to shape HRBs, i.e. physical access, psychosocial stress and social group habitus. Ideally more refined measures would have been used. For example, psychosocial stress due to occupation may be better captured through a specific measure of job strain (Kivimäki et al., 2013) and it may be possible to measure social group habitus by capturing more detailed information on social norms (Reid et al., 2010). The weak direct path between pre-adolescent SEP and HRB cluster membership identified in this chapter may also be a reflection of inadequacies in SEP measurement.

The absence of information during pre-adolescence pertaining to the dietary intake and physical activity of the cohort members, exposure to parental health-damaging HRBs in the household and information on the participants' attitudes and beliefs towards smoking and alcohol consumption is a clear limitation. Access to this information would serve to strengthen the hypothesis of the effect of pre-adolescent familial SEP on embedding HRBs that emerge in mid-adulthood.

Furthermore, in mid-adulthood measures that explicitly capture participants' physical access to health-promoting HRBs, levels of psychosocial stress and cultural attitudes towards HRBs would strengthen the assertion made here of these pathways reinforcing previously embedded HRBs and shaping HRB in mid-adulthood.

Nevertheless, whilst identifying the independent effects of each dimension of SEP on HRB cluster membership would have been of scientific interest, this does not retract from the findings of the path models which are based on a uni-dimensional construct of SEP at each age. It may be argued that the overarching message of this work would remain largely unchanged even if dimensions of SEP had been separate from one another. This assertion is supported by the results of the bivariate analyses (see section 7.3.1) which found consistent associations between observed indicator variables and HRB cluster membership (i.e. being socially advantaged was associated with 'Mainstream' cluster membership). Furthermore, the results of the CFA models (see section 7.3.2) demonstrated that all indicators for each dimension of SEP meaningfully contributed to an underlying 'combined' construct of SEP which in the path models was found to be salient in predicting HRB cluster membership. Thus, these constructs are conceived to effectively capture the complex interrelationship between dimensions of SEP, parsimoniously explained by an underlying construct of SEP at each age.

Care was taken to only include SEP indictors at ages 11 and 33 in the NCDS and ages 10 and 34 in the BCS70 although there were some exceptions (i.e. car ownership was taken at age 30 in the BCS70, information on age cohort member left full-time education was taken from age 42 in the NCDS and information on parent's education was taken from birth and age 7 in the NCDS and age 5 in the BCS70). Despite this, we consider it reasonable to assume that responses to these indicators were very similar during the relatively short periods between the ages of data collection and the ages of interest. Whilst material and cultural dimensions of SEP were included in both pre-adolescence and mid-adulthood, the occupational dimension of SEP was only considered in mid-adulthood. This is because the occupational dimension is considered to influence HRBs through employment relations equating to differential levels of job strain, experienced at the individual level (Sacker et al., 2001). Thus, occupation is not a plausible dimension of SEP in pre-adolescence given that children in Britain at this age do not work and are not influenced by employment relations.

The HRB outcome variable was treated as observed in the path analysis. Participants were 'modally assigned' (Heron et al., 2015) to their most likely HRB cluster, based on the measurement model responsible for the derivation of the HRB clusters (see chapter 6, section 6.3.2.3). This approach does not consider classification error and can lead to an underestimation of standard errors of regression coefficients (Clark and Muthén, 2009). However, this was circumvented by applying a higher alpha threshold ($p<0.01$ instead of $p<0.05$) to determine statistical significance (Clark and Muthén, 2009) in the path models and using probabilities from measurement models with high entropy ($p>0.8$) (Clark and Muthén, 2009).

Additionally, the estimation of bootstrapped 95% confidence intervals with 10,000 iterations in the SEMs, as recommended (Hayes, 2013), considers asymmetry around the coefficients and increases statistical power and may also serve to mitigate the potential underestimation of the standard errors caused by modal assignment.

Furthermore, the sensitivity analysis, comparing estimates from multinomial logistic regression models with and without participant classification error, found that whilst the confidence intervals were wider when HRB cluster classification error was incorporated into the regression models, there was little change to the regression coefficients.

On balance, the modal assignment of participants to their most likely latent class and thus the omission of classification error in the outcome were not considered to alter the conclusions drawn from the path model results.

There was very little difference in estimates from path models with and without the inclusion of additional variables previously found to predict missingness and attrition in these cohort studies (Atherton et al., 2008; Hawkes and Plewis, 2006; Mostafa and Wiggins, 2014). Not only did this indicate that the pairwise approach employed to manage missing data appeared to be robust, it also provided some evidence that attrition between pre-adolescence and mid-adulthood had not introduced bias. Furthermore, whilst it is not possible to say for certain that bias has not occurred, research conducted in other cohorts has found that even with considerable attrition over time (i.e. more than 50%) the effects of SEP differentials on a number of outcomes, including behavioural, did not influence the direction of the results or the overall conclusions (Howe et al., 2013b).

Replicating the path models in two cohorts purposefully similar in design (Ekinsmyth et al., 1992), enabled a meaningful comparison of research findings. Descriptive analyses identified cohort differences in the distribution of the SEP indicator variables. Participants in the later-born cohort tended to be more advantaged than NCDS participants in relation to their income which has been identified elsewhere (Bukodi and Goldthorpe, 2011). Moreover, consistent with previous research (Willson et al., 2007), academic achievement was found to be more homogeneous in the BCS70 compared to the NCDS.

Despite these differences the results of the analyses were found to be similar across the two cohorts. Identifying a similar relationship between social circumstances in pre-adolescence and mid-adulthood and HRB cluster membership not only adds support to the assertion that these HRB clusters are equivalent across the two cohorts and may therefore exist in later-born cohorts (see chapter6, section 6.4), it also indicates that the relationship between HRB cluster membership and social circumstances in pre-adolescence and mid-adulthood may also be generalised to later-born British cohorts.

8 Chapter 8: Transitions in HRB cluster membership over time and the role of mid-adulthood SEP (objective 3)

8.1 Introduction

This chapter addresses the third research objective and associated research hypotheses (see chapter 4). The objective is to ascertain the stability of HRB clustering during mid-adulthood and identify the relationship between HRB clustering stability and social circumstances during mid-adulthood.

It is hypothesised that HRB cluster membership in mid-life will be relatively stable (>70% probability). At the same time, some change in HRB cluster membership will occur (approximately 20% probability). This change will be in a positive direction (i.e. moving to clusters characterised by more positive HRBs than the one they leave behind).

In addition, social circumstances in mid-adulthood are considered likely to influence the extent to which participants move from one HRB cluster to another. More advantaged SEP in mid-adulthood will be associated with transitions to HRB clusters characterised by more positive HRBs than the one they leave behind.

Investigating change in HRB cluster membership during mid-life may elucidate 'natural fluctuations' (Mulder et al., 1998) that provide opportunities to modify lifestyles. Interventions that can enact HRB change in mid-life may have long-term health benefits given evidence that positive change in health lifestyles during mid-age can increase life expectancy (Berstad et al., 2016) and reduce morbidity later in life (Artaud et al., 2016).

To address the third research objective, only information from the NCDS was used due to a lack of equivalence in HRB measurement at age 34 and 42 in the BCS70.

8.2 Methods

8.2.1 Analytical sample

Data were taken from the National Child Development Study (NCDS) when participants were age 33 (data collected in 1991) and age 42 (data collected in 2000). The analytical sample included participants who had information on at least one of four HRBs (smoking, alcohol, diet, physical activity) at either age 33 or 42 (n=50 removed) and information on at least one SEP indicator from either childhood (age 11) or mid-adulthood (age 33) (n=163 removed). This equated to a final analytical sample of 12,784. Appendix 8.1 provides a flowchart outlining the derivation of the final analytical sample.

The data were collected in line with ethical approval procedures at both time points (Shepherd, 2012a). The data were anonymised prior to being used in this doctoral work. Therefore, ethical approval was not required in order to conduct research on this data.

8.2.2 Measures

8.2.2.1 *Outcome: HRB cluster membership*

The four HRBs of interest: Smoking, alcohol, diet and physical activity were measured at ages 33 and 42 in the NCDS using six variables: cigarette smoking, alcohol unit consumption, the frequency of fruit and vegetable consumption, fried food consumption and sweet food consumption and leisure-time physical activity frequency. A full description on the derivation of these measures is given in chapter 5 (see section 5.2.1).

8.2.2.2 *Predictor: Socio-economic position (SEP)*

SEP for this work is a uni-dimensional construct incorporating material, occupational and cultural dimensions. The same indicator variables used to capture participants' SEP in mid-adulthood at age 33 in the work presented in chapter 7 (see section 7.2.2) were used here. A full description on the derivation of these SEP indicator variables is given in chapter 5 (see section 5.2.2).

8.2.3 Statistical analysis

8.2.3.1 *Descriptive*

The distribution of the HRB variables and SEP indicators at each time point (i.e. ages 33 and 42) were examined. Wald test p values were calculated, determining statistically significant differences ($p<0.05$) between individual HRB parameters (i.e. means and proportions) across the two time points.

8.2.3.2 *Confirmatory Factor Analysis (CFA)*

Confirmatory Factor Analysis (CFA) was undertaken to derive a uni-dimensional construct of SEP using material, occupational and cultural indicators at age 33. The same CFA model and observed indicators used to derive the SEP at age 33 construct in chapter 7 were used in this analysis (see section 7.2.3.2). However, unlike the CFA model in the previous chapter, which only used information at age 33, this CFA model incorporated information from participants at age 33 and age 42. This analysis employed the WLSMV estimator in Mplus Version 7 (Muthén, 2014).

Out of the 12,784 participants in the analytical sample, 162 (1.3%) had no data pertaining to the SEP indicators at age 33. As mentioned in chapter 7 (see section 7.2.3.2), in the absence of covariates in the model, the WLSMV estimator manages missing data by considering all available information for each pair of variables.[15] Therefore, information regarding SEP at age 11 (using the same pre-adolescent SEP indicators as those identified in chapter 7 (see section 7.2.2.2)), was also incorporated into the CFA models on the premise that any missing information for SEP at age 33 could be explained by indicators for SEP at age 11 (i.e. during pre-adolescence). By doing so, all participants could be retained in the analytical sample (N=12,784) because they all had information on at least one SEP indicator from either age 11 or age 33. The management of missing data in the CFA models using the WLSMV estimator was considered adequate based on the results of the sensitivity analysis conducted in chapter 7 (see section 7.3.5), which found the method to be robust. Missing data patterns for SEP indicator variables at age 11 and age 33 in the final analytical sample (N=12,784) are presented in Appendix 8.2**.**

Standardised factor loadings, standard errors and p values for the SEP indicators as well as model fit indices from the CFA models were examined. Indicator variables with factor loadings >0.32 and p values <0.05 were considered to contribute moderately to the SEP construct (Tabachnick et al., 2001). The RMSEA and CFI, used in chapter 7 (see section 7.2.3.2) to ascertain model fit, were again employed here. In order to improve model fit, the model modification indices from the CFA models were examined and correlations between pairs of SEP indicators, thought to be theoretically plausible (i.e. such as variables which were similar in wording and measurement), were free to correlate.

8.2.3.3 *Latent Transition Analysis (LTA)*

8.2.3.3.1 Operationalising the LTA model

Latent Transition Analysis (LTA) was undertaken in order to identify transitions between HRB clusters at ages 33 and 42, thus determining stability during mid-life. LTA is a longitudinal extension of Latent Profile Analysis (LPA) (Collins and Lanza, 2010) and was used here to

---

[15] This method is less restrictive than a missing completely at random assumption yet more restrictive than FIML (only available with an MLR estimator) which uses all available information under a missing at random assumption (Wang and Wang, 2012).

methodologically build upon the work from chapter 6, which used LPA to detect three distinct HRB clusters across two cohorts (see section 6.3.2.3).

The LTA model consists of three parameters (Collins and Lanza, 2010). The first is the probability of responses to the observed HRB variables (item response probabilities) given cluster membership at each time point thus assessing the degree of error in each observed indicator in capturing the latent variable. The second is the proportion of participants in each cluster at ages 33 and 42. The third was the probability of transitioning to a cluster at age 42 given cluster membership at age 33. The MLR estimator was used to identify the maximum likelihood solution in Mplus Version 7 (Muthén, 2014). The maximum likelihood estimate was identified by running multiple sets of random starts (4,000 partially, 1,000 fully).

Similar to the LPA models outlined fully in chapter 6 (see section 6.2.3), in all the LTA models the assumption of conditional independence (i.e. any association between observed variables can be explained by the latent variable) (Collins and Lanza, 2010) was relaxed for the three diet variables at each time point because it was considered likely that they would be associated over and above what could be explained by the latent variable. Furthermore, the mean and variance of smoking in the largest cluster was again fixed at zero in order to aid model convergence given the very low prevalence of smoking previously identified in this cluster in chapter 6 (see section 6.2.3).

To manage the modest amount of missing data for each HRB variable (the largest proportion being smoking at age 42; n=523, 4.1%) in the LTA models, the Full Information Maximum Likelihood function in Mplus Version 7 (Muthén, 2014) was employed. This approach uses all available information in the data under a missing at random (MAR) assumption (i.e. that missing data in one variable can be explained by other variables in the model) (Enders, 2010).

8.2.3.3.2 A comparison of exploratory and confirmatory LTA model selection

Two approaches to LTA model selection were considered. The first could be described as exploratory in nature and the second confirmatory. The exploratory approach mimics the analysis undertaken in chapter 6 whereby there is no *a priori* hypothesis on the optimal number of HRB clusters that may exist and can be considered purely data driven. In contrast, the confirmatory approach has substantive underpinnings whereby an *a priori* hypothesis on the number of clusters at each time point is proposed based on prior knowledge. In this case, the hypothesis would be that the three distinct HRB clusters, ‘Risky, ‘Moderate Smokers’ and ‘Mainstream’, identified at age 33 in the NCDS in chapter 6, would once again emerge at age 42.

The exploratory approach is considered given that this analysis is based on a new analytical sample consisting of participants at both age 33 and age 42. The inclusion of information from an additional time point and an increased sample size of 1,411 participants (N=12,784), may lead to an increase in statistical power, allowing for the detection of previously unidentified clustered patterns (Collins and Lanza, 2010). In consequence, whilst the 3-cluster LPA model was considered optimal based upon information at age 33 (see chapter 6), an LTA mode which is based on a new sample that incorporates information at two time points may lead to a different optimal solution.

On the other hand, as mentioned in chapter 5 (see section 5.3) whilst data-driven approaches, such as LPA and LTA, are considered useful tools in elucidating behavioural patterns over time (Collins and Lanza, 2010; Reboussin et al., 1998) they are criticised for being largely subjective (Martinez et al., 1998; McAloney et al., 2013). In consequence, scholars suggest that model selection is based upon interpretability and relevant theory as well as model fit statistics (Nylund et al., 2007; Wang and Wang, 2012) and that exploratory approaches to latent variable modelling that do not take into account prior knowledge may be inefficient (Finch and Bronk, 2011). It could therefore be surmised, given the prior knowledge that now exists as a result of the analyses undertaken in chapter 6, a confirmatory approach to LTA model selection is more appropriate.

To employ an LTA data-driven approach, LPA models were first run for each time point in order to ascertain whether the same number of HRB clusters were found at ages 33 and 42. In this work, LPA models with between 3 and 6 HRB clusters were considered. LPA model selection was determined by examining several model fit statistics including the likelihood ratio chi-squared test, entropy, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), adjusted Bayesian Information Criterion (aBIC) and the Lo-Mendell Rubin likelihood ratio test (LMR LRT) (Collins and Lanza, 2010). Similarly to chapter 6 (see section 6.2.3), emphasis was placed upon the information criterion statistics (a lower number indicates an improvement in model fit) which balance model fit and parsimony. It was not possible to estimate bootstrapped LRM LRTs in these models despite following recommended procedures and using a high number of LRT starts.

The second stage was to run a series of LTA models in order to elucidate whether the additional statistical power did in fact allow for the detection of previously unidentified clusters. Here, the

AIC, BIC and aBIC statistics from LTA models with between 3 and 7 clusters were compared in order to determine the optimal LTA model. Estimates from the LTA models were also examined to ascertain cluster interpretability.

8.2.3.3.3 Measurement invariance over time

In all of the LTA models (i.e. exploratory and confirmatory), measurement invariance across the two time points was assessed. Measurement invariance is imposed in LTA models in order to ensure identification of the LTA model and aid interpretation of class membership transitions (Collins and Lanza, 2010). Scholars recommend that measurement invariance is imposed whenever it can be reasonably assumed (Collins and Lanza, 2010).

Measurement invariance over time was assessed both statistically and substantively. Measurement invariance assumptions were tested statistically using chi-square difference tests to compare the log-likelihood from LTA models allowing parameters to be freely estimated across time points (i.e. rejecting measurement invariance assumption) with those restricted to be equal over time (i.e. imposing measurement invariance).

Measurement invariance was also evaluated in substantive terms, given the very high likelihood of a significant log-likelihood chi-square difference test ($p<0.05$) due to the large sample sizes ($N>2000$); thus, very minor differences will be detected (Meade and Lautenschlager, 2004). The substantive investigation of measurement invariance over time determined whether the nature of the HRB clusters were similar at both time points (Collins and Lanza, 2010). Substantive investigations involved the examination of the observed item indicator means (for continuous variables) and thresholds (for categorical variables) in the LPA models run separately at age 33 and age 42 and model estimates from the best fitting LTA model whereby parameters were freely estimated across time points.

By their very nature, LTA models are considered to be autoregressive, i.e. the measurement model at one time point is related to the measurement model at the following time point (Sutton et al., 2004). This is appropriate in this context, given evidence suggesting that an HRB at one time point strongly predicts that same HRB at a second time point (Laaksonen et al., 2002; Paavola et al., 2004). In consequence, whilst the LPA model estimates were considered informative, emphasis was placed on the estimates from the freely estimated LTA models, which incorporate longitudinal information, in order to ascertain whether measurement invariance could be assumed over time.

8.2.3.4 *Latent Transition Analysis with a covariate*

Once the optimal LTA model had been identified for men and women, a covariate was incorporated into the model representing SEP at age 33. This 'LTA with a covariate' model tested the extent to which SEP at age 33 predicted HRB cluster membership at age 33 and the influence of SEP at age 33 on HRB cluster transitions between ages 33 and 42.

Two models were run, separately for men and women. The first model estimated the crude effect of one unit increase of age 33 SEP (a higher score=more disadvantaged) on the likelihood of HRB cluster membership at age 33 (see Figure 8.1). The second model investigated moderation, estimating the effect of one unit increase of age 33 SEP on the likelihood of transitioning from one HRB cluster at age 33 to a different cluster at age 42 (see Figure 8.2).

The 'LTA with a covariate' model is a type of SEM (Hoyle, 2012) and was operationalised in Mplus Version 7 (Muthén, 2014). The measurement model part of the SEM was the derivation of the HRB clusters at each time point (taken from the preferable LTA model) and the structural part was the transitions between the HRB clusters at ages 33 and 42 and the path between the covariate (i.e. SEP at age 33) and HRB cluster membership at age 33 (see Figure 8.1) and the path between the covariate and transitions in HRB cluster membership over time (Figure 8.2).

Akin to chapter 7 (see section 7.2.3.7), the MLR estimator and 'mixture' type command in Mplus Version 7 (Muthén, 2014) was required to run the LTA models. The latent SEP constructs could not be derived in the same model because they are estimated using CFA models and the WLSMV estimator. Consequently, the 'LTA with a covariate' models incorporated the latent continuous uni-dimensional construct of SEP at age 33 as an observed variable. This was done by estimating and saving the factor scores (a higher score=more disadvantaged) using the regression method (DiStefano et al., 2009). As mentioned in chapter 7 (see section 7.2.3.5), this method of factor score estimation provides a high correlation between the observed estimated factor score and the latent construct (DiStefano et al., 2009).

In these 'LTA with a covariate' models SEP at age 33 was conceptualised as a predictor of HRB clustering and therefore deemed to be independent of the derivation of the HRB clusters (i.e. the LTA measurement model). Given this premise, item means and thresholds for the measurement model were fixed (i.e. using the starting values from the optimal LTA measurement model). This meant that the original measurement model would be retained following the inclusion of the covariate (Collins and Lanza, 2010).

Figure 8.1 A 'LTA with covariate' model testing the effect of SEP at age 33 on HRB cluster membership at age 33

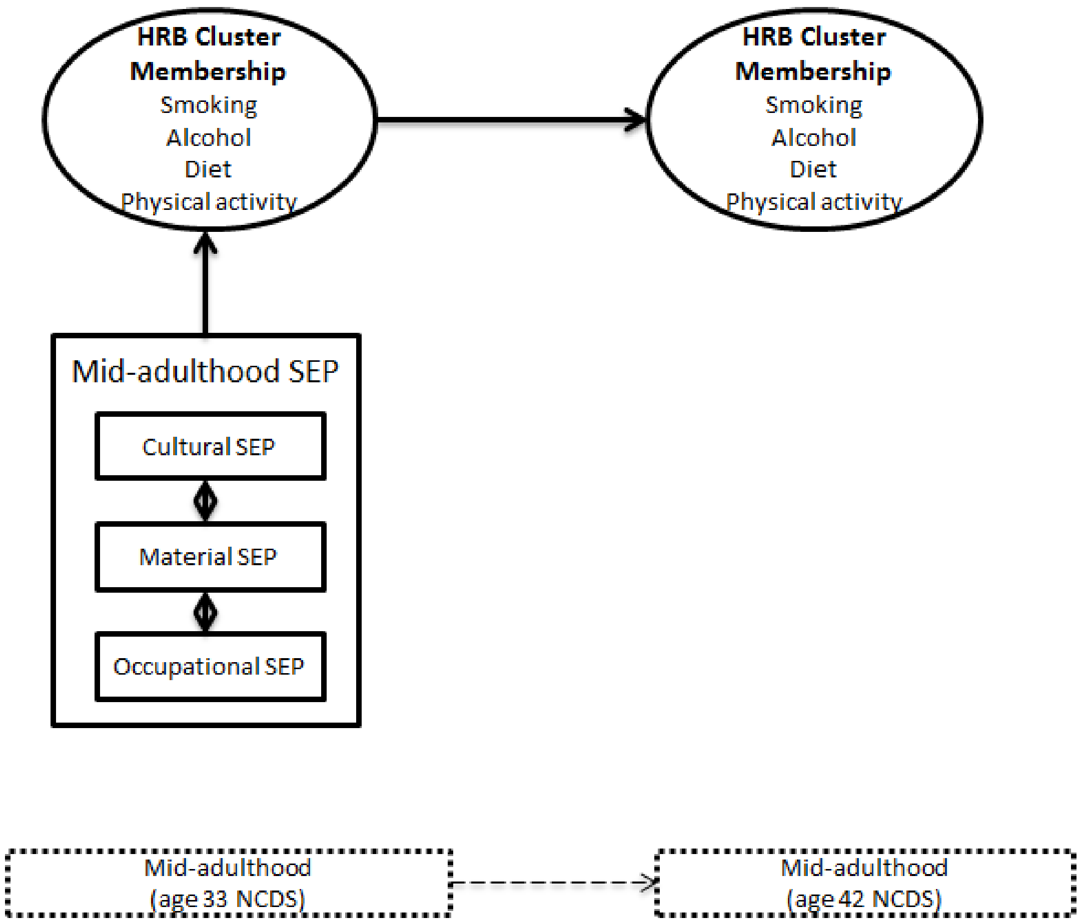
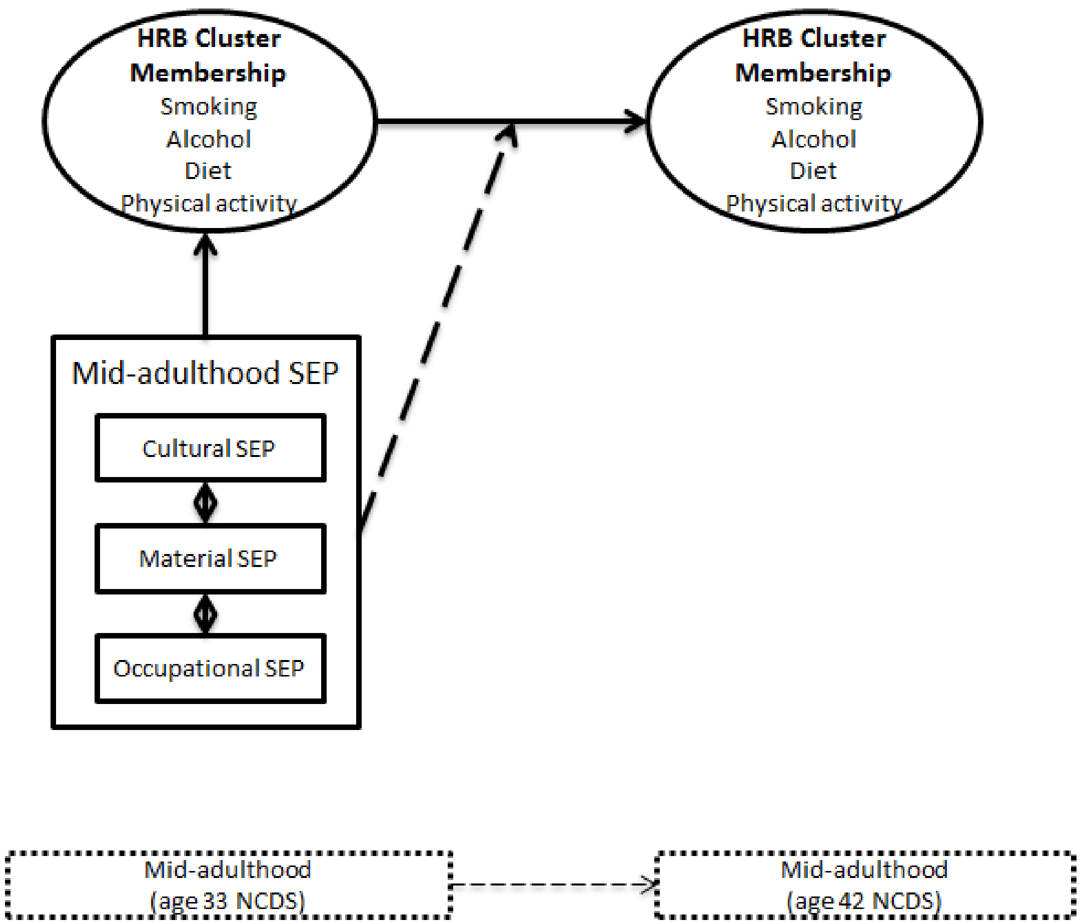


Figure 8.2 A 'LTA with covariate' model testing the effect of SEP at age 33 on transitions in HRB cluster membership between age 33 and age 42.



8.3 Results

8.3.1 Descriptive

Table 8.1, present the descriptive statistics for the HRB variables at age 33 and age 42. Wald tests indicated that smoking and diet behaviours tended to be more beneficial for health at age 42 in comparison to age 33. For both genders, the proportion of smokers was lower at age 42 in comparison to age 33 (age 33=31.5%, age 42=26.9%, $p<0.001$), although the average number of cigarettes smoked per day was not found to be statistically different (age 33 mean=17.35 (sd 8.9), age 42 mean=17.45 (sd 8.5), $p>0.05$). In comparison to age 33, at age 42 there was a higher mean frequency of fruit and vegetable consumption (age 33 mean=4.97 (sd 2.1), age 42 mean=5.48 (sd 2.2), $p<0.001$), a lower mean frequency of fried food consumption ($p<0.001$) and a lower mean frequency of sweet food consumption (age 33 mean=3.14 (sd 1.6), age 42 mean=2.54 (sd 1.1), $p<0.001$). In contrast, physical activity levels and alcohol consumption tended to be less beneficial for health at age 42 in comparison to age 33. For men and women in comparison to age 33, the proportion participating in physical activity ≤3 times a week (age 33=31.4 %, age 42=34.6%) and the proportion consuming alcohol above recommended limits (age 33=17.8%, age 42=27.4%) were both higher at age 42 ($p<0.001$).

Table 8.2 and Table 8.3, present descriptive statistics of the SEP indicators measured at age 11 and age 33 for participants in the final analytical sample (N=12,784). The frequency and distribution of the SEP indicator variables at age 11 and 33, among NCDS participants, were considered to be very similar to those presented in chapter 7 (see section 7.3.1).

Table 8.1 Descriptive statistics for HRB indicator variables at age 33 and age 42

HRB cluster indicator variables†	Total age 33	Men age 33	Women age 33	Total age 42	Men age 42	Women age 42
	Mean (sd)	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**
Number of cigarettes smoked per day [a]	17.35 (8.9)	18.51 (9.52)	16.22 (8.20)	17.45 (8.52)	18.75 (9.02)	16.21 (7.82)
Frequency of fruit and vegetable consumption [b]	4.97 (2.1)	4.40 (2.01)	5.51 (2.00)	5.48 (2.24)	4.67 (2.14)	5.97 (2.22)
Frequency of fried food consumption [b]	3.14 (1.6)	3.58 (1.57)	2.73 (1.46)	2.54 (1.08)	2.79 (1.10)	2.30 (0.96)
Frequency of sweet food consumption [b]	4.59 (2.3)	4.53 (2.28)	4.65 (2.34)	4.37 (2.29)	4.40 (2.27)	4.35 (2.30)
Proportion smoking cigarettes daily	11,330 (100%)	5,560 (100%)	5,770 (100%)	10,717 (100%)	5,266 (100%)	5,451 (100%)
0	7,761 (68.5%)	3,797 (68.3%)	3,964 (68.7%)	7,830 (73.1%)	3,855 (73.2%)	3,975 (72.9%)
1–10	1,031 (9.1%)	458 (8.2%)	573 (9.9%)	790 (7.4%)	333 (6.3%)	457 (8.4%)
11–20	1,896 (16.7%)	912 (16.4%)	984 (17.1%)	1,582 (14.8%)	740 (14.1%)	842 (15.5%)
21+	642 (5.7%)	393 (7.1%)	249 (4.3%)	515 (4.8%)	338 (6.4%)	177 (3.3%)
Frequency of leisure-time physical activity	11,311 (100%)	5,561 (100%)	5,750 (100%)	11,208 (100%)	5,527 (100%)	5,681 (100%)
≤3 times a month	3,548 (31.4%)	1,773 (31.9%)	1,775 (30.9%)	3,877 (34.6%)	1,895 (34.3%)	1,982 (34.9%)
Once a week	2,480 (21.9%)	1,166 (21.0%)	1,314 (22.9%)	2,022 (18.0%)	1,080 (19.5%)	942 (16.6%)
2–3 days a week	2,402 (21.2%)	1,292 (23.2%)	1,110 (19.3%)	2,377 (21.2%)	1,193 (21.6%)	1,184 (20.8%)
4–7 days a week	2,881 (25.5%)	1,330 (23.9%)	1,551 (27.0%)	2,932 (26.2%)	1,359 (24.6%)	1,573 (27.7%)
Alcohol units consumed in the previous week [c]	11,367 (100%)	5,583 (100%)	5,784 (100%)	11,194 (100%)	5,518 (100%)	5,676 (100%)
No units	2,424 (21.3%)	754 (13.5%)	1,670 (28.9%)	2,065 (18.5%)	712 (12.9%)	1,353 (23.8%)
Within limits (≤14 units women, ≤21 units men)	6,920 (60.9%)	3,280 (58.8%)	3,640 (62.9%)	6,062 (54.2%)	2,746 (49.8%)	3,316 (58.4%)
Above limits (≥15 units women, ≥22 units men)	2,023 (17.8%)	1,549 (27.7%)	474 (8.2%)	3,067 (27.4%)	2,060 (37.3%)	1,007 (17.7%)
Missing data	**N=12,784 (100%)**	**N=6,396 (100%)**	**N=6,388 (100%)**	**N=12,784 (100%)**	**N=6,396 (100%)**	**N=6,388 (100%)**
Smoking						
Item missing	104 (0.8%)	57 (0.9%)	47 (0.7%)	523 (4.1%)	280 (4.4%)	177 (2.8%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)	N/A	N/A	N/A
No data at age 42	N/A	N/A	N/A	1,544 (12.1%)	850 (13.3%)	694 (10.9%)
Diet						
Fruit and vegetable consumption item missing	68 (0.5%)	36 (0.6%)	32 (0.5%)	32 (0.3%)	19 (0.3%)	13 (0.2%)
Fried food consumption item missing	79 (0.6%)	40 (0.6%)	39 (0.6%)	41 (0.3%)	29 (0.5%)	12 (0.2%)
Sweet food consumption item missing	80 (0.6%)	38 (0.6%)	42 (0.7%)	31 (0.2%)	19 (0.3%)	12 (0.2%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (9.0%)	N/A	N/A	N/A
No data at age 42	N/A	N/A	N/A	1,544 (12.08%)	850 (13.3%)	694 (10.9%)
Frequency of leisure-time physical activity						
Item missing	123 (1.0%)	56 (0.9%)	67 (1.1%)	32 (0.3%)	19 (0.3%)	13 (0.2%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)	N/A	N/A	N/A
No data at age 42	N/A	N/A	N/A	1,544 (12.08%)	850 (13.3%)	694 (10.9%)
Alcohol units consumed in the previous week						
Item missing	67 (0.5%)	34 (0.5%)	33 (0.5%)	46 (0.4%)	28 (0.4%)	18 (0.3%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)	N/A	N/A	N/A
No data at age 42	N/A	N/A	N/A	1,544 (12.1%)	850 (13.3%)	694 (10.9%)

† Proportions excluding missing data.

a. Range 1–80 age 33. Range 1–70 age 42.

b. A Higher score indicates a higher consumption frequency. Range 0–10. Diet score equivalent (rounded to zero decimal places): ‘never’ [0] ‘occasionally /less than 1 day a week’ [1–2] ‘1–2 days a week’ [3–4] ‘3–6 days a week’ [5–6] once a day’ [7–8] ‘more than once a day’ [9–10].

c. ‘No units’ category includes never drinkers and non-frequent drinkers who report 0 units in the previous week. Frequent drinkers who report 0 units in the previous week have been placed in category ‘within limits’.

N/A = not applicable.

Table 8.2 Descriptive statistics for age 11 SEP indicator variables

Age 11 SEP indicator variables	**Total N=12,784 (100%)**	**Men n=6,396 (100%)**	**Women n=6,388 (100%)**
Housing tenure			
Owner occupied/private rent/tied to occupation/other	6,422 (50.2%)	3,251 (50.8%)	3,171 (49.6%)
Council rented	4,489 (35.1%)	2,214 (34.6%)	2,275 (35.6%)
Item missing	1,052 (8.2%)	531 (8.3%)	521 (8.2%)
No data age 0–16	821 (0.6%)	400 (0.6%)	421 (0.6%)
Overcrowding			
<1 person per room	4,034 (31.6%)	2,033 (31.8%)	2,001 (31.3%)
1 person per room	2,716 (21.3%)	1,364 (21.3%)	1,352 (21.2%)
>1 to 1.5 people per room	2,941 (23.0%)	1,461 (22.8%)	1,480 (23.2%)
>1.5 to 2 people per room	987 (7.7%)	482 (7.5%)	505 (7.9%)
>2 people per room	231 (1.8%)	122 (1.9%)	109 (1.7%)
Item missing	1,054 (8.2%)	534 (8.3%)	520 (8.1%)
No data age 0–16	821 (0.6%)	400 (0.6%)	421 (0.6%)
Free school meals			
No	9,758 (76.3%)	4,916 (76.9%)	4,842 (75.8%)
Yes	1,026 (8.0%)	485 (7.6%)	541 (8.5%)
Item missing	1,179 (9.2%)	595 (9.3%)	584 (6.70%)
No data age 0–16	821 (6.4%)	400 (6.3%)	421 (6.6%)
Benefits received *(unemployment benefit, income support, family income supplement)*			
No benefits	8,088 (63.3%)	4,041 (63.2%)	4,047 (63.4%)
=>1 benefits	839 (6.6%)	411 (6.4%)	428 (6.7%)
Item missing	3,036 (23.7%)	1,544 (24.1%)	1,492 (23.3%)
No data age 0–16	821 (6.4%)	400 (6.3%)	421 (6.6%)
Mothers education			
Stayed past minimum school-leaving age	3,027 (23.7%)	1,517 (23.7%)	1,510 (23.6%)
Did not stay past minimum school-leaving age	8,744 (68.4%)	4,386 (68.6%)	4,358 (68.8%)
Item missing	192 (1.5%)	93 (1.5%)	99 (1.5%)
No data age 0–16	821 (6.4%)	400 (6.3%)	421 (6.6%)
Fathers education			
Stayed past minimum school-leaving age	2,693 (21.1%)	1,335 (20.9%)	1,358 (21.3%)
Did not stay past minimum school-leaving age	8,481 (66.3%)	4,271 (66.8%)	4,210 (65.9%)
Item missing	789 (6.2%)	390 (6.1%)	399 (6.2%)
No data age 0–16	821 (6.4%)	400 (6.3%)	421 (6.6%)
Parental interest in education			
Over concerned	233 (1.8%)	120 (1.9%)	113 (1.8%)
Very interested	3,639 (28.5%)	1,755 (27.4%)	1,884 (29.5%)
Some interest	4,086 (32.0%)	2,069 (32.4%)	2,017 (31.6%)
Little interest	2,031 (15.9%)	1,080 (16.9%)	951 (14.9%)
Item missing	1,974 (15.4%)	972 (15.2%)	1,002 (15.7%)
No data age 0–16	821 (6.4%)	400 (6.3%)	421 (6.6%)

Note: Information for mother's education taken at cohort member's birth and information for father's education taken at age 7.

Table 8.3 Descriptive statistics for age 33 SEP indicator variables

Age 33 SEP indicator variables	**Total N=12,784 (100%)**	**Men n=6,396 (100%)**	**Women n=6,388 (100%)**
Housing tenure			
Owner occupied/private rent/other	8,851 (69.2%)	4,326 (67.6%)	4,525 (70.8%)
Council rented	1,588 (12.4%)	669 (10.5%)	919 (14.4%)
Item missing	995 (7.8%)	622 (9.7%)	373 (5.8%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)
Overcrowding			
<1 person per room	7,488 (58.6%)	3,715 (58.1%)	3,772 (59.1%)
1 person per room	2,366 (18.5%)	1,153 (18.03%)	1,213 (19.0%)
>1 to 1.5 people per room	1,142 (8.9%)	512 (8.0%)	630 (9.9%)
>1.5 people per room	149 (1.2%)	69 (1.1%)	80 (1.3%)
Item missing	289 (2.3%)	168 (2.6%)	121 (1.9%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)
Car ownership			
Yes	9,658 (75.6%)	4,759 (74.4%)	4,899 (76.7%)
No	1,604 (12.6%)	779 (12.2%)	825 (12.9%)
Item missing	172 (1.3%)	79 (1.2%)	93 (1.5%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)
Benefits received *(income support, unemployment benefit, housing benefit)*			
No benefits	10,160 (79.5%)	5,097 (79.7%)	5,064 (79.3%)
=>1 benefits	1,165 (9.1%)	467 (7.3%)	698 (11.0%)
Item missing	108 (0.8%)	53 (0.8%)	55 (0.9%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)
NS-SEC			
Higher managerial, administrative and professional	3,578 (28.0%)	2,008 (31.4%)	1,570 (24.6%)
Intermediate	2,680 (21.0%)	1,044 (16.3%)	1,636 (25.6%)
Routine and manual	4,483 (35.1%)	2,332 (36.5%)	2,151 (33.7%)
Item missing	693 (5.4%)	233 (3.6%)	460 (7.2%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)
Access to employer pension scheme			
No	2,606 (20.4%)	998 (15.6%)	1,608 (25.2%)
Yes	4,797 (37.5%)	2,887 (45.1%)	1,910 (29.9%)
Item missing	4,031 (31.5%)	1,732 (27.7%)	2,299 (36.0%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)
Chance to buy shares (NCDS only)			
No	5,562 (43.5%)	2,701 (42.2%)	2,861 (44.8%)
Yes	1,841 (14.4%)	1,184 (18.5%)	657 (10.3%)
Item missing	4,031 (31.5%)	1,732 (27.1%)	2,299 (36.0%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)

Table 8.3 Descriptive statistics for age 33 SEP indicator variables (continued)

Age 33 SEP indicator variables	**Total N=12,784 (100%)**	**Men n=6,396 (100%)**	**Women n=6,388 (100%)**
Access to company car (NCDS only)			
No	6,107 (47.8%)	2,826 (44.2%)	3,281 (51.4%)
Yes	1,296 (10.1%)	1,059 (16.6%)	237 (3.7%)
Item missing	4,031 (31.5%)	1,732 (27.1%)	2,299 (36.0%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)
Private medical insurance (NCDS only)			
No	6,037 (47.2%)	2,921 (45.7%)	3,116 (48.8%)
Yes	1,366 (10.7%)	964 (15.1%)	402 (6.3%)
Item missing	4,031 (31.5%)	1,732 (27.1%)	2,299 (36.0%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)
Highest held qualification			
No qualifications	1,402 (11.0%)	619 (9.7%)	783 (12.3%)
CSE 2–5/NVQ1	1,387 (10.9%)	607 (9.5%)	780 (12.2%)
O Level/ NVQ2	3,803 (29.8%)	1,669 (26.1%)	2,134 (33.4%)
A Level/NVQ3	1,568 (12.3%)	1,001 (15.7%)	567 (8.9%)
Diploma or higher qualification below degree/NVQ4	1,577 (12.3%)	785 (12.3%)	792 (12.4%)
Degree or higher/NVQ5 or 6	1,401 (11.0%)	770 (12.0%)	631 (9.9%)
Item missing	296 (2.3%)	166 (2.6%)	130 (2.0%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)
	Mean(sd)	**Mean(sd)**	**Mean(sd)**
Age left full-time education (range 14 to 33)	17.2 (2.1)	17.2 (2.2)	17.2 (2.0)
Item missing	464 (3.6%)	267 (4.2%)	197 (3.1%)
Highest household Cambridge scale (range 10 to 99)	55.3 (14.8)	54.0 (14.6)	56.7 (14.9)
Item missing	675 (5.3%)	221 (3.5%)	454 (7.1%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)
Weekly net household income adjusted for household size (range £0 to £90,000)	195.0 (1090.7)	209.8 (1516.0)	181.7 (419.1)
Item missing	1,742 (13.6%)	967 (15.1%)	775 (12.1%)
No data at age 33	1,350 (10.6%)	779 (12.2%)	571 (8.9%)

Note: Information for 'age left full-time education' taken from age 42.

8.3.2 Confirmatory Factor Analysis

Table 8.4 presents the results of the CFA model run separately for men and women, including indicators of SEP at age 11 and at age 33. The results of the CFA model, incorporated six correlations (see note 2 under Table 8.4) between pairs of indicator variables, all of which were theoretically plausible. These six correlations were the same as those used to derive the SEP construct at age 33 in chapter 7 (see section 7.3.2).

The inclusion of these correlations improved the CFA model fit substantially from being below recommended thresholds, i.e. CFI <0.9, RMSEA >0.05 and RMSEA close fit test (p<0.05) (Wang and Wang, 2012), to being adequate, for both men (CFI = 0.936, RMSEA = 0.047 (95% CI = 0.045, 0.048), RMSEA close fit test p value = 1.000) and women (CFI = 0.943, RMSEA = 0.047 (95% CI = 0.046, 0.049), RMSEA close fit test p value = 0.997). All indicators loaded significantly onto their respective SEP construct (p<0.001) with the exception of the 'chance to buy shares' indicator variable amongst women (p>0.05). However, this variable was retained due to being statistically significant (p<0.05) amongst men. The majority of the SEP indicators were deemed to contribute at least moderately to their respective SEP construct (factor loading >0.32).

Each CFA model produced two continuous latent constructs, one capturing SEP at age 11 the other capturing SEP at age 33. These constructs were considered to be uni-dimensional, capturing material and cultural dimensions of SEP at age 11 and material, occupational and cultural dimensions of SEP at age 33. A higher score on each construct indicated more disadvantaged SEP at that age. Each latent construct had a mean of zero and an estimated variance. In order to treat the latent construct for SEP at age 33 as observed in further analyses, the estimated factor scores from the latent constructs for SEP at ages 11 and 33 were saved. Descriptive analyses of these estimated factor scores found them to be normally distributed with a mean very close to 0. The descriptive statistics for the latent constructs for SEP at age 11 and at age 33 and the corresponding observed estimated factor scores are presented in Appendix 8.3.

Table 8.4 Estimates from 'combined' CFA models incorporating both pre-adolescent and mid-adulthood SEP indicator variables in the NCDS

Pre-adolescent Indicator Variables	**NCDS Men (n=6,396)** **Estimate (S.E)**	**NCDS Women (n=6,388)** **Estimate (S.E)**
Housing tenure	0.71 (0.02)	0.69 (0.02)
Overcrowding	0.57 (0.01)	0.58 (0.01)
Free school meals	0.65 (0.02)	0.59 (0.03)
Benefits received	0.57 (0.03)	0.56 (0.03)
Mothers education	0.65 (0.02)	0.65 (0.02)
Fathers education	0.70 (0.02)	0.71 (0.02)
Parental interest in education	0.64 (0.02)	0.63 (0.01)
Mid-adulthood Indicator Variables	**NCDS Men (n=6,396)** **Estimate (S.E)**	**NCDS Women (n=6,388)** **Estimate (S.E)**
Housing tenure	0.78 (0.02)	0.78 (0.02)
Overcrowding	0.48 (0.02)	0.45 (0.02)
Car ownership	0.30 (0.02)	0.45 (0.02)
Benefits received	0.52 (0.02)	0.61 (0.02)
Household income	-0.48 (0.01)	-0.54 (0.01)
NS-SEC	0.80 (0.01)	0.75 (0.01)
Access to employer pension scheme	-0.28 (0.02)	-0.49 (0.02)
Chance to buy shares	-0.19 (0.02)	0.01 (0.03)*
Access to company car	-0.41 (0.02)	-0.28 (0.02)
Private medical insurance	-0.43 (0.02)	-0.23 (0.02)
Highest held qualification	-0.76 (0.01)	-0.83 (0.01)
Age left full-time education	-0.63 (0.01)	-0.65 (0.01)
Highest household Cambridge scale	-0.73 (0.01)	-0.81 (0.01)

Note 1: Standardised factor loadings. All loadings statistically significant (p<0.001) except superscripts; *=p>0.05.
Note 2: Six correlations between indicator measurement errors included in the model based on modification indices 1) NS-SEC and highest household Cambridge scale; 2) Access to a company car and employer offering private medical insurance; 3) household income and employer offering private medical insurance; 4) the chance to buy shares in company and employer offering a pension scheme; 5) being in receipt of benefits and car ownership; 6) number of years in education and highest held qualification.

8.3.3 Latent Transition Analysis

8.3.3.1 *Exploratory LTA model selection*

Table 8.5 and Table 8.6, present the model fit statistics for LPA models each with 3 to 6 HRB clusters, run separately at each time point and according to gender. Table 8.7 presents the model fit statistics for a series of LTA models with numbers of HRB clusters increasing from 3 to 7.

As previously identified in chapter 6 (see section 6.3.2.1), based solely upon the model fit statistics from the LPA models at age 33, a 4-cluster model was found to be optimal for men and women at age 33. Examining the model fit statistics for men and women at age 42 also found that amongst men a 4-cluster solution was preferred. For men and women at age 33 and men at age 42, the 5-cluster LPA model BIC statistic is higher than that of the 4-cluster LPA model and the LMR LRT p value is not statistically significant (p>0.1). For women at age 42, the 5-cluster LPA model is preferable to the 4-cluster LPA model, given that there is a lower BIC statistic than the 4- and 6-cluster models, a significant LRM LRT p value (<0.1) and entropy >0.9.

Table 8.5 Model fit statistics for NCDS LPA models at age 33

MEN (n=5,586)	**Log-likelihood**	**AIC**	**BIC**	**aBIC**	**Lo-Mendell-Rubin LRT p value**	**Entropy**
3 cluster	38917.438	77916.877	78188.626	78058.340	0.03	0.978
4 cluster	**38850.141**	**77808.282**	**78166.195**	**77994.599**	**0.06**	**0.917**
5 cluster	38811.209	77756.417	78200.494	77987.589	0.15	0.861
6 cluster	38778.255	77716.510	78246.751	77992.536	0.76	0.874
WOMEN (n=5,787)	**Log-likelihood**	**AIC**	**BIC**	**aBIC**	**Lo-Mendell-Rubin LRT p value**	**Entropy**
3 cluster	38915.657	77913.313	78186.511	78056.225	<0.001	0.899
4 cluster	**38849.292**	**77806.585**	**78166.407**	**77994.811**	**0.01**	**0.905**
5 cluster	38817.567	77769.134	78215.580	78002.673	0.23	0.903
6 cluster	38790.693	77741.385	78274.455	78020.238	0.81	0.884

Note: Preferred model highlighted in bold. Bootstrapped Lo-Mendell-Rubin LRT p values could not be estimated.

Table 8.6 Model fit statistics for NCDS LPA models at age 42

MEN (n=5,529)	Log-likelihood	AIC	BIC	aBIC	Lo-Mendell-Rubin LRT p value	Entropy
3 cluster	35811.814	71705.629	71976.957	71846.671	0.01	0.960
4 cluster	**35726.248**	**71560.496**	**71917.855**	**71746.259**	**<0.001**	**0.947**
5 cluster	35681.915	71497.829	71941.219	71728.314	0.20	0.907
6 cluster	35649.885	71459.771	71989.192	71734.976	0.51	0.919
WOMEN (n=5,683)	**Log-likelihood**	**AIC**	**BIC**	**aBIC**	**Lo-Mendell-Rubin LRT p value**	**Entropy**
3 cluster	36042.687	72167.375	72439.830	72309.544	0.05	0.928
4 cluster	35940.479	71988.959	72347.801	72176.206	0.01	0.964
5 cluster	**35873.780**	**71881.560**	**72326.791**	**72113.884**	**0.02**	**0.906**
6 cluster	35826.518	71813.035	72344.654	72090.438	0.10	0.898

Note: Preferred model highlighted in bold. Bootstrapped Lo-Mendell-Rubin LRT p values could not be estimated.

The model fit statistics comparing LTA models, incorporating information from both time points, suggest the 6-cluster LTA model as the optimal solution. This model has a lower BIC than LTA models with 5 and 7 clusters, and the 7-cluster LTA model for men is not identified.

Table 8.7 Model fit statistics for NCDS LTA models

MEN (n=6,396)	Log-likelihood	AIC	BIC	aBIC	Entropy
3 cluster	74322.630	148757.260	149136.012	148958.059	0.869
4 cluster	74067.639	148291.277	148818.825	148570.961	0.888
5 cluster	73906.364	148016.729	148706.598	148382.469	0.890
6 cluster	**73763.618**	**147783.237**	**148648.955**	**148242.205**	**0.887**
7 cluster [a]	73673.958	147659.915	148715.010	148219.282	0.902
WOMEN (n=6,388)	**Log-likelihood**	**AIC**	**BIC**	**aBIC**	**Entropy**
3 cluster	74093.473	148298.947	148677.628	148499.675	0.904
4 cluster	73845.118	147846.236	148373.686	148125.822	0.901
5 cluster	73670.379	147544.758	148234.500	147910.371	0.903
6 cluster	**73507.250**	**147270.500**	**148136.059**	**147729.308**	**0.905**
7 cluster [b]	73422.916	147157.832	148212.731	147717.004	0.905

Note: Measurement invariance assumed over time. Preferred model highlighted in bold.
a The model was unidentified the log-likelihood did not replicate. Out of 1,000 fully run models with random starts 635 did not converge. Fit statistics may not be reliable but are included for comparative purposes.
b The model was identified the log-likelihood was replicated. Out of 1,000 fully run models with random starts 470 did not converge.

Based on these model fit statistics, it would appear that the 6-cluster LTA should be selected. However, as recommended (Collins and Lanza, 2010), estimates from the 4-, 5- and 6-cluster LTA models were examined to ascertain whether cluster size and interpretability were adequate. Additionally, given that LTA models rely heavily on the measurement invariance assumption (Collins and Lanza, 2010), estimates from 4-, 5- and 6-cluster LPA models at each time point were also examined to ascertain whether measurement invariance could be reasonably assumed. The estimates from these exploratory 4-, 5- and 6-cluster LTA and LPA models are presented in Appendix 8.4.

In all of the LTA models, the original three cluster patterns (‘Risky’, ‘Moderate Smokers’ and ‘Mainstream’) appeared to be stable, although there was some variation in cluster prevalence. These cluster patterns were also stable in the LPA models at ages 33 and 42. The 4-, 5- and 6-cluster LTA models (and to some extent the LPA models) elucidated additional clustered patterns. These additional clustered patterns are described fully in Appendix 8.4.

For men and women, the cluster patterns found in the 4-, 5- and 6-cluster LPA models were not consistent with those identified in the 4-, 5- and 6-cluster LTA models. This suggested that there was little substantive evidence of measurement invariance holding for each of these models.

8.3.3.2 *Confirmatory LTA model selection*

Table 8.8 and Table 8.9 present the results of a 3-cluster LTA model, run separately for men and women. These models can be considered as ‘fixed’ models because they assume measurement invariance over time – i.e. that the nature of the three previously derived clusters, ‘Risky’, ‘Moderate Smokers’ and ‘Mainstream’, are the same at ages 33 and 42.

In order to assess measurement invariance, the ‘fixed’ 3-cluster LTA models were compared with estimates from ‘free’ LTA models, whereby cluster patterns were free to vary at each time point. Table 8.10 presents the results of a statistical test for measurement invariance (i.e. chi-square difference test of LTA model log-likelihoods).

Comparisons of the ‘fixed’ and ‘free’ 3-cluster LTA models (see Table 8.10) found the chi-square difference test to be significant ($p<0.001$) suggesting that measurement invariance does not hold quantitatively for men or women. This is to be expected given the large sample sizes. However, the measurement invariance assumption of the 'fixed' 3 cluster LTA models could be considered reasonable based on a substantive examination of the LPAs run separately at each time point for each gender (see Appendix 8.5).

Table 8.8 Estimates from 'fixed' 3-cluster LTA model in the NCDS (men)

NCDS Men Total N=6,396	**Cluster 1 'Risky'**	**Cluster 2 'Moderate Smokers'**	**Cluster 3 'Mainstream'**
Latent status prevalence	**n (%)**	**n (%)**	**n (%)**
Time 1 (Age 33)	849 (13.2%)	1,247 (19.5%)	4,301 (67.2%)
Time 2 (Age 42)	195 (3.0%)	1,616 (25.3%)	4,584 (71.7%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	22.38 (14.5)	17.16 (0.67)	0
Frequency of fruit and vegetable consumption	3.06 (0.21)	4.38 (0.14)	4.97 (0.03)
Frequency of fried food consumption	5.10 (0.25)	3.00 (0.05)	3.03 (0.02)
Frequency of sweet food consumption	3.61 (0.20)	4.31 (0.10)	4.64 (0.03)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.50 (0.04)	0.41 (0.02)	0.29 (0.01)
Once a week	0.17 (0.02)	0.19 (0.01)	0.21 (0.01)
2–3 days a week	0.15 (0.02)	0.19 (0.01)	0.25 (0.01)
4–7 days a week	0.18 (0.02)	0.21 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week			
No units	0.14 (0.02)	0.16 (0.01)	0.12 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.39 (0.04)	0.46 (0.02)	0.59 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.48 (0.04)	0.38 (0.02)	0.29 (0.01)
Transition probabilities (standard errors) from age 33 (rows) to age 42 (columns)			
Cluster 1 'Risky'	**0.23 (0.03)**	0.60 (0.03)	0.17 (0.02)
Cluster 2 'Moderate Smokers'	0[a]	**0.72 (0.02)**	0.27 (0.02)
Cluster 3 'Mainstream'	0[a]	0.05 (<0.01)	**0.95 (<0.01)**

Note: Cluster prevalence based on estimated model. Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.9 Estimates from 'fixed' 3-cluster LTA model in the NCDS (women)

NCDS Women Total N=6,388	**Cluster 1 'Risky'**	**Cluster 2 'Moderate Smokers'**	**Cluster 6 'Mainstream'**
Latent status prevalence	**n (%)**	**n (%)**	**n (%)**
Time 1 (Age 33)	725 (11.4%)	1,318 (20.6%)	4,345 (68.0%)
Time 2 (Age 42)	524 (8.2%)	1,289 (20.1%)	4,575 (71.6%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	21.05 (0.07)	13.88 (0.28)	0
Frequency of fruit and vegetable consumption	3.44 (0.09)	5.66 (0.11)	6.07 (0.03)
Frequency of fried food consumption	3.67 (0.17)	2.37 (0.05)	2.41 (0.02)
Frequency of sweet food consumption	4.04 (0.21)	4.16 (0.10)	4.67 (0.03)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.63 (0.02)	0.33 (0.02)	0.29 (0.01)
Once a week	0.13 (0.01)	0.18 (0.01)	0.21 (0.01)
2–3 days a week	0.07 (0.01)	0.19 (0.01)	0.22 (0.01)
4–7 days a week	0.17 (0.02)	0.30 (0.01)	0.28 (0.01)
Alcohol units consumed in the previous week			
No units	0.30 (0.03)	0.28 (0.01)	0.25 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.47 (0.02)	0.56 (0.02)	0.64 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.23 (0.02)	0.15 (0.01)	0.11 (0.01)
Transition probabilities (standard errors) from age 33 (rows) to age 42 (columns)			
Cluster 1 'Risky'	**0.72 (0.04)**	0.19 (0.04)	0.09 (0.02)
Cluster 2 'Moderate Smokers'	0[a]	**0.75 (0.03)**	0.25 (0.02)
Cluster 3 'Mainstream'	0.001 (<0.01)	0.04 (<0.01)	**0.96 (<0.01)**

Note: Cluster prevalence based on estimated model. Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.10 Chi-square difference tests for LTA model log-likelihoods (fixed vs free)

NCDS Men [a] (n=6,396)	Log-likelihood	No. of parameters estimated	Scaling correction factor	Difference test *p* value [c]	AIC	BIC	aBIC	Entropy
Model 1[a]	-74322.662	56	1.3541	<0.001	148757.324	149136.076	148958.122	0.869
Model 2[b]	-73715.035	82	1.1565		147594.070	148148.671	147888.097	0.907
NCDS Women [a] (n=6,388)	**Log-likelihood**	**No. of parameters estimated**	**Scaling correction factor**	**Difference test *p* value [c]**	**AIC**	**BIC**	**aBIC**	**Entropy**
Model 1[a]	-74093.473	56	1.2841	<0.001	148298.947	148677.628	148499.675	0.904
Model 2[b]	-73606.027	82	1.1385		147376.053	147930.552	147669.977	0.923

a Model 1 = cluster parameter estimates (item means/thresholds) are restricted to be equal over time.
b Model 2 = cluster parameter estimates (item means/thresholds) can vary over time.
c Difference test *p* value for nested models (comparing fixed and free).

Estimates from the ‘fixed’ LTA models were compared with the ‘free’ LTA models (see Table 8.11 to Table 8.16). Unlike the LPA models run separately at each time point (see Appendix 8.5), the ‘free’ LTA model incorporated longitudinal information to influence the measurement model. As mentioned above (see section 8.2.3.3), given that LTA models are inherently autoregressive, emphasis was placed on the estimates from the ‘free’ LTA model, incorporating information from both time points, in comparison with estimates from the LPA models run separately at each time point in order to ascertain whether measurement invariance could be reasonably assumed.

The item means and thresholds from these ‘free’ LTA models (tables Table 8.11 to Table 8.14) suggested that for both genders the cluster patterns for the ‘Moderate Smokers’ and ‘Mainstream’ clusters were similar across time points. Moreover, the transition probabilities for these two clusters were similar to those found in the ‘fixed’ LTA models (tables Table 8.15 to Table 8.16). This suggested that for men and women measurement invariance could be reasonably assumed for the ‘Moderate Smokers’ and ‘Mainstream’ clusters.

However, the ‘free’ LTA model estimates indicated that cluster patterns were less consistent for the ‘Risky’ cluster at each time point. Examination of the parameters (item means/thresholds) and standard errors elucidated that the ‘Risky’ cluster patterns differed for men in relation to fried food consumption (age 33 mean=5.15 (sd 0.17), age 42 mean=3.54 (sd 0.14), $p<0.05$) and for women in relation to fried food consumption (age 33 mean=4.13 (sd 0.11), age 42 mean=2.83 (sd

0.10), p<0.05) and physical activity (≤3 times a month age 33=57%, age 42=65%, overall p<0.05). Therefore, the difference in the clustered patterns and transitions for the ‘Risky’ cluster in the ‘fixed’ LTA model compared to the ‘free’ LTA model suggested that the assumption of measurement invariance over time did not fully hold for members of the ‘Risky’ cluster. This indicates partial measurement invariance over time for the ‘Risky’ cluster which is further investigated in section 8.3.3.4.

Table 8.11 Parameter estimates from 'free' 3-cluster LTA model at age 33 in the NCDS (men)

NCDS Men Total N=6,396	**Cluster 1 'Risky' n=602 (9.4%)**	**Cluster 2 'Moderate Smokers' n=1,493 (23.4%)**	**Cluster 3 'Mainstream' n=4,300 (67.2%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	25.20 (1.50)	15.87 (0.35)	0
Frequency of fruit and vegetable consumption	2.78 (0.11)	4.30 (0.12)	4.65 (0.03)
Frequency of fried food consumption	5.15 (0.17)	3.60 (0.08)	3.37 (0.02)
Frequency of sweet food consumption	3.44 (0.23)	4.41 (0.13)	4.71 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.54 (0.04)	0.35 (0.02)	0.28 (0.01)
Once a week	0.17 (0.03)	0.22 (0.01)	0.21 (0.01)
2–3 days a week	0.14 (0.02)	0.21 (0.01)	0.25 (0.01)
4–7 days a week	0.16 (0.02)	0.22 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week			
No units	0.15 (0.03)	0.14 (0.01)	0.13 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.30 (0.04)	0.56 (0.02)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.55 (0.04)	0.30 (0.02)	0.24 (0.01)

Note: Cluster prevalence based on estimated model.

Table 8.12 Parameter estimates from 'free' 3-cluster LTA model at age 42 in the NCDS (men)

NCDS Men Total N=6,396	**Cluster 1 'Risky' n=535 (8.3%)**	**Cluster 2 'Moderate Smokers' n=1,274 (19.9%)**	**Cluster 3 'Mainstream' n=4,587 (71.7%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	25.25 (1.55)	16.02 (0.36)	0
Frequency of fruit and vegetable consumption	2.94 (0.12)	4.67 (0.14)	5.27 (0.03)
Frequency of fried food consumption	3.54 (0.14)	2.79 (0.04)	2.71 (0.02)
Frequency of sweet food consumption	3.41 (0.24)	4.39 (0.15)	4.50 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.62 (0.04)	0.40 (0.02)	0.30 (0.01)
Once a week	0.10 (0.03)	0.19 (0.01)	0.21 (0.01)
2–3 days a week	0.09 (0.02)	0.19 (0.01)	0.24 (0.01)
4–7 days a week	0.19 (0.03)	0.22 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week			
No units	0.18 (0.03)	0.16 (0.02)	0.11 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.23 (0.03)	0.43 (0.02)	0.54 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.59 (0.04)	0.40 (0.02)	0.34 (0.01)

Note: Cluster prevalence based on estimated model.

Table 8.13 Parameter estimates from 'free' 3-cluster LTA model at age 33 in the NCDS (women)

NCDS Women Total N=6,388	**Cluster 1 'Risky'** **n=750 (11.7%)**	**Cluster 2 'Moderate Smokers'** **n=1,295 (20.3%)**	**Cluster 3 'Mainstream'** **n=4,343 (68.0%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	20.62 (0.67)	13.72 (0.27)	0
Frequency of fruit and vegetable consumption	3.52 (0.10)	5.67 (0.10)	5.79 (0.03)
Frequency of fried food consumption	4.13 (0.11)	2.54 (0.07)	2.55 (0.02)
Frequency of sweet food consumption	4.14 (0.20)	4.25 (0.11)	4.85 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.57 (0.02)	0.29 (0.02)	0.27 (0.01)
Once a week	0.19 (0.02)	0.22 (0.01)	0.24 (0.01)
2–3 days a week	0.09 (0.01)	0.20 (0.01)	0.21 (0.01)
4–7 days a week	0.18 (0.02)	0.31 (0.02)	0.28 (0.01)
Alcohol units consumed in the previous week			
No units	0.29 (0.03)	0.29 (0.02)	0.29 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.54 (0.03)	0.62 (0.02)	0.65 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.17 (0.02)	0.09 (0.01)	0.07 (0.01)

Note: Cluster prevalence based on estimated model.

Table 8.14 Parameter estimates from 'free' 3-cluster LTA model at age 42 in the NCDS (women)

NCDS Women Total N=6,388	**Cluster 1 'Risky'** **n=700 (11.0%)**	**Cluster 2 'Moderate Smokers'** **n=1,111 (17.4%)**	**Cluster 3 'Mainstream'** **n=4,578 (71.7%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	20.12 (0.81)	13.70 (0.03)	0
Frequency of fruit and vegetable consumption	3.50 (0.15)	5.86 (0.17)	6.35 (0.03)
Frequency of fried food consumption	2.83 (0.10)	2.16 (0.04)	2.26 (0.02)
Frequency of sweet food consumption	3.98 (0.20)	4.03 (0.12)	4.47 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.65 (0.03)	0.34 (0.03)	0.31 (0.01)
Once a week	0.09 (0.01)	0.15 (0.01)	0.18 (0.01)
2–3 days a week	0.07 (0.01)	0.20 (0.02)	0.23 (0.01)
4–7 days a week	0.20 (0.02)	0.31 (0.02)	0.28 (0.01)
Alcohol units consumed in the previous week			
No units	0.33 (0.03)	0.27 (0.02)	0.22 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.38 (0.03)	0.52 (0.02)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.29 (0.03)	0.21 (0.02)	0.15 (0.01)

Note: Cluster prevalence based on estimated model.

Table 8.15 Transition probabilities from 'free' 3-cluster LTA model in the NCDS (men)

Transition probabilities (standard errors) from age 33 (rows) to age 42 (columns)			
Cluster 1 'Risky'	**0.88 (0.03)**	0[b]	0.11 (0.02)
Cluster 2 'Moderate Smokers'	0[a]	**0.72 (0.02)**	0.28 (0.02)
Cluster 3 'Mainstream'	0.001 (<0.01)	0.04 (<0.01)	**0.96 (<0.01)**

Note: Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.16 Transition probabilities from 'free' 3-cluster LTA model in the NCDS (women)

Transition probabilities (standard errors) from age 33 (rows) to age 42 (columns)			
Cluster 1 'Risky'	**0.92 (0.03)**	0[a]	0.08 (0.02)
Cluster 2 'Moderate Smokers'	0[a]	**0.74 (0.02)**	0.26 (0.02)
Cluster 3 'Mainstream'	0.003 (<0.01)	0.04 (<0.01)	**0.96 (<0.01)**

Note: Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

8.3.3.3 *LTA model selection summary*

Based on a comparison of exploratory and confirmatory approaches to LTA model selection, the confirmatory approach (i.e. taking into consideration prior knowledge on the existence of 3 HRB clusters at age 33) was considered superior.

Based on model fit statistics, the 6-cluster solution appeared optimal. However, further investigations comparing LTA and LPA estimates found the three HRB cluster patterns identified in the 3-cluster model to be present and stable in the 4-, 5- and 6-cluster models. Moreover, there was little substantive evidence of measurement invariance holding for the 4-, 5- and 6-cluster models.

In contrast, the confirmatory 3-cluster LTA model not only incorporated prior knowledge gained through the exploratory LPA conducted to achieve objective one of this thesis (see chapter 6), there was stronger evidence of measurement invariance holding in the 3-cluster model compared to the 4-, 5- and 6-cluster models. The clustered patterns of the 'fixed' 3-cluster LTAs were similar to those of the 'free' LTA models and the LPAs at age 33, although there was evidence of partial measurement invariance for the 'Risky' cluster.

In consequence, the LTA models with four, five and six clusters, identified through exploratory investigations, were not felt to be optimal in comparison to the 3-cluster LTA models. This adds support to arguments that the measurement invariance assumption which underpins LTA models should be explicitly tested (Collins and Lanza, 2010) and that prior knowledge should be drawn upon when considering LTA model selection (Finch and Bronk, 2011).

The next step was further examinations of partial measurement invariance identified for the ‘Risky’ cluster in the 3-cluster LTA models.

8.3.3.4 *Partial measurement invariance for the 3-cluster LTA model*

Measurement invariance analysis undertaken in section 8.3.3.2 found partial measurement invariance for the ‘Risky’ cluster in the 3-cluster LTA models. Partial measurement invariance related to fried food consumption for men and women and leisure-time physical activity for women. This implies that change between ages 33 and 42 for these HRBs is over and above what can be captured by the HRB cluster transitions estimated in the LTA model and that change for these HRBs extends beyond the interrelationship of the four HRBs.

In order to further explore partial measurement invariance for the ‘Risky’ cluster, ‘partial’ LTA models were run. These models used the starting values from the ‘fixed’ LTA models to constrain item means and thresholds to be the same for the ‘Moderate Smokers’ and ‘Mainstream’ clusters over time. Item means and thresholds for the ‘Risky’ cluster were also constrained to be the same over time, except for fried food consumption amongst men and leisure-time physical activity amongst women. The fried food consumption mean (amongst men) and leisure-time physical activity thresholds (amongst women) were freely estimated over time. Whilst fried food consumption amongst women was also found to be different at each time point this was fixed to be the same over time on the basis that the mean consumption frequency at age 33 rounded to 4 and at age 42 rounded to 3. For women, these values of 3 and 4 equate to the same frequency category over time ‘1–2 days per week’. By contrast, for men, the mean fried food consumption at age 33 was 5 and a mean of 4 at age 42. These values, 4 and 5, equate to ‘1–2 days per week’ and ‘3–6 days per week’ respectively.

For men and women, statistical investigations found these ‘partial’ LTA models to have a preferable model fit in comparison to the ‘fixed’ LTA model. However, the fit remained worse than that of the ‘free’ LTA models (see Table 8.17). As expected, for men and women the cluster probability transitions for the ‘Risky’ cluster in the ‘partial’ LTA model were found to lie between those estimated in the ‘fixed’ and ‘free’ LTA models. This indicates that imposing differing levels of measurement invariance in the models introduces differing levels of bias in the ‘Risky’ cluster patterns and transitions.

Table 8.17 Model fit statistics for ‘fixed’, ‘partial’ and ‘free’ LTA model log-likelihoods

NCDS Men[a] (n=6,396)	**Log-likelihood**	**AIC**	**BIC**	**aBIC**	**Entropy**
Fixed model	-74322.662	148757.324	149136.076	148958.122	0.869
Partial model	-74233.196	148490.393	148571.554	148533.421	0.870
Free model	-73715.035	147594.070	148148.671	147888.097	0.907
NCDS Women[a] (n=6,388)	**Log-likelihood**	**AIC**	**BIC**	**aBIC**	**Entropy**
Fixed model	-74093.473	148298.947	148677.628	148499.675	0.904
Partial model	-74080.213	148188.425	148283.095	148238.607	0.902
Free model	-73606.027	147376.053	147930.552	147669.977	0.923

Note: ‘Fixed’ LTA model (full measurement invariance). Item mean and probabilities fixed to be the same across all clusters at age 33 and age 42.
‘Partial’ LTA model (partial measurement invariance). Starting values from ‘fixed’ LTA model used to constrain parameter estimates (item means/thresholds) for the ‘Moderate Smokers’ and ‘Mainstream’ clusters. Parameter estimates (item means/thresholds) also constrained for the ‘Risky’ cluster except for fried food consumption in men and physical activity in women.
‘Free’ LTA model (no measurement invariance). Cluster parameter estimates (item means/thresholds) can vary over time.

On the basis of the above analyses, for men and women the ‘partial’ LTA models were selected. Whilst a ‘free’ LTA model had better model fit, on balance the ‘partial’ LTA model was preferred because it resulted in a more parsimonious and interpretable solution (Collins and Lanza, 2010). Thus, the ‘partial’ LTA models were considered to be a reasonable compromise, taking into account that full measurement invariance cannot be reasonably assumed for the ‘Risky’ cluster.

Table 8.20 to Table 8.23 provide the HRB cluster prevalence and patterns in the ‘partial’ LTA models at each age. Notably, ‘Risky’ cluster membership for men at age 33 (13.1%) is larger in the LTA models than those found in the LPA models for NCDS men at age 33 (1.7%) which are described in Chapter 6 (see Table 6.12). This can be explained by the different samples used to derive the models, the LTA models in this chapter include information on HRBs at age 33 and age 42 (see section 8.2.1) whereas the LPA models in chapter 6 are based solely on information at age 33.

8.3.3.5 *Transitions in HRB cluster membership during mid-life based on LTA models with partial measurement invariance over time*

Table 8.18 to Table 8.19 provide the probabilities for transitions in HRB cluster membership between ages 33 and 42.

Together, these findings are similar to those identified in the ‘fixed’ LTA models which imposed full measurement invariance (see Table 8.8 and Table 8.9), suggesting that HRBs are relatively stable during mid-age. This is indicated by a relatively high probability (>67%, p<0.001) of remaining in the same cluster, whether it be ‘Risky’, ‘Moderate Smokers’ or ‘Mainstream’, men and women.

At the same time, these results demonstrate that some participants did transition to another HRB cluster at age 42. Amongst those that did move, membership of HRB clusters at age 42 tending to be characterised by HRBs that were more beneficial for health in comparison to HRB cluster membership at age 33. For example, amongst men and women there was more than a 25% probability (men=27%, women=26%, p<0.001) of transitioning from the ‘Moderate Smokers’ to the ‘Mainstream’ cluster between ages 33 and 42 and more than 9% probability (men=17%, women=9%, p<0.001) probability of transitioning from the ‘Risky’ to the ‘Mainstream’ cluster. Given that members of the ‘Mainstream’ cluster were non-smokers, had a higher frequency of fruit and vegetable consumption, a lower frequency of fried food consumption and a higher frequency of leisure-time physical activity in comparison to the ‘Risky’ and ‘Moderate Smokers’ clusters (p<0.001), this movement suggests that during mid-life HRBs tend to improve.

Appendix 8.6 shows how the individual behaviours at the two time points differ for participants who moved to a different HRB cluster between ages 33 and 42 in comparison to those who remained in the same HRB cluster over time.

Table 8.18 Transition probabilities from 'partial' 3 cluster LTA model in the NCDS (men)

Transition probabilities (standard errors) from age 33 (rows) to age 42 (columns)			
Cluster 1 'Risky'	**0.67 (0.08)**	0.16 (0.08)	0.17 (0.02)
Cluster 2 'Moderate Smokers'	0[a]	**0.73 (0.16)**	0.27 (0.03)
Cluster 3 'Mainstream'	0[a]	0.04 (<0.01)	**0.96 (<0.01)**

Note: Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.19 Transition probabilities from 'partial' 3 cluster LTA model in the NCDS (women)

Transition probabilities (standard errors) from age 33 (rows) to age 42 (columns)			
Cluster 1 'Risky'	**0.70 (0.04)**	0.21 (0.04)	0.09 (0.02)
Cluster 2 'Moderate Smokers'	0[a]	**0.74 (0.03)**	0.26 (0.02)
Cluster 3 'Mainstream'	0.002 (<0.01)	0.04 (<0.01)	**0.96 (<0.01)**

Note: Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.20 Parameter estimates from 'partial' 3-cluster LTA model at age 33 in the NCDS (men)

NCDS Men Total N=6,396	**Cluster 1 'Risky' n=838 (13.1%)**	**Cluster 2 'Moderate Smokers' n=1,259 (19.7%)**	**Cluster 3 'Mainstream' n=4,299 (67.2%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	22.38	17.16	0
Frequency of fruit and vegetable consumption	3.06	4.38	4.97
Frequency of fried food consumption	5.14 (0.09)	3.00	3.03
Frequency of sweet food consumption	3.61	4.31	4.64
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.50	0.41	0.29
Once a week	0.17	0.19	0.21
2–3 days a week	0.15	0.19	0.25
4–7 days a week	0.18	0.21	0.26
Alcohol units consumed in the previous week			
No units	0.14	0.16	0.12
Within limits (≤14 units women, ≤21 units men)	0.39	0.46	0.59
Above limits (≥15 units women, ≥22 units men)	0.48	0.38	0.29

Note: Cluster prevalence based on estimated model. Standard errors not estimated for parameters constrained using starting values from 'fixed' LTA model (see table 8.8).

Table 8.21 Parameter estimates from 'partial' 3-cluster LTA model at age 42 in the NCDS (men)

NCDS Men Total N=6,396	**Cluster 1 'Risky' n=559 (8.7%)**	**Cluster 2 'Moderate Smokers' n=1,251 (19.6%)**	**Cluster 3 'Mainstream' n=4,586 (71.7%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	22.38	17.16	0
Frequency of fruit and vegetable consumption	3.06	4.38	4.97
Frequency of fried food consumption	3.53 (0.20)	3.00	3.03
Frequency of sweet food consumption	3.61	4.31	4.64
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.50	0.41	0.29
Once a week	0.17	0.19	0.21
2–3 days a week	0.15	0.19	0.25
4–7 days a week	0.18	0.21	0.26
Alcohol units consumed in the previous week			
No units	0.14	0.16	0.12
Within limits (≤14 units women, ≤21 units men)	0.39	0.46	0.59
Above limits (≥15 units women, ≥22 units men)	0.48	0.38	0.29

Note: Cluster prevalence based on estimated model. Standard errors not estimated for parameters constrained using starting values from 'fixed' LTA model (see table 8.8).

Table 8.22 Parameter estimates from 'partial' 3-cluster LTA model at age 33 in the NCDS (women)

NCDS Women Total N=6,388	**Cluster 1 'Risky'** **n=737 (11.5%)**	**Cluster 2 'Moderate Smokers'** **n=1,306 (20.5%)**	**Cluster 3 'Mainstream'** **n=4,345 (68.0%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	21.05	13.88	0
Frequency of fruit and vegetable consumption	3.44	5.66	6.07
Frequency of fried food consumption	3.67	2.37	2.41
Frequency of sweet food consumption	4.04	4.16	4.67
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.57 (0.02)	0.33	0.29
Once a week	0.18 (0.01)	0.18	0.21
2–3 days a week	0.09 (0.01)	0.19	0.22
4–7 days a week	0.17 (0.02)	0.30	0.28
Alcohol units consumed in the previous week			
No units	0.30	0.28	0.25
Within limits (≤14 units women, ≤21 units men)	0.47	0.56	0.64
Above limits (≥15 units women, ≥22 units men)	0.23	0.15	0.11

Note: Cluster prevalence based on estimated model. Standard errors not estimated for parameters constrained using starting values from 'fixed' 3 cluster LTA model (see table 8.9).

Table 8.23 Parameter estimates from 'partial' 3-cluster LTA model at age 42 in the NCDS (women)

NCDS Women Total N=6,388	**Cluster 1 'Risky'** **n=516 (8.1%)**	**Cluster 2 'Moderate Smokers'** **n=1,297 (20.3%)**	**Cluster 3 'Mainstream'** **n=4,575 (71.6%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	21.05	13.88	0
Frequency of fruit and vegetable consumption	3.44	5.66	6.07
Frequency of fried food consumption	3.67	2.37	2.41
Frequency of sweet food consumption	4.04	4.16	4.67
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.72 (0.03)	0.33	0.29
Once a week	0.07 (0.02)	0.18	0.21
2–3 days a week	0.05 (0.01)	0.19	0.22
4–7 days a week	0.16 (0.02)	0.30	0.28
Alcohol units consumed in the previous week			
No units	0.30	0.28	0.25
Within limits (≤14 units women, ≤21 units men)	0.47	0.56	0.64
Above limits (≥15 units women, ≥22 units men)	0.23	0.15	0.11

Note: Cluster prevalence based on estimated model. Standard errors not estimated for parameters constrained using starting values from 'fixed' 3 cluster LTA model (see table 8.9).

8.3.4 Latent Transition Analysis with a covariate

Based on the results of the measurement invariance analysis outlined above, the 3-cluster LTA models with 'partial' measurement invariance were taken forward and incorporated in to a 'LTA with a covariate' model. As mentioned in section 8.2.3.4, this model was operationalised using structural equation models (SEMs) in Mplus Version 7 (Muthén, 2014). The measurement part of each SEM was the derivation of the HRB clusters at ages 33 and 42 (using the gender-specific 'partial' 3-cluster LTA model). This produced a three-category latent dependent variable, capturing 'Risky', 'Moderate Smokers' and 'Mainstream' HRB clusters over time. The structural part of each SEM was the transition between HRB clusters at ages 33 and 42, the path between an observed continuous variable capturing SEP at age 33 (higher scores =more disadvantaged) and HRB cluster membership at age 33 (see Figure 8.1) and the path between SEP at age 33 and transitions in HRB cluster membership (see Figure 8.2).

Table 8.24, presents the model fit indices from the 'LTA with a covariate' models. Model 1 estimated the crude effect of SEP at age 33 on HRB cluster membership at age 33 (see Figure 8.1). Model 2 investigated moderation to ascertain the effect of SEP at age 33 on HRB cluster membership transitions (see Figure 8.2). The BIC and aBIC fit indices (which balance model fit and parsimony) were found to be higher for model 2 in comparison to model 1 implying a poorer model fit, this may be due to model 1 being more parsimonious than model 2. In contrast, the AIC (which also balances model fit and parsimony) was found to be lower for model 2 and the log-likelihood chi-square difference test was significant (p<0.001), both suggesting improved model fit.

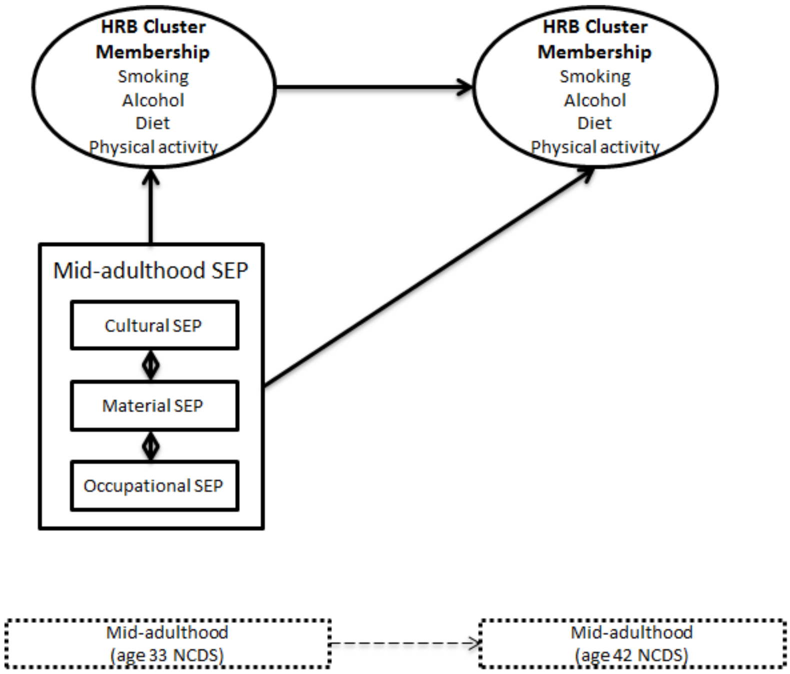
Table 8.24 Model fit statistics from 'LTA with a covariate' models 1 and 2

MEN (n=6,396)	**Log-likelihood**	**Scaling correction factor**	**Number of parameters**	**Difference test *p* value [a]**	**AIC**	**BIC**	**aBIC**	**Entropy**
Model 1	-73878.926	0.7872	10	<0.001	147777.853	147845.487	147813.710	0.876
Model 2	-73865.112	0.7260	16		147762.224	147870.439	147819.595	0.875
WOMEN (n=6,388)	**Log-likelihood**	**Scaling correction factor**	**Number of parameters**	**Difference test *p* value [a]**	**AIC**	**BIC**	**aBIC**	**Entropy**
Model 1	-73626.269	0.8963	10	<0.001	147272.539	147340.160	147308.383	0.906
Model 2	-73611.784	0.8261	16		147255.568	147363.763	147312.919	0.906

Note: Partial measurement invariance over time for fried food consumption in the 'Risky' cluster for men. Partial measurement invariance over time for physical activity in the 'Risky' cluster for women.
Model 1=Effect of SEP on HRB cluster membership age 33.
Model 2=Moderation model, effect of SEP on transitions in HRB cluster membership over time.
a= Difference test *p* value for nested models (comparing model 1 and model 2).

Because of the ambiguity of the model fit indices, further models were run in order to explore whether SEP at age 33 had a moderating effect on transitions in HRB cluster membership. If moderation were present (i.e. the effect of SEP at age 33 predicts transitions in HRB cluster membership over time) we would expect to see the effect of SEP at age 33 on HRB cluster membership at age 42 stratified by HRB cluster membership at age 33 (model 2, see Figure 8.2) to be different to the effect of SEP at age 33 on HRB cluster membership at age 42 adjusting for HRB cluster membership at age 33 (hereafter named model 3, see Figure 8.3).

Figure 8.3 A 'LTA with covariate model' testing the effect of SEP at age 33 on HRB cluster membership at age 42 adjusting for HRB cluster membership at age 33



The results presented in Table 8.25, suggest no significant difference in the coefficients for models 2 and 3 for the effect of SEP at age 33 on HRB cluster membership at age 42 when comparing ‘Moderate Smokers’ and ‘Mainstream’ cluster membership (see ‘cluster 2’ column), indicated by confidence intervals that overlap (NCDS men model 2 cluster 2 coefficient=0.40 (0.16, 0.63), model 2 cluster 3 coefficient=0.52 (0.31, 0.72), model 3 coefficient=0.43 (95% CI=0.27, 0.58); NCDS women model 2 cluster 2 coefficient=0.47 (0.25, 0.69), model 2 cluster 3 coefficient=0.43 (0.19, 0.67), model 3 coefficient=0.46 (95% CI=0.31, 0.62)). This implies that whilst SEP at age 33 appears to influence HRB cluster membership it does not appear to predict transitions between HRB clusters at each age.

The effect of SEP at age 33 on ‘Risky’ vs ‘Mainstream’ cluster membership at age 42 is different across models 2 and 3, and is largely non-significant (p<0.05). The coefficients are considered to be somewhat unstable, due to the small cell sizes, and in consequence these effects provide little evidence of moderation.

Due to the limited evidence of a moderating effect, model 1 was selected. This implies that whilst SEP at age 33 appears to influence HRB cluster membership at age 33 it does not predict transitions between HRB clusters at each age.

Table 8.25 Regression coefficients for the effect of SEP at age 33 on HRB cluster membership at age 42 from ‘LTA with a covariate’ models 2 and 3

NCDS Men Total N=6,396	**Cluster 1 ‘Risky’ age 42**	**Cluster 2 ‘Moderate Smokers’ age 42**	**Cluster 3 ‘Mainstream’ age 42**
	Logit coefficient (95% CI)	**Logit coefficient (95% CI)**	**Logit coefficient (95% CI)**
Model 2 (cluster 1 age 33)	0.38 (-0.14, 0.91)	-0.48 (-1.46, 0.49)	Reference
Model 2 (cluster 2 age 33)	2.63[a]	0.40 (0.16, 0.63)*	Reference
Model 2 (cluster 3 age 33)	3.63[a]	0.52 (0.31, 0.72)*	Reference
Model 3	0.44 (-0.07, 0.96)	0.43 (0.27, 0.58)*	Reference
NCDS Women Total N=6,388	**Cluster 1 ‘Risky’ age 42**	**Cluster 2 ‘Moderate Smokers’ age 42**	**Cluster 3 ‘Mainstream’ age 42**
	Logit coefficient (95% CI)	**Logit coefficient (95% CI)**	**Logit coefficient (95% CI)**
Model 2 (cluster 1 age 33)	0.05 (-0.61, 0.71)	0.28 (-0.52, 1.09)	Reference
Model 2 (cluster 2 age 33)	5.40[a]	0.47 (0.25, 0.69)*	Reference
Model 2 (cluster 3 age 33)	1.76 (0.69, 2.83)	0.43 (0.19, 0.67)*	Reference
Model 3	0.20 (-0.42, 0.83)	0.46 (0.31, 0.62)*	Reference

Note: Partial measurement invariance over time for fried food consumption in the ‘Risky’ cluster for men. Partial measurement invariance over time for physical activity in the ‘Risky’ cluster for women. SEP=socio-economic position at age 33, CI=95% confidence interval, *p≤0.01.
Superscript a = 95% CI not estimated in model, p value fixed at 0.999 in Mplus Version 7.
Model 2 = Moderation model, effect of SEP on transitions in HRB cluster membership over time (the effect of SEP at age 33 on HRB cluster membership at age 42 stratified by HRB cluster membership at age 33).
Model 3 = Effect of SEP on HRB cluster membership at age 42 (adjusting for HRB cluster membership at age 33).

The regression coefficients for model 1 (Table 8.26) suggest that for both genders a one unit increase in SEP at age 33 (equating to being more disadvantaged) was significantly associated ($p<0.001$) with HRB cluster membership at age 33. These results indicate that those experiencing more social disadvantage at age 33 had a higher likelihood of membership of the 'Moderate Smokers' cluster and to the 'Risky' cluster in comparison to the 'Mainstream' cluster.

The coefficients presented in Table 8.26 reflect a one unit increase in SEP at age 33 on HRB cluster membership at age 33. Given that SEP at age 33 ranges from -1.97 to 2.10 amongst men and -2.91 to 2.12 amongst women (see Appendix 8.3 for a full description of the SEP variable distributions) the coefficients pertaining to the 'Risky' cluster could be considered very large. This is likely the result of data sparseness due to the smaller size of the 'Risky' cluster.

Table 8.26 Regression coefficients for the effect of SEP at age 33 on HRB cluster membership at age 33 from 'LTA with a covariate' model 1

NCDS Men Total N=6,396	**Cluster 1 'Risky' age 33**	**Cluster 2 'Moderate Smokers' age 33**	**Cluster 3 'Mainstream' age 33**
	Logit coefficient (95% CI)	**Logit coefficient (95% CI)**	**Logit coefficient (95% CI)**
Model 1	2.00 (1.84, 2.15)*	0.46 (0.37, 0.56)*	Reference
NCDS Women Total N=6,388	**Cluster 1 'Risky' age 33**	**Cluster 2 'Moderate Smokers' age 33**	**Cluster 3 'Mainstream' age 33**
	Logit coefficient (95% CI)	**Logit coefficient (95% CI)**	**Logit coefficient (95% CI)**
Model 1	2.29 (2.12, 2.46)*	0.57 (0.47, 0.66)*	Reference

Note: Partial measurement invariance over time for fried food consumption in the 'Risky' cluster for men. Partial measurement invariance over time for physical activity in the 'Risky' cluster for women.
SEP=socio-economic position at age 33, CI=95% confidence interval, *p≤0.01.
Model 1= Effect of SEP on HRB cluster membership at age 33.

8.4 Summary of findings

Using prospectively collected data from a cohort of participants born in 1958, this chapter investigated the stability of HRB cluster membership between ages 33 and 42. Latent Transition Analysis (LTA) was undertaken to identify transitions between three HRB clusters: 'Risky', 'Moderate Smokers' and 'Mainstream' during mid-life. The hypothesis that HRBs are relatively stable during mid-adulthood (men=67% probability; women=70% probability) but that some change will occur, largely in a positive direction (men=27% probability; women=26% probability) was confirmed. The hypothesis that mid-adulthood SEP would influence change in HRB over time did not hold, there was no evidence of a moderating effect of mid-adulthood SEP on transitions in HRB cluster membership between ages 33 and 42.

Comparing confirmatory and exploratory approaches to LTA model selection was insightful. This analysis suggested that taking a confirmatory approach, based upon prior knowledge of 3 HRB clusters being present at age 33, was more efficient in comparison to an exploratory approach which was purely data driven. By taking a confirmatory approach, it was concluded that the 3-cluster LTA model was preferred for men and women.

Measurement invariance analysis, examining whether the nature of the HRB clusters were equivalent over time, found that full measurement invariance did not hold for the 3-cluster LTA model. Instead, models with 'partial' measurement invariance for fried food consumption amongst men and leisure-time physical activity amongst women in the 'Risky' cluster were selected. These 'partial' LTA models were considered to provide a better representation of the extent to which HRB cluster membership changed between ages 33 and 42.

Results from the 'partial' LTA models suggested that for men and women HRB cluster membership was relatively stable during mid-adulthood with a large proportion of participants (>67%, $p<0.001$) remaining in the same cluster at ages 33 and 42. At the same time this stability was not universal, and there was evidence of significant ($p<0.001$) and substantial transitions between HRB clusters during mid-life (the probability of movement from the 'Risky' and 'Moderate Smokers' clusters to the 'Mainstream' cluster being >9% and >25% respectively). This implies that assumptions regarding HRB stability in mid-adulthood do not hold and should be challenged.

The LTA models showed movement from HRB clusters characterised by multiple negative HRBs to clusters characterised by HRB patterns tending to be more beneficial for health. For example, amongst men and women there were transitions from the 'Risky' cluster and 'Moderate Smokers' cluster at age 33 to the 'Mainstream' cluster at age 42. Akin to the results in chapter 6 (see section 6.3.2.3), members of the 'Mainstream' cluster were non-smokers, had a higher frequency of fruit and vegetable consumption, a lower frequency of fried food consumption and a higher frequency of leisure-time physical activity in comparison to the 'Risky' and 'Moderate Smokers' clusters. Again, similarly to chapter 6 (see section 6.3.2.3) there was a higher frequency of sweet food consumption in the 'Mainstream' cluster in comparison to the other two clusters.

It should be noted that a transition in HRB cluster membership between ages 33 and 42 does not imply that an individual changed all four HRBs. Instead, movement to a cluster characterised by more positive HRBs than the one left behind suggests an average improvement in HRB patterns over time. This average improvement in HRBs over time was reflected, to some extent, in the descriptive statistics comparing individuals HRBs (i.e. smoking, alcohol, three diet variables and physical activity) at ages 33 and 42. Wald tests indicted a lower proportion smoking cigarettes ($p<0.001$), a higher mean frequency of fruit and vegetable consumption ($p<0.001$) and a lower mean frequency of fried food consumption ($p<0.001$).

Alongside these general improvements in HRB patterns, the LTA models indicated a higher proportion of women who never drink alcohol in the 'Risky' cluster compared to the 'Mainstream' cluster. Thus, a shift to the 'Mainstream' cluster between ages 33 and 42 implies that some women may increase their alcohol consumption from never drinking alcohol to drinking alcohol within recommended limits during mid-life. Moreover, for men and women movement to the 'Mainstream' cluster, characterised by a higher frequency of sweet food consumption in comparison to the other two clusters, implies the average frequency of sweet food consumption increased.

Interestingly, transitions between the 'Risky' cluster at age 33 and the 'Mainstream' cluster at age 42 were larger for men in comparison to women. This may be explained by differences in the clustered patterns of HRBs according to gender. For example, the patterns of alcohol consumption across the three HRB clusters are different for women compared to men. The proportion drinking above recommended limits is higher for men in all three clusters compared to women. Moreover, the observed difference across the three clusters in relation to the proportion drinking above recommended limits was more marked amongst men. In consequence, women in the 'Risky' cluster at age 33 would need to change their alcohol consumption substantially in order for them to transition to the 'Mainstream' HRB cluster at age 42. In contrast, men in the 'Risky' cluster at age 33 may only need to reduce their alcohol consumption slightly in order to make a transition to the 'Mainstream' cluster at age 42. It is possible that this is an artefact of the slight over-estimation of beer drinkers identified in Appendix 5.3, impacting on men's alcohol consumption more than women's. However, these gender differences in alcohol consumption are considered to be reliable, given evidence that men do drink more units of alcohol then women (Meng et al., 2014) which may be explained by pregnancy and motherhood (Tran et al., 2015).

Results from the 'partial' LTA models should be interpreted with consideration to the measurement variability identified in relation to particular behaviours in the 'Risky' cluster (i.e. fried food for men and leisure-time physical activity for women). This implies that change between ages 33 and 42 for these particular HRBs is over and above what can be captured by the HRB cluster transitions estimated in the model. Such an interpretation indicates changes in fried food consumption for men and leisure-time physical activity for women extend beyond the interrelationship of the four HRBs.

LTA models including SEP at age 33 as a covariate found this uni-dimensional construct, derived through CFA models, to predict HRB cluster membership at age 33. However, there was a lack of evidence of an effect of SEP at age 33 on transitions in HRB cluster membership during mid-adulthood. These results suggest that social circumstances at age 33 may influence lifestyle at age 33 but not with change in HRBs over time.

8.5 Strengths and limitations

A major strength of this work was the utilisation of longitudinal data in order to investigate the extent to which HRB cluster membership changes during mid-life and thus address the commonly held assumption that HRBs are stable during mid-life.

The application of LTA to identify the progression of HRBs over time is considered a powerful tool and superior to other methods such as index scoring (Reboussin et al., 1998) and generalised estimation equations (Yeh et al., 2012). Using LTA to examine HRB clustering over time allowed for a person-centred exploration of HRB cluster stability during mid-life accommodating multiple HRB variables to derive HRB cluster patterns representing distinct lifestyles shared by subgroups of British middle-aged adults.

Treating HRBs as time-varying builds upon the cross-sectional examination of the relationship between HRB cluster membership and transitions and SEP advances our understanding of how clustered patterns persist or decline over time (Blue et al., 2014) and provides further insights into the development of SEP inequalities in HRBs and subsequent health outcomes (Oude Groeniger and van Lenthe, 2016). Moreover, using information for HRBs at two time points from this prospectively collected data from this British birth cohort study ensures temporality and minimises recall bias (Cohen et al., 2010) as well as controlling for cohort effects and thus allowing for an investigation of period and age (Schoon, 2006).

However, the consideration of HRB cluster membership and transitions at two points is also a limitation of this study. LTA requires the same measures to be used at both time points in order to impose measurement invariance (Collins and Lanza, 2010). Therefore, only information at two time points could be incorporated into the LTA models. Ideally, information pertaining to all four HRBs from more than two time points during mid-life would have been included, allowing for a more detailed description of the transitions that may have occurred during this nine-year period (i.e. ages 33 to 42). Using repeated HRB measures at other time points may also elucidate 'natural fluctuations' in HRB cluster patterns across the lifecourse which could be optimal points for multiple HRB interventions and maximise their efficacy (Mulder et al., 1998). Moreover, the inclusion of information on all four HRBs at multiple time points may have improved statistical power (Collins and Lanza, 2010), thus increasing the ability to detect the effect of SEP at age 33 on HRB cluster transitions over time.

Additionally, including SEP at other points in the lifecourse would be insightful given the relationship between multiple risky HRBs and SEP in adolescence (Hale and Viner, 2016) and at older ages (Shankar et al., 2010) and research suggesting SEP predicts change in HRBs during earlier periods of the lifecourse (de Winter et al., 2016; White et al., 2016) and the later stages of working life (Harrington et al., 2014). Using time-varying measures of HRBs and SEP would also serve to highlight the extent to which social inequalities in HRBs may be widening over time (Oude Groeniger and van Lenthe, 2016).

Furthermore, a cross-cohort comparison with participants from the BCS70 would have been useful both contributing to disentangling age and period effects and improving the generalisability of the results to later-born cohorts. However, the same measures for three diet variables used at age 34 in the BCS70 (see section Age 33 NCDS and age 34 NCDS) were not available at any other time points during mid-age and therefore the LTA models could not be replicated in this cohort.

Information for all four HRBs was self-reported by participants and therefore could be subject to bias in regard to their measurement (Conry et al., 2011; Heroux et al., 2012; Schneider et al., 2009). For example, self-report bias has been found to influence the observed relationship between SEP and alcohol consumption (Devaux and Sassi, 2015). Social desirability bias

(Crowne and Marlowe, 1960) is possible when using information pertaining to HRBs, potentially leading to an underestimation of health-damaging HRBs and overestimating of health-promoting HRBs. This may not only explain the larger prevalence of the 'Mainstream' cluster in comparison to the 'Risky' and 'Moderate Smokers' clusters at ages 33 and 42 but also influence the transitions to the 'Mainstream' cluster. For example, persistent public health campaigns and policies have led to smoking cigarettes becoming increasingly socially unacceptable over time (Graham, 2012) and therefore could lead to increases in the under-reporting of smoking over time (Stait and Calnan, 2016).

As mentioned in chapter 6 (see section 6.5) the inclusion of additional measures of the HRBs, such as other occupational and commuting physical activity and the tendency to binge drink, would have been insightful considering their differential relationship with SEP. For example, leisure-time physical activity has been shown to be positively associated with increased educational attainment whereas the opposite appears to be true for occupational physical activity (Wallmann-Sperlich and Froboese, 2014). Moreover, there is research to suggest that, whilst those in more disadvantaged social circumstances tend to consume fewer alcohol units overall, they are more likely to partake in heavy episodic use (i.e. binge drinking) (Bellis et al., 2016).

Appendix 5.3 outlines the issues in measurement for 'beer consumption' at age 42. Although this error introduced the potential for an over-estimation of 'beer' consumption, subsequent corrections were made using alcohol consumption in relation to the remaining alcohol beverage categories at age 42 and all alcohol beverage categories at age 33. Comparing the distribution of alcohol consumption before and after these corrections and undertaking sensitivity analyses using external datasets validated this work. The analyses demonstrated that whilst beer consumption may remain slightly over-estimated in the NCDS at age 42 when compared to beer-specific consumption in external datasets, following these corrections, total alcohol consumption at age 42 did not appear to be over-estimated.

The measure of SEP at age 33 was considered to strengthen this work. The continuous construct derived using multiple SEP indicators is considered to parsimoniously capture three dimensions of SEP (material, cultural and occupational). However, unlike in chapter 7 (see section 7.3.3) the estimated factor scores were used in the 'LTA with a covariate' models meaning that SEP was included as an observed, rather than latent variable. Although it would be preferable to use the latent variable in the models, the results are not considered to be heavily influenced by the inclusion of SEP at age 33 as observed given that the regression method in estimating factor scores is considered robust (DiStefano et al., 2009) and that these models have been found to adequately estimate the relationship between an observed predictor and latent outcome (Muthén and Hsu, 1993). Moreover, sensitivity analysis carried out in chapter 7 (see section 7.3.5) found the inclusion of SEP at age 33 as both a latent and observed variable led to very similar results and did not affect the overall conclusions.

The attrition between age 33 and age 42 amongst NCDS participants has the potential to bias the results. However, the management of missing data in the LTA models was considered adequate based on the sensitivity analysis undertaken in chapter 7 (see section 7.3.5). In addition, it has been found in research investigating SEP inequalities for a range of outcomes (including smoking during pregnancy and educational attainment) that, even with high levels of attrition (>50%), conclusions regarding the direction and size of the effect did not change (Howe et al., 2013b).

9 Chapter 9: Discussion

This doctoral project makes novel and valid contributions to existing evidence in three important ways: 1) by elucidating distinct clustered patterns of HRBs from two British cohorts born 12 years apart (i.e. in 1958 and 1970); 2) by determining the predictive role of pre-adolescent social circumstances in HRB cluster membership in mid-adulthood; 3) by providing evidence that suggests stability in HRB cluster membership during mid-life.

This chapter demonstrates this contribution further by firstly providing highlights of the findings, secondly discussing these findings within the context of existing evidence, thirdly raising some key points that have emerged as a product of critiquing the analytical methods and research findings, fourthly highlighting the potential policy implications of this work and, finally, suggesting the direction of future research in HRB clustering.

9.1 Highlights of findings

This section provides highlights of the thesis' findings, consolidating the summarised results from chapters 6, 7 and 8.

9.1.1 The cross-cohort comparison of HRB clustering

The results of chapter 6 supported the first hypothesis relating to objective 1, finding three distinct HRB clusters which were labelled 'Risky', 'Moderate Smokers' and 'Mainstream'. The 'Mainstream' cluster largely consisted of multiple positive HRBs, the 'Risky' cluster was largely characterised by multiple negative HRBs and the 'Moderate Smokers' cluster was a mixture of both. The latter HRB cluster consisted of HRB patterns that are considered to be more beneficial to health than the other two clusters (i.e. not smoking, eating fruit and vegetables more frequently, eating chips and fried food less frequently and being more active). However, the frequency of sweet food consumption was generally higher in the 'Mainstream' cluster. Moreover, the hypothesis that heavy smokers and drinkers would also be more physically activity was not supported with members of the 'Risky' cluster having lower levels of physical activity in comparison to the other clusters.

The results also supported the hypothesis of cohort and gender differences in HRB cluster patterns and membership. A larger proportion of BCS70 participants were members of the 'Mainstream' cluster, compared to NCDS participants. HRB cluster patterns were found to be similar for men and women in the two cohorts in relation to smoking, fruit and vegetable consumption, fried food consumption and physical activity. There was a slight divergence for sweet food consumption amongst men in the later-born cohort who, unlike men and women in the NCDS and women in the BCS70, had the same frequency in the 'Risky' and 'Mainstream' cluster. Cohort differences were notable for alcohol consumption amongst women. For women in both cohorts the proportion drinking above recommended guidelines was highest in the 'Risky' cluster followed by the 'Moderate Smokers' and 'Mainstream' clusters respectively. However, women in the later-born cohort drank more overall. The proportion of women in the later-born cohort drinking above recommended guidelines was almost double that of their predecessors across all three clusters, with a sizeable proportion (>10%) of women in the BCS70 drinking above recommended levels in the 'Mainstream' cluster. At the same time, sweet food consumption in the 'Mainstream' cluster, whilst being higher than the 'Risky' and 'Moderate Smokers' clusters, was less frequent amongst later-born women compared to those born earlier.

9.1.2 Pre-adolescent SEP predicting HRB cluster membership in mid-adulthood

The research findings presented in chapter 7 (objective 2) provide support for some of the hypothesised paths between pre-adolescence SEP and mid-adulthood HRB cluster membership. The findings showed how social circumstances in pre-adolescence and mid-adulthood, captured by material, occupational and cultural dimensions of individual SEP, are associated with HRB cluster membership in mid-adulthood.

There was stronger evidence of an indirect effect of pre-adolescent SEP on HRB cluster membership through mid-adulthood SEP in comparison to a direct effect. For men and women in both cohorts, a substantial proportion of the total effect of pre-adolescent SEP on mid-adulthood HRB occurred via mid-adulthood SEP. Disadvantaged pre-adolescence SEP strongly predicted disadvantaged mid-adulthood SEP which was, in turn, strongly associated with 'Risky' and 'Moderate Smokers' cluster membership in mid-adulthood. This work found that differentials in HRB patterns associated with social circumstances in mid-adulthood are strengthened through the persistence of either advantaged or disadvantaged social circumstances since pre-adolescence. This was shown through analyses indicating that pre-adolescent SEP strengthened the association between mid-adulthood SEP and adult HRB cluster membership.

Together these results indicate that pre-adolescent SEP is less likely to directly shape HRB patterns in mid-adulthood independent of mid-adulthood SEP. Therefore, SEP has a contemporaneous influence on HRB clustering in mid-adulthood, dictating the proximal environment in which people live their lives strongly influencing their lifestyles. Consequently, these findings suggest a social rather than behavioural pathway between pre-adolescence and mid-adulthood HRB clustering. The behavioural pathway implies that HRBs are embedded in pre-adolescence and re-emerge in mid-adulthood whereas the social pathway implies that the accumulation of resources from pre-adolescence to mid-adulthood dictates the social circumstances in which people live their lives in mid-adulthood and consequently their lifestyles.

9.1.3 Transitions in HRB cluster membership over time and the role of mid-adulthood SEP

The findings from chapter 8 supported the first hypothesis for objective 3 (see section 4.3), showing that, whilst HRB cluster membership was relatively stable during mid-life, some transitions did occur. Participants tended to move to an HRB cluster characterised by more positive HRBs than the one they had left behind. Movement to the 'Mainstream' cluster was both statistically significant ($p<0.001$) and substantial ('Risky' cluster >9% probability, 'Moderate Smokers' cluster >25% probability), suggesting that assumptions regarding HRB stability in mid-adulthood do not necessarily hold and should be challenged.

The analyses in chapter 8 (objective 3) used the information from the cohort born in 1958, following their HRBs from age 33 to the second time point nine years later at age 42. The same three HRB clusters, identified in chapter 6 (objective 1), were found once more, despite the incorporation of additional information from the second time point.

Sensitivity analyses (see section 8.3.3.4) suggested that the nature of the HRB patterns in the 'Mainstream' and 'Moderate Smokers' clusters were somewhat equivalent across the two time points. However, for the 'Risky' cluster, differences in the nature of the HRB patterns over time were identified for some HRBs (i.e. partial measurement invariance). The sensitivity analyses suggested that changes in fried food consumption for men and leisure-time physical activity for women during the subsequent nine years extended beyond the interrelationship of the four HRBs.

There was little evidence to support the hypothesis that social circumstances at age 33 influence transitions in HRB cluster membership during mid-life. The findings indicated HRB cluster membership at age 33 was associated with social circumstances at age 33. Those in more disadvantaged social circumstances at age 33 had a higher probability of membership of the 'Risky' and 'Moderate Smokers' clusters in comparison to the 'Mainstream' cluster at age 33. However, transitions in HRB cluster membership thereafter (i.e. between ages 33 and 42) were not found to be related to social circumstances at age 33.

9.2 Limitations of self-reported HRB measures

As mentioned in chapters 6 (see section 6.5) and 8 (see section 8.5) the HRBs measures used in this thesis are self-reported and therefore social desirability bias (Crowne and Marlowe, 1960) is possible. This bias may lead to an underestimation of health-damaging HRBs and overestimating of health-promoting HRBs and may explain the larger prevalence of the 'Mainstream' cluster in comparison to the 'Risky' and 'Moderate Smokers' clusters at ages 33 and 42 as well as influence the transitions to the 'Mainstream' cluster.

For example, persistent public health campaigns and policies have led to smoking cigarettes becoming increasingly socially unacceptable over time (Graham, 2012) and therefore could lead to increases in the under-reporting of smoking (Stait and Calnan, 2016). Similarly, awareness of the 'five a day' public health message (FSA, 2008), physical activity recommendations (Stamatakis et al., 2007) and alcohol consumption guidelines (DOH, 1995) could result in over-reporting of these behaviours.

Moreover, such bias may differ according to SEP given the increased receptivity to public health messages amongst more advantaged SEP groups (Cutler and Lleras-Muney, 2010; Schooling and Kuh, 2002).

9.3 Limitations of the clustering approach.

The purpose of this thesis was to empirically derive clustered patterns of HRBs in two British birth cohorts in order to identify a complex web of HRB patterns that suggest distinct typologies are practised by different types of people providing a person-centred understanding of lifestyle. However, it is acknowledged that these HRB clusters provide a summary of the different ways in which HRB patterns may occur in these cohorts. The large size of the 'Mainstream' cluster suggests that they are likely to be a heterogeneous group which should be taken into consideration when interpreting these results.

The inclusion of additional measures of the four HRBs would have been insightful and may have uncovered some of the heterogeneity that is likely to exist within the large 'Mainstream' cluster. For example, the intensity of leisure-time physical activity could have been measured using metabolic equivalents (Conry et al., 2011; van Nieuwenhuijzen et al., 2009).

9.4 Contribution to the existing evidence

This section shows how the findings from chapters 6, 7 and 8 addressing research objectives 1, 2 and 3 respectively add stronger evidence to the existing body of literature outlined in parts 1, 2 and 3 of chapter 2, and demonstrates the novel contribution of these findings to HRB clustering research.

9.4.1 Contribution of the cross-cohort comparison of HRB clustering

The results from chapter 6 strengthen existing research findings from the appraised studies investigating HRB clustering (see section 2.2) and the two literature reviews of HRB clustering research (Meader et al., 2016; Noble et al., 2015), all of which find that HRBs are interrelated, by addressing some key methodological limitations of existing HRB-clustering research (see section 2.5). These include avoiding dichotomised HRB measures and using advanced person-centred techniques to elucidate HRB cluster patterns.

Whilst the findings between studies investigating HRB clustering are difficult to compare due to differences in methodological approach (outlined fully in section 2.2.1), the HRB patterns empirically identified in chapter 6 were suggested previously. For example, previous HRB clustering research suggests that people who smoke daily tend to have both lower fruit and vegetable consumption and lower levels of physical activity (Conry et al., 2011; De Vries et al., 2008; Laaksonen et al., 2001; Schneider et al., 2009; Schuit et al., 2002; Verger et al., 2009; Vermeulen-Smit et al., 2015), people who drink alcohol heavily are more likely to smoke (Conry et al., 2011; Schneider et al., 2009; Verger et al., 2009; Vermeulen-Smit et al., 2015) and people who smoke and drink heavily are more likely to consume fried food and less likely to consume sweet snacks (Maibach et al., 1996; Vermeulen-Smit et al., 2015).

Results from chapter 6 identified that whilst all four of the HRBs made a meaningful contribution to the HRB clusters, in both cohorts smoking contributed the most (see section 6.3.2.2). This is in keeping with empirical research that has found smoking to play a central role in the formation of HRB patterns in adulthood (Chiolero et al., 2006; Paavola et al., 2004).

The finding that HRB patterns did not differ for men and women with respect to smoking, fruit and vegetable consumption, fried food consumption and physical activity, lends empirical support to the conclusions through the critical appraisal of HRB cluster research by Meader et al. (2016), that the relationship between gender and the presence of multiple HRBs was likely to be relatively weak.

This work contributes to the existing evidence on HRB clustering by identifying both similarities and differences in HRB cluster patterns and membership across two British cohorts, providing useful insights and elucidating complexity. Detecting three HRB clusters suggests distinct patterns of HRBs are practised by different types of people providing a person-centred understanding. Largely consistent HRB cluster patterns were identified across the two cohorts for smoking, fruit and vegetable consumption, fried food consumption and physical activity, with only a slight divergence for sweet food consumption, suggesting these HRB patterns may persist across time. In regard to alcohol consumption, differences were found across the cohorts, particularly amongst women. Results indicated a convergence in alcohol consumption for men and women over time, which was consistent in previous work (Elliott et al., 2007; Keyes et al., 2011; Meng et al., 2014; Purshouse et al., 2017; Schoon and Parsons, 2003; Slade et al., 2016).

A higher membership of the 'Mainstream' cluster in the later-born cohort was found. This latter finding supports a previous comparison of these two cohorts in relation to individual HRBs (Schoon and Parsons, 2003) which found improvements in smoking and diet amongst those in the later-born cohort. These findings are also consistent with observed trends in HRBs in the United Kingdom over time. For example, a large proportion of participants in the 'Mainstream' cluster did not smoke cigarettes, corresponding with declines in the prevalence of smoking over the past 50 years (RCP, 2012). Moreover, a higher frequency of sweet food consumption in the 'Mainstream' cluster appears to coincide with global trends of increasing sugar consumption since the early 1990s (Chang et al., 2017; Singh et al., 2015; WHO, 2015).

This thesis provides some evidence of generalisability of these HRB clusters to later-born cohorts, with respect to smoking, fruit and vegetable consumption, fried food consumption and physical activity. Despite the intervening years (i.e. years between being age 33 in the NCDS (1991), age 34 in the BCS70 (2004) and age 33/34 today (2017)) resulting in later-born cohorts of mid-age adults being exposed from an earlier age to interventions that may have influenced their HRBs in mid-adulthood (i.e. smoke-free legislation implemented in 2007 (Bauld, 2011) and recent healthy lifestyle campaigns, such as 'Change4life' (DOH, 2011)), similarities in HRB cluster patterns for these HRBs across the cohorts indicate that these clusters could, to some extent, be generalised to individuals in mid-adulthood in Britain today.

Whilst it is acknowledged that these HRB clusters provide a summary of the different ways in which HRB patterns may occur in these cohorts and further replication using similar analyses in other datasets on mid-adults living in Britain today would be advantageous, the assertion of generalisability is strengthened by previous research (Graham et al., 2016; Watts et al., 2015). These studies have detected similar HRB cluster patterns to those found in this thesis using more recent information on these four HRBs in samples of adults in the United Kingdom, albeit not population-based (Graham et al., 2016; Watts et al., 2015), lending further support to the view that the distinct HRB clusters found in this thesis remain relevant. For example, whilst not directly comparable, both studies and the results of this thesis identified a cluster characterised by heavy smoking and alcohol consumption and lower intakes of fruit and vegetables and levels of physical activity. Moreover, the results of this thesis and these two studies also detected a cluster characterised by not smoking, moderate alcohol consumption, higher intakes of fruit and vegetables and higher levels of physical activity.

9.4.2 Contribution of investigating the role of pre-adolescent SEP in predicting HRB cluster membership in mid-adulthood

The cross-cohort comparison undertaken in chapter 7 lends support to the assertion made in the previous section (9.4.1), that these three distinct HRB patterns may exist in later-born cohorts. The results pertaining to the second research objective identified similarities across the two cohorts in the relationship between HRB cluster membership and social circumstances in pre-adolescence and mid-adulthood. Such similarities provide further evidence that these HRB clusters are equivalent across the cohorts and thus generalisable to those born later. Moreover, this indicates that the relationship between HRB cluster membership and social circumstances in pre-adolescence and mid-adulthood may also be generalised to those born later.

The findings presented in chapter 7 show a clear relationship, across cohort and gender groups, between disadvantaged social circumstances in mid-adulthood and an increased likelihood of membership of HRB clusters characterised by multiple negative HRBs. This is consistent with the two literature reviews of HRB clustering research (Meader et al., 2016; Noble et al., 2015) as well as a number of empirical studies (Bécue-Bertaut et al., 2008; Berrigan et al., 2003; Bondy and Rehm, 1998; Conry et al., 2011; De Vries et al., 2008; Falkstedt et al., 2016; French et al., 2008; Patterson et al., 1994; Poortinga, 2007; Schneider et al., 2009; Schuit et al., 2002; Silva et al., 2013; Slater and Flora, 1991; Vermeulen-Smit et al., 2015) that suggest a strong cross-sectional relationship between disadvantaged SEP and membership of HRB clusters characterised by multiple negative HRBs. Consequently, this work provides further evidence of distinct and complex HRB cluster patterns which are shared by individuals with similar social circumstances (Evans-Polce et al., 2016; Stapinski et al., 2016). This provides a person-centred understanding of health lifestyles and lends support to theories that suggest an interplay between social structure and individual agency (Cockerham, 2005).

The finding that pre-adolescent SEP is linked to HRB cluster membership in mid-adulthood through its influence on SEP in mid-adulthood confirms those of other studies which have found a ‘chain of risk’ (Ben-Shlomo and Kuh, 2002: 287), between child SEP, adult SEP and individual HRBs. Empirical evidence suggests that HRBs in adulthood are largely determined by SEP in adulthood which is predicted by SEP in childhood (Kamphuis et al., 2013; Paavola et al., 2004; Pudrovska and Anishkin, 2013; Yang et al., 2008). These results strengthen this existing evidence by addressing some key methodological limitations (see section 2.5). These include incorporating a range of measures to capture material, occupational and cultural dimensions of SEP and using prospectively collected data on pre-adolescent SEP.

The direct effect of pre-adolescent SEP on HRB clustering was negligible relative to the indirect effect described in the preceding paragraph. This may relate to how each HRB contributes to the derivation of the HRB clusters. For example, van de Mheen et al. (1998) found that smoking and alcohol consumption are strongly associated with adulthood SEP compared to diet and physical activity which have stronger links to childhood SEP. In this context, it may be that the hypothesised direct path from pre-adolescent SEP to HRBs for diet and physical activity may be present but overshadowed by the influence of mid-adulthood SEP on smoking which makes the greatest contribution in shaping mid-adulthood lifestyles.

The novelty of the research undertaken in chapter 7 is the application of a social determinants lifecourse framework to investigate the influence of social circumstances in pre-adolescence on HRB clustering in mid-adulthood. The results provide evidence of a total effect of pre-adolescent SEP on mid-adulthood through a social pathway. Social differentials in HRB clustering are associated with social circumstances in mid-adulthood which are strengthened through the persistence of either advantaged or disadvantaged social circumstances since pre-adolescence. This is in line with previous research, using these two cohorts, identifying that disadvantaged social circumstances in childhood tend to accumulate and therefore dictate disadvantaged social circumstances in adulthood (Anders and Dorsett, 2017; Breen and Goldthorpe, 2001; Bukodi and Goldthorpe, 2012).

Identification of a contemporaneous effect of SEP on HRB cluster membership using a uni-dimensional measure of SEP which is underpinned by prior knowledge can contribute to a better understanding of the pathways through which social circumstances in mid-adulthood influence health, which to date remains elusive (Øversveen et al., 2017). For example, the contemporaneous influence of social circumstances in mid-adulthood suggests that it is circumstantial structural factors rather than, as previously proposed, behavioural embedding processes earlier in life that determine HRB cluster membership in mid-adulthood. These mid-adulthood structural factors are conceived to influence HRB cluster membership via physical access, psychosocial stress and social group habitus.

For example, both 'Risky' and 'Moderate Smokers' clusters are characterised by lower levels of physical activity and consumption of fruit and vegetables in comparison to the 'Mainstream' cluster. This may be the result of differentials in purchasing power dictating the ability to buy fresh fruit and vegetables (Jones et al., 2014; Kelly et al., 2016; Pechey et al., 2013) and participate in leisure-time physical activity (Beenackers et al., 2012; Chinn et al., 1999; Kelly et al., 2016; Parry, 2013). Similarly, membership of the 'Risky' and 'Moderate Smokers' clusters implies smoking and heavier alcohol consumption which could occur through exposure to material hardship (Hoek and Smith, 2016; Lindström et al., 2013; Pampel et al., 2010; Twyman et al., 2016) or job strain (Brunner et al., 2007; Heikkilä et al., 2012; Kivimäki et al., 2013; Lallukka et al., 2008; Nyberg et al., 2015). Moreover, both positive and negative HRBs may be compounded through frequent interactions with others who are experiencing similar social circumstances, thus making these HRBs socially normative (Bann et al., 2016; Bartley, 2016a) and linking them to an individual's identity as well as their overall lifestyle (Fennis et al., 2015; Geronimus et al., 2016; Hedegaard, 2016; Kelly et al., 2016; Reid et al., 2010).

9.4.3 Contribution of examining transitions in HRB clustering over time and the role of mid-adulthood SEP

The results presented in chapter 8 lend support to previous arguments that assuming HRBs to be stable during mid-adulthood may not be valid (Laaksonen et al., 2002; Oude Groeniger and van Lenthe, 2016). Whilst the results are consistent with existing findings that these four HRBs are relatively stable in mid-life (Benzies et al., 2008; Mulder et al., 1998), this stability was not found to be absolute, there being a substantial shift to the 'Mainstream' cluster (see section 8.3.3.5). This finding is consistent with other studies demonstrating positive change in HRBs during mid-life (Backett and Davison, 1995; Benzies et al., 2008; Mulder et al., 1998) and suggests that assumptions regarding stability of HRBs during mid-life do not necessarily hold and should be challenged.

Again, these findings are considered to strengthen existing research evidence by considering HRB clustering, as opposed to individual HRBs, and using advanced person-centred techniques to elucidate HRB cluster patterns over time.

Whilst a transition in HRB cluster membership during mid-life does not necessitate an individual changing all four HRBs, movement to a cluster characterised by more positive HRBs in comparison from the one left behind is consistent with research suggesting that, on average, individuals tend to improve their HRBs during mid-adulthood (Artaud et al., 2016; Backett and Davison, 1995; Benzies et al., 2008; Mulder et al., 1998; Sijtsma et al., 2012; Britton et al., 2015). Furthermore, these results reflect HRB trends observed in the UK population since the 1990s, such as decreases in smoking prevalence seen over the last 50 years (RCP, 2012), increases in fruit and vegetable consumption and reductions in fat intake (Ezzati et al., 2015). These findings are also consistent with general increases in leisure-time physical activity seen across developed countries since the 1990s (An et al., 2016; Stamatakis et al., 2007).

This shift to the 'Mainstream' cluster, characterised by a higher frequency of sweet food consumption in comparison to the other two clusters, implies the average frequency of sweet food consumption increased. This is consistent with evidence of increasing sugar consumption since the early 1990s (Chang et al., 2017; Singh et al., 2015; WHO, 2015). Consequently, increases in the frequency of sweet food consumption identified in the 'Mainstream' cluster may reflect a period effect. For example, Sijtsma et al. (2012) found that, whilst age tended to be associated with improved diet quality, age-adjusted time trends showed a general deterioration in diet quality over time.

However, general increases in sweet food consumption, indicated by a shift to the 'Mainstream' cluster contrasts with research suggesting no change in sweet snack consumption amongst mid-age adults (Mulder et al., 1998). Furthermore, Parsons et al. (2006) found, in the same cohort, that between ages 33 and 42 the frequency in which 'sweets or chocolate' and 'biscuits' were consumed decreased, as did fried food consumption. Disparities may be due to differences in measurement, thus making comparability difficult. For example, in chapter 8 fried food and sweet food consumption were both captured based upon derived measures that combined 'sweets and chocolate' and 'biscuits' to become sweet foods and 'foods fried in fat' and 'chips' to become fried foods (see section 5.2.1.1), whereas Parsons et al. (2006) examined these diet variables separately. Moreover, Mulder et al. (1998) captured sugar consumption as eating more than four sweet snacks during the day.

On the other hand, such disparities between the results of chapter 8 in relation to fried food and sweet food consumption and those of Mulder et al. (1998) and Parsons et al. (2006) may be explained by a lack of consideration by these studies to the ways in which HRBs interrelate. For example, these studies do not focus on HRB clustering and in consequence they do not consider the relationship which has been found to exist between smoking and sugar consumption (Crawley and While, 1996; Méjean et al., 2011; O'Doherty et al., 2011; Whichelow et al., 1991). Research has found that smokers and heavy drinkers are less likely to consume sweet snacks (Maibach et al., 1996; Vermeulen-Smit et al., 2015) and are more likely to prefer fatty and salty tastes (Lampure et al., 2014), in comparison to those who do not smoke and do not exceed recommended levels of alcohol consumption. Smoking may have an effect on dietary preferences (Colditz et al., 1991; Lampure et al., 2014), by influencing taste buds (Iredale et al., 2016), while the inverse pattern of alcohol consumption and sugar may reflect a replacement of sugar intake with alcohol use (Colditz et al., 1991).

In light of the above, I would suggest that sugar consumption has increased over time and is interrelated with reductions in smoking and heavy alcohol consumption. This demonstrates how consideration of the ways in which HRBs interrelate over time may provide a richer and person-centred perspective of lifestyles in mid-adulthood compared to treating them as non-related entities (Buck and Frosini, 2012; Filippidis et al., 2016; Noble et al., 2015).

This research work has also uncovered gender differences in HRB clustering over time which provides a meaningful contribution to the existing evidence base. For example, the shift to the 'Mainstream' cluster between ages 33 and 42 also has implications for alcohol use amongst women. The findings suggest that some women may increase their alcohol consumption from drinking no alcohol to drinking within recommended limits during mid-life. This finding offers empirical support to the theoretical assumption that women's alcohol consumption increases as their children grow older, though it would be in the range of low to moderate levels (Tran et al., 2015).

The sensitivity analyses pertaining to the nature of the HRB clusters at each time point (i.e. measurement invariance, see section 9.5.3) also elucidated gender differences. This indicated that for members of the 'Risky' cluster changes in fried food consumption for men and leisure-time physical activity for women extended beyond the interrelationship of the four HRBs.

The finding of reduced fried food intake for men that could not be fully explained by the interrelationship between the HRBs mimics general decreases in fried food consumption during this period. For example, analysis using data on British food purchasing and consumption suggests that since the 1980s there has been a reduction in the purchase of traditional fats, such as butter (Leicester, 2004), and fat intake (Allender et al., 2006; Prentice and Jebb, 1995). During this period, there has also been an increased availability of low-fat alternatives (Rolls and Miller, 1997). Findings from a sample of US adults showed that over a 20-year period there were reductions in butter and fried potatoes (Sijtsma et al., 2012). Although not directly comparable, these US trends in diet provide some indication of future UK trends, given that the UK is considered to be 10–15 years behind the US with respect to obesity (McAuley et al., 2016) which is closely related to fat intake (Drewnowski et al., 1992). The 'partial' LTA models outlined in chapter 8 (see section 8.3.3.4) suggest fried food decreased only in men but not women. This may relate to women having more consistent and lower overall consumption of fried foods than men to begin with. It is argued that women have been subject to higher levels of social pressure to be slim for a longer time period in comparison to men whose bodies have become increasingly objectivised only recently (Schuster et al., 2013).

Consequently, any changes in fried food consumption were adequately captured by the 'Risky' clustered patterns for women, whereas for men, increased avoidance of fried foods in this 'Risky' cluster may have been more recent and consequently changes in their consumption could not be fully captured by the clustered patterns. Moreover, members of the 'Mainstream' cluster had higher frequencies of fried food consumption in comparison to men in the other two clusters; therefore, whilst change may have occurred for all men, it could have been more dramatic for those in this smallest cluster and therefore better detected in the analyses.

By contrast, amongst women in the 'Risky' cluster, leisure-time physical activity appeared to decrease for the 'Risky' cluster and this change could not be fully explained by the interrelationship of the four HRBs. Research suggests that whilst physical activity tends to reduce during mid-life for both genders (Allender et al., 2008b; Artaud et al., 2016; Corder et al., 2009; Mulder et al., 1998; Wannamethee et al., 1998) decreases may be more likely amongst women (Hunt et al., 2001; Parsons et al., 2006). This is considered to be primarily due to a scarcity of free time, which is one often cited reason for physical inactivity (Kelly et al., 2016). Whilst increased responsibilities in mid-life impact both men and women (Lachman et al., 2015), these may disproportionately affect women due to their increased participation in the labour market in recent decades whilst continuing to undertake a larger share of caring responsibilities and domestic chores (Chou et al., 2004; Nomaguchi and Bianchi, 2004).

Moreover, the findings from chapter 8 suggest that those in more disadvantaged social circumstances were more likely to be members of the 'Risky' cluster compared to the 'Mainstream' cluster. Therefore, on average, women in the 'Risky' cluster were more socially disadvantaged. Socially disadvantaged women are more likely to have caring responsibilities and associated household chores in comparison to their more socially advantaged peers because they are unable to afford child care or care homes for elderly relatives (Bruhn and Rebach, 2014). In consequence, additional responsibilities implicated with mid-life (Lachman et al., 2015) may be particularly burdensome for disadvantaged women, thus further restricting their time and their ability to undertake leisure-time physical activity.

The findings provide a meaningful contribution to HRB cluster research by identifying a strong cross-sectional association between social circumstances in mid-adulthood and HRB cluster membership at the same age yet a lack of effect on change thereafter. The detection of a proximal relationship between SEP and HRB cluster membership lends support to the findings of chapter 7, which also found a contemporaneous effect of mid-adulthood SEP on HRB clustering, relating to the social circumstances in which people live their lives as opposed to embedded behavioural processes.

The finding of a strong contemporaneous effect of social circumstances on HRB cluster membership at age 33, but not change, implies that SEP may shape lifestyles at the beginning of mid-adulthood which remained stable thereafter for reasons that extend beyond social circumstances. For example, as mentioned in the preceding section (section 9.4.2), the contemporaneous effect of SEP on HRB cluster membership might occur through physical access, psychosocial stress and social group habitus mechanisms. However, it may be that upon reaching mid-life there is little chance for HRB change due to the adoption of adult roles, such as raising children and employment (Benzies et al., 2008; Mulder et al., 1998), resulting in increased regularity and routine (Backett and Davison, 1995), through which health lifestyles become embedded (Blue et al., 2014). Following the embedding of these lifestyles, other factors unrelated to SEP may dictate their persistence during mid-life. As mentioned above, increased demands placed upon individuals in mid-adulthood (Lachman et al., 2015) are likely to lead to a lack of time and energy. These are both considered barriers to effective HRB change (Kelly et al., 2016), reducing motivation and capacity to alter HRB patterns (Borland, 2013b), thus contributing to the stability of HRB patterns for a large proportion of mid-age adults, that were shaped by social circumstances in early mid-adulthood.

Therefore, physical access, psychosocial stress and social group habitus may play a role in embedding lifestyles as individuals enter the third decade of life, and factors such as time and motivation, posing barriers to a wider group of mid-age adults, could be responsible for their continuation thereafter. Consequently, the inverse may explain the smaller proportion whose HRBs do appear to change during mid-life, tending to be in a positive direction. For example, two factors, more free time and an awareness of ageing, are more commonly associated with older age groups and may explain increases in physical activity found amongst retirees in comparison to those of working age (Feng et al., 2016; Menai et al., 2014). It may be the case that the small proportion who change HRB cluster membership during mid-life have higher than average amounts of free time in comparison to their peers, perhaps due to an absence of caring and/or employment demands. This may allow them greater opportunities to engage in positive HRBs, such as physical activity and diet (Kelly et al., 2016), and increased capacity to refrain from negative HRBs, such as smoking and alcohol consumption (Borland, 2013b). On the other hand, motivation to change HRBs in mid-life for this minority of individuals may be rooted in their premature awareness of ageing and the consequences of lifestyle on health subsequently leading to higher levels of dedication to change HRBs (Backett and Davison, 1995).

Another explanation for the lack of effect of social circumstances on HRB change may relate, once again, to identity. Scholars suggest that as humans we strive to maintain our identities (Fennis et al., 2015) which are strongly tied to both lifestyles (Hedegaard, 2016; Pampel et al., 2010) and SEP group identity (Bartley, 2016a; Pampel et al., 2010). Adult lifestyles are thought to fulfil self-identity and meaning, and reflect the consumption patterns of particular social groups (Blue et al., 2014; Cockerham, 2005). As a result of occupational maturity and the economic returns associated with mid-life (Bukodi and Goldthorpe, 2011; Herd et al., 2007), the attainment of social position is likely to have been achieved by early mid-adulthood which will be tied to a particular identity and lifestyle (Fennis et al., 2015; Geronimus et al., 2016; Hedegaard, 2016; Kelly et al., 2016; Reid et al., 2010). The solidification of a mid-adulthood identity in the third decade of life will result in lifestyle stability during mid-adulthood with evidence suggesting that identity is a barrier to HRB change (Kelly et al., 2016).

This influence of identity, proposed to partially explain a lack of effect of social circumstances on HRB change, is consistent with previous work which suggests a ‘lock in’ of HRBs according to social circumstances, captured by social class, for alcohol and smoking over the lifecourse and to a lesser extent physical activity (Jones et al., 2011). As mentioned in chapter 2 (see section 2.3.2), it is theorised that social differentials in the persistence of health-damaging HRBs, such as smoking, despite awareness of the harms of smoking (Blaxter, 1990), are the product of unequal 'power and privilege' (Geronimus et al., 2016) and tied to an identity of non-conformity and toughness amongst some social groups (Pampel et al., 2010). This 'lock in' effect (Jones et al., 2011) may partially explain higher proportions of individuals who continue to smoke within more disadvantaged SEP groups when compared to the rest of the population (Jefferis et al., 2004).

However, there is compelling evidence of SEP differentials in multiple HRB change among adults over time (Buck and Frosini, 2012; Ding et al., 2015). It may therefore be the case that SEP does play some role in HRB change during mid-life but, given the relatively small number of participants who do change their HRBs, there is a lack of statistical power to detect an effect. Consequently, the existence of an effect between social circumstances in early mid-life and change in HRB cluster membership during the subsequent nine years cannot be completely ruled out.

9.5 Further insights from the research findings and methodology

Sections 9.1 to 9.4.3 have discussed the main findings and contributions of chapters 6, 7 and 8, addressing research objectives 1, 2 and 3 respectively. This section consolidates the thesis findings by further linking the chapters to one another.

9.5.1 The possible period effects for sweet food and alcohol consumption

As mentioned in section 1.4, there is a period of overlap across the two cohorts. The period covers the early 2000s, when BCS70 participants were aged 30 (2000) and 34 (2004) and NCDS participants were aged 42 (2000). Consequently, similarities in findings at these ages may, to some extent, elucidate period effects. The cohort comparison of participants of the same age, undertaken in chapter 6, controls for age and highlights a combination of period and cohort effects. Chapter 8 used information from one cohort at two different ages, controlling for cohort effects and therefore elucidating the combined effects of period and age (Schoon, 2006).

By identifying general increases in moderate alcohol consumption and sweet food consumption in both chapters 6 and 8, it could be deduced that these changes are due to period effects. The results of chapters 6 and 8 identified shifts in membership of the 'Mainstream' cluster. In chapter 6, there was a higher membership of the 'Mainstream' cluster in the later-born cohort at age 34 when compared to the earlier-born cohort at age 33. In chapter 8, there was movement from the 'Risky' and 'Moderate Smokers' clusters to the 'Mainstream' cluster between ages 33 and 42.

Whilst increased membership of the 'Mainstream' cluster implies overall improvements for some HRBs (i.e. members were non-smokers, had higher frequencies of fruit and vegetable consumption, lower frequencies of fried food consumption and higher levels of leisure-time physical activity, compared to members of the 'Risky' and 'Moderate Smokers' clusters), an increased membership of the 'Mainstream' cluster has possible implications for the frequency of sweet food consumption, tending to be higher in this cluster compared to the other two. Moreover, a shift to the 'Mainstream' cluster indicates general increases in moderate alcohol consumption and for women in the later-born cohort drinking above recommended limits.

The potential period effects for increased moderate alcohol consumption and a higher frequency of sweet food consumption may relate to declines in the prevalence of smoking which both cohorts have been exposed to over the past 50 years (RCP, 2012). As mentioned in section 9.4.3, there is evidence of an association between smoking and sugar consumption (Maibach et al., 1996; Vermeulen-Smit et al., 2015; Méjean et al., 2011). Furthermore, research investigating alcohol typologies over time (see section 2.4.2) found that, whilst the prevalence of smoking was higher in drinking typologies characterised by heavier alcohol consumption, smoking reduced between 1978 and 2010 for all typologies (Purshouse et al., 2017).

Recent research (conducted in 2014–15) suggests that the majority of people have a 'vice' (described as a minor bad habit) and that eating sweets is amongst the most commonly reported (Dale et al., 2016). Moreover, research suggests individuals 'balance out' positive and negative HRBs (Backett and Davison, 1995) through compensatory health beliefs (Knäuper et al., 2004) (i.e. individuals compensate for participating in some negative HRBs by partaking in some positive ones). In light of this, it may be that smoking was a common 'vice' (Dale et al., 2016) in the past, but that, as a result of public health efforts, individuals have subsequently quit smoking and replaced it with a new 'vice' (Dale et al., 2016) of sweet food. This replacement may be an unintended consequence of anti-smoking campaigns (Chou et al., 2004).

At the same time, it may be that for later-born women this sweet food 'vice' (Dale et al., 2016) may have been replaced once more by heavier alcohol consumption. For example, whilst still high compared to the other HRB clusters, the lower frequency of sweet food consumption in the ‘Mainstream’ cluster for women in the later-born cohort compared to the earlier may reflect a replacement of sugar intake with alcohol use amongst women which has been observed elsewhere (Colditz et al., 1991). Heavier alcohol use amongst women born later may in turn reduce their sweet food consumption by satisfying the need for sugar through the consumption of sugar-laden and calorie-dense alcoholic beverages (Sayon-Orea et al., 2011) or via compensatory health beliefs (Knäuper et al., 2004) mentioned above or due to changes in dietary preferences that have been found to occur amongst heavier drinkers (Lampure et al., 2014).

A general preference for heavier alcohol consumption over sweet food consumption observed for women members of the ‘Mainstream’ cluster in the later-born cohort compared to their predecessors may relate to delays in childbirth and stronger ties to the labour market increasing disposable income and the opportunity to consume alcohol (Keyes et al., 2011; Slade et al., 2016). By contrast, increases in moderate consumption between ages 33 and 42 for women born in 1958 may be due to having fewer childcare responsibilities as their children have become older (Tran et al., 2015).

It is not possible from my results to isolate a period effect for alcohol and sweet food consumption. However, for women in both cohorts, alcohol consumption has become socially permissible due to changes in gender relations (Keyes et al., 2011; Slade et al., 2016) and increased exposure to alcohol product marketing and advertising aimed specifically at women (Bosque-Prous et al., 2015; Keyes et al., 2011; Purshouse et al., 2017). These factors correspond to trends in women consuming more alcohol over time (Britton et al., 2015; Davies, 2016; Meng et al., 2014; Purshouse et al., 2017) and converging with men’s alcohol consumption (Elliott et al., 2007; Keyes et al., 2011; Meng et al., 2014; Purshouse et al., 2017; Schoon and Parsons, 2003; Slade et al., 2016). Moreover, it is predicted that moderate alcohol consumption among individuals born before 1980 is set to increase (Kraus et al., 2015; Meng et al., 2014; Molander et al., 2010). Additionally, increases in sugar consumption are in line with global trends (Chang et al., 2017; Singh et al., 2015; WHO, 2015) and age-adjusted time trends that have shown a general deterioration in diet quality over time (Sijtsma et al., 2012).

In sum, it is plausible that these trends in alcohol consumption and sweet food consumption, captured in the findings of chapters 6 and 8, may be due to an overlap in the period of life shared by the cohorts (Britton et al., 2015). The policy implications of these possible period effects on alcohol and sweet food consumption are elaborated upon in section 9.6 of this chapter.

9.5.2 The subjectivity of latent variable model selection

Chapters 6 and 8 highlighted the issue of subjectivity in selecting the optimal number of clusters amongst studies using data reduction techniques (see sections 6.2.3 and 8.2.3). Whilst these data reduction techniques, used to elucidate HRB clusters, have a number of advantages in comparison to other methods (McAloney et al., 2013), this subjectivity in model selection has been raised as a potential limitation (Martinez et al., 1998; McAloney et al., 2013).

Consequently, both chapters 6 and 8 sought to minimise the potential for subjectivity in model selection. The replication of the analysis within two populations in chapter 6 addressed concerns that subjectivity in model selection may reduce comparability across population samples (Martinez et al., 1998). Furthermore, comparing exploratory and confirmatory approaches to LTA model selection in chapter 8 demonstrated that, whilst latent variable modelling is a data-driven method, model selection should also be guided by prior knowledge (Finch and Bronk, 2011) and relevant theory (Nylund et al., 2007; Wang and Wang, 2012).

9.5.3 The investigation of measurement invariance

The measurement invariance analysis, used to ascertain latent cluster equivalence across known subgroups (see section 5.3.1.1), highlights the salience of thorough examinations of measurement invariance over time when using latent variable models (Collins and Lanza, 2010; Finch, 2015). Measurement invariance analysis was employed in chapter 6 to determine whether the nature of the three HRB clusters – 'Risky', 'Moderate Smokers' and 'Mainstream' – was similar across the two birth cohort studies. In chapter 8, measurement invariance analysis was used to determine HRB cluster equivalence over time (i.e. ages 33 and 42) in the earlier-born cohort. Such investigations are crucial when there is theoretical reason to believe that the nature of the HRB clustered patterns are either likely to be the same or likely to be different according to known subgroups in the populations under investigation (Finch, 2015) or at different points in time (Collins and Lanza, 2010).

In both chapters 6 and 8, an assumption of full measurement invariance could not be verified; instead, partial measurement invariance was found in these latent variable models. In chapter 6, HRB cluster patterns were somewhat equivalent across the two cohorts within each gender, except in relation to alcohol consumption, particularly amongst women. The detection of partial measurement invariance in these latent variable models was important in the interpretation of the models in two respects. Firstly, as mentioned above, it allowed for some generalisations of these HRB cluster patterns to later-born cohorts in mid-adulthood today. Secondly, identifying differences in relation to women's drinking patterns across the two cohorts provided useful insights into gender convergence in alcohol consumption amongst those born in 1970, which may have been missed otherwise.

In chapter 8, full measurement invariance over time could not be reasonably assumed for the ‘Risky’ cluster. Certain HRBs (fried food consumption amongst men and leisure-time physical activity amongst women) were found to substantially differ at each age. These partial measurement invariant models provided a better representation of the extent to which HRB cluster membership changed between ages 33 and 42. These models indicated that changes observed for these specific HRBS over time extended beyond the interrelationship of the four HRBs.

9.6 Policy implications

The policy implications of this doctoral project are far reaching and can contribute to the development of policies and interventions that seek to improve lifestyles in mid-adulthood. It is clear that a joint approach is required in order to improve HRBs, consisting of policies and interventions aimed at the individual level, known as ‘downstream’, and those aimed at a population level, known as ‘upstream’ (Short and Mollborn, 2015).

Whilst downstream interventions are considered to be effective in improving the HRBs of individuals (Bambra et al., 2015; Okechukwu et al., 2014), they have been criticised for placing emphasis on individual responsibility and agency and perpetuating social differentials in HRBs (Katikireddi et al., 2013). Individual-level interventions have the potential to be inequitable because more disadvantaged groups are less able to effectively engage in and respond to HRB interventions in comparison with their more advantaged peers. Individuals experiencing more disadvantaged social circumstances may have diminished motivation and capacity to change their lifestyle as a result of disadvantaged social circumstances – e.g. a greater amount of time and energy is spent on managing financial difficulties (Borland, 2013c).

Consequently, upstream interventions can serve to mitigate social inequalities that are associated with downstream interventions, by tackling the ‘causes of the causes’ (Marmot, 2005: 1101) of health-damaging HRBs in mid-adulthood. Moreover, upstream interventions target populations as opposed to high-risk groups, reaching a larger number of people simultaneously and are thus a more powerful strategy to improve HRBs (Rose, 1989).

However, upstream interventions are difficult to implement, requiring political will and collaborations that extend beyond the remit of health into a wide range of areas, such as education, economics, housing, town planning, food and agriculture and social services (Catford, 2006; Hayward and Gorman, 2004; Koh et al., 2011).

The policy implications of this work are discussed according to these two approaches. This section first considers how the findings of this thesis can inform downstream interventions to improve HRBs before turning attention to upstream interventions. The final section discusses how these two policy options may be combined to create a joint approach known as 'Proportionate Universalism' (Marmot, 2010: 16).

9.6.1 Downstream interventions

9.6.1.1 *Multiple HRB change*

The findings from chapters 6 and 8 of this thesis combined with the results of previous HRB clustering research (see section 2.2) provide strong indications that smoking, alcohol, diet and physical activity interrelate and tend to cluster. Therefore, these findings support the assertion made by others (Buck and Frosini, 2012; Filippidis et al., 2016; Noble et al., 2015) that the current policy approach, which tackles these four HRBs as individual entities, may not be appropriate. Consideration of the ways in which HRBs cluster together can be used to refine existing HRB policies and interventions, by making them more person-centred. Consistent HRB cluster patterns should be considered alongside patterns that differ according to subgroups, to develop public health interventions that reflect the individual's experience of HRB (Buck and Frosini, 2012; Watts et al., 2015).

Such 'downstream' (Short and Mollborn, 2015) interventions that focus on lifestyle (Ashra et al., 2015; Younge et al., 2015) could have long-term benefits for health outcomes. As mentioned in section 1.1, the presence of multiple negative HRBs has been found to be strongly related to the risks of both mortality (Khaw et al., 2008; Kvaavik et al., 2010; Martin-Diener et al., 2014) and morbidity (Alageel et al., 2016; Artaud et al., 2016; Chow et al., 2010; Sabia et al., 2012).

Whilst a review of studies evaluating the efficacy of multiple HRB interventions found that research in this area is limited (King et al., 2015) and the comparability of the samples from the reviewed studies with the two birth cohorts used in my doctoral project is questionable, research does suggest that interventions addressing multiple HRBs simultaneously may be more successful (Ashra et al., 2015; Goldstein et al., 2004; Hale et al., 2014; Hyman et al., 2007; King et al., 2013; Koshy et al., 2012; Nigg et al., 2002; Steptoe, 2007) and cost effective (Prochaska et al., 2008b) than targeting HRBs independently. However, it is argued that attention must also be paid to the content of these HRB interventions as much as the number of HRBs that they target (Mc Sharry et al., 2015). Such interventions are based upon the premise of 'behavioural spillover' (Poortinga et al., 2013) whereby change in one HRB can lead to change in others (Prochaska et al., 2008a).

The findings from chapters 6 and 8 reinforce this evidence and could inform practices to improve HRBs by developing person-centred care plans based on the three distinct clustered patterns of HRBs found in this work. For example, the HRB cluster patterns suggest that members in the two clusters characterised by smoking (i.e. 'Risky' and 'Moderate Smokers') tended to consume fruits and vegetables less frequently, consume fried food more frequently and do less physical activity than members in the cluster characterised by non-smoking (i.e. 'Mainstream'). Therefore, any interventions to quit smoking could employ person-centred strategies to improve diets and exercise, in order to achieve a healthier lifestyle.

At the same time, members of the 'Risky' cluster smoke more cigarettes per day and have a lower probability of transitioning to the 'Mainstream' cluster in comparison to the 'Moderate Smokers' cluster, whose other HRBs are more aligned with the 'Mainstream' cluster. Moreover, the findings demonstrate that members of the 'Risky' cluster are more socially disadvantaged than members of the 'Moderate Smokers' and 'Mainstream' clusters.

This implies that members of the 'Risky' cluster may be experiencing higher levels of nicotine addiction which is interrelated with other aspects of their lifestyle, i.e. diet, physical activity and alcohol consumption. Given their more disadvantaged social circumstances, they may be less able to respond to traditional downstream interventions (e.g. smoking cessation) which do not resonate with their everyday experience of HRBs or their social circumstances. In comparison, members of the 'Moderate Smokers' cluster, who are more socially advantaged, are already making positive changes to their HRBs.

Consequently, members of the 'Risky' cluster would benefit from targeted lifestyle person-centred interventions which take into account how their contemporaneous social circumstances may undermine their ability to change negative HRBs and start to unpick how these HRBs interrelate. On the basis of these intensive consultations, the clinician and the individual can together develop a realistic person-centred care plan which resonates with the individual's everyday experience of HRBs. By contrast, members of the 'Moderate Smokers' cluster may not need such specialist support; instead, they will have a greater capacity to make positive lifestyle changes and be motivated at the individual level by a brief intervention with a clinician or through their growing awareness of the impact of negative HRBs on their health (e.g. increased breathlessness when walking upstairs).

9.6.2 Upstream interventions

As demonstrated above, the findings of this thesis can contribute to the development of downstream interventions to improve negative HRBs at the individual level, particularly amongst the most socially disadvantaged who are disproportionately members of the 'Risky' cluster. However, this thesis points implies that wider social circumstances also influence lifestyles of members of the 'Moderate Smokers' cluster. Furthermore, the findings imply that a population strategy is required in order to improve HRBs for members of the 'Mainstream' cluster. Therefore, emphasis must also be placed on upstream interventions in order to tackle multiple negative HRBs at a population level.

9.6.2.1 *Social differentials in HRB patterns*

The findings highlight social differentials in HRB cluster membership, providing a person-centred understanding of which subgroups of the population are likely to share clustered patterns of multiple negative HRBs. Investigating the relationship between social determinants across the lifespan and HRB clustering provides insights into the unique social context in which different social groups experience HRBs in their everyday lives and consequently deepens our understanding of how SEP differentials in mid-adulthood lifestyle contribute to health inequalities previously observed in these two cohort studies (Ferri et al., 2003; Mensah and Hobcraft, 2008; Parsons et al., 2013).

This thesis demonstrates a contemporaneous effect of mid-adulthood SEP on HRB cluster membership, which is strongly influenced by SEP in pre-adolescence. In consequence, it is imperative that policies and interventions designed to tackle HRBs consider both the accumulation of resources across the lifecourse (Hayward and Gorman, 2004) and the proximal social circumstances in which adults live their lives (Short and Mollborn, 2015). This implies that a population strategy is required, consisting of 'upstream' policies and interventions that focus on the social structure (Cleland et al., 2010; McCartney et al., 2011; Okechukwu et al., 2014; Short and Mollborn, 2015).

The findings presented in chapter 7 suggest a 'chain of risk' (Ben-Shlomo and Kuh, 2002: 287), with pre-adolescent SEP predicting mid-adulthood SEP, consisting of material, occupational and cultural dimensions, which has a contemporaneous influence on HRB cluster membership, conceived to occur via physical access, psychosocial stress and social group habitus mechanisms. Considering the total effect, whilst identifying the contribution of the indirect effect, provides an insight into the ways in which policymakers can intervene (Green and Popham, 2016). The first part of the chain points to an accumulation of resources between pre-adolescence and mid-adulthood, while the second part of the chain identifies the contemporaneous effect of social circumstances in mid-adulthood on lifestyle at the same age. Thus, effective policies and interventions seeking to break these links in the chain may prevent the formation of clustered patterns of health-damaging HRBs in mid-adulthood.

For example, the first link in the chain implies that policies and interventions are required to mitigate differentials in the accumulation of resources between childhood and adulthood (Cohen et al., 2010) and thus 'give every child the best start in life' (Marmot, 2010: 15). The most disadvantaged children with little or no access to material and cultural resources in early life are considered to have the least chance of accruing them later (Heinz and Marshall, 2003). Therefore, policies to ensure access to resources that meet a sufficient threshold for all children are important. For example, whilst education is free to all children, which can aid social levelling (Georg, 2016), inequalities do exist in the British schooling system (Singleton et al., 2011; Smithers and Robinson, 2010).

Policies to improve material and cultural resources in childhood are set out in the Marmot review entitled 'Fair Society, Healthy Lives' (Marmot, 2010). Material aspects include providing families with a living wage, good quality housing and generous welfare provisions. Cultural aspects include access to good quality education and apprenticeships. Such policies equip children with skills and resources, setting them on more advantaged SEP trajectories (Cohen et al., 2010). This equates to higher material, occupational and cultural dimensions of SEP in adulthood and in turn improved HRBs, representing the second link in the chain.

These mechanisms reflect the contemporaneous influence of social circumstances on lifestyles, driving everyday decisions on whether to participate in or refrain from particular HRBs (Borland, 2013c). This implies that HRBs in mid-adulthood can be modified through contemporaneous social factors rather than social circumstances in childhood, which is promising from an interventionist perspective. Altering physical access, psychosocial stress and social group habitus processes may be one way in which policymakers can intervene to improve HRB cluster patterns.

'Upstream' policies tackling psychosocial stress and physical access mechanisms may be enacted through legislation and taxation. A package of 'upstream' policies and interventions to improve HRB clustering in mid-adulthood is required in order to maximise their efficacy

(Okechukwu et al., 2014) and counteract potential unintended consequences associated with different approaches (Frohlich and Potvin, 2008). Empirical evidence demonstrates that taxation (Bhimjiyani et al., 2016; Blakely et al., 2015; Burton et al., 2016; Dragone et al., 2015; Hu et al., 2016; McAuley et al., 2016; McGill et al., 2015; Meier et al., 2010) and legislation (Bauld, 2011; Patterson and Elinder, 2015; Robertson-Wilson et al., 2012) are effective in tackling HRBs at a population level.

For example, workplaces that guarantee decent employment conditions and a fair wage to workers and welfare provisions that ensure all people, including the most vulnerable, have an adequate income to meet basic needs are recommended (Marmot, 2010). These are likely to reduce sources of occupational and economic psychosocial stress thought to lead to the perpetuation of smoking (Garrett et al., 2015; Hoek and Smith, 2016; Pampel et al., 2010) and alcohol dependence (Head et al., 2004) amongst those in more disadvantaged social circumstances. Moreover, research suggests that workplace policies that promote better psychosocial work environments and challenge negative beliefs regarding physical activity may improve the uptake of exercise amongst workers (Andersen, 2011).

Revenue incurred through taxation on foods high in sugar and fat could be used to increase access to health-promoting HRBs, such as healthier diets and physical activity (Dragone et al., 2015). For example, this revenue could subsidise other products, such as fresh fruit and vegetables (Cobiac et al., 2017), which are often too expensive for those experiencing financial hardship (Jones et al., 2014; McGill et al., 2015). Funds could be used to provide local and low-cost opportunities for physical activity, given that a lack of money and transport are two frequently cited barriers to a lack of participation in exercise for those in more disadvantaged social circumstances (Burton et al., 2012; Chinn et al., 1999; Kelly et al., 2016). Such activities could be better tailored at a local level according to the preferences of the beneficiaries (Burton et al., 2012; Hunt et al., 2001). For example, team-based physical activities that incorporate a social element alongside exercise have been found to be preferred by those with lower compared to higher incomes (Burton et al., 2012).

Whilst effective, both taxation and legislation have caveats that should be acknowledged and mitigated. Taxation on smoking and alcohol can be regressive, disproportionately affecting those at the lower end of the SEP distribution (Blakely et al., 2015; Dragone et al., 2015; McGill et al., 2015) and may have implications for multiple HRBs. For example, individuals in financial difficulties may have legitimate reasons, given their social circumstances, to be less concerned with their long-term health and thus partake in health-damaging HRBs (Birch, 2010). Yet by doing so, they spend more of their income on cigarettes and alcohol which further restricts their ability to buy healthier food and take part in physical activity. This point demonstrates the need to take a person-centred approach in the development of policies and interventions to improve HRBs.

Moreover, social differentials could provide further insight into whether declines in smoking are related to increases in obesity which are thus an unintended consequence of anti-smoking campaigns (Chou et al., 2004). As mentioned above (see section 9.5.1), it is suggested that smoking has been replaced by a higher calorie intake (Chou et al., 2004). At the same time, others suggest that such campaigns have in fact led to reductions in obesity by improving diets and physical activity (Dragone et al., 2015). It may be that those in more advantaged social circumstances replace smoking with positive HRBs, whereas those in less advantaged social circumstances replace smoking with negative HRBs. For example, research investigating HRB clustering in Australian older adults (Griffin et al., 2014) identified an ex-smoker cluster with more negative HRBs, labelled 'higher risk ex-smokers' and an ex-smoker cluster with fewer negative HRBs, labelled 'lower risk ex-smokers'. Whilst both clusters were found to have lower levels of educational attainment compared to the reference cluster, labelled 'active non-smokers', the estimates from this study indicate that the 'higher risk ex-smokers' were less educated compared to the 'lower risk ex-smokers'. The 'lower risk ex-smokers' cluster was found to be more physically active than the 'higher risk ex-smokers'. At the same time, the 'higher risk ex-smokers' had poorer diets than the 'lower risk ex-smokers'.

'Upstream' policies, such as those involving taxation and legislation, are considered important levers with which to improve HRBs in the population but only in the presence of changes to social norms (Blue et al., 2014; Hargreaves et al., 2010). There are a number of strategies that may be used to change social norms at the population level, conceived to influence HRB change via influencing social environments and the spread of positive HRBs from one individual to another through behavioural modelling (Perry et al., 2016) and diffusion (Dixon and Banwell, 2009).

However, the extent to which public health campaigns, such as raising awareness of recent changes to alcohol guidelines (DOH, 2016), are effective depends on sustained activity and the ability to effectively target all social groups (Holmes et al., 2016). Consideration of the unique social context in which different social groups experience HRBs in their everyday lives is also required for public health messages to be relevant and resonant (Buck and Frosini, 2012; Garrett et al., 2015; Reid et al., 2010; Watts et al., 2015). Sophisticated messages may serve to counteract the ‘lock in’ of identity which has been found amongst some social groups (Jones et al., 2011) and considered to perpetuate negative HRBs amongst those in more disadvantaged social circumstances (Jefferis et al., 2004; Pampel et al., 2010).

Moreover, diffusion theory suggests that HRBs in the most advantaged SEP group are imitated by other SEP groups, which may have had positive effects on smoking behaviour over time (Dixon and Banwell, 2009). For example, historically, advantaged SEP women were the first to start and the first to stop smoking (Schooling and Kuh, 2002). At the same time, these diffusion processes could become problematic in the future if we consider alcohol consumption, given evidence that increases in alcohol consumption amongst women in recent decades can be largely explained by increases amongst those who are in more advantaged social circumstances (Purshouse et al., 2017). Thus, whilst public health efforts may serve to reduce alcohol consumption amongst women in more advantaged social circumstances, given their increased receptivity to public health messages (Cutler and Lleras-Muney, 2010; Schooling and Kuh, 2002), they may not adequately address the consumption amongst more disadvantaged women who may increasingly adopt these behaviours but within a different context. For example, women who are considered part of the ‘ladette culture’ are often portrayed as working-class (Jackson and Tinkler, 2007). This ‘ladette culture’ rose in the 1990s and early 2000s amongst young women and is associated with heavier episodic drinking and the consumption of beverages other than wine (Day et al., 2004).

This section has demonstrated how social differentials in HRB clustering may be addressed through upstream interventions that tackle the root cause of multiple negative HRBs for those who are more socially disadvantaged. However, alongside reducing social inequalities in HRB clustering, population strategies that seek to improve the HRBs for every citizen must also be implemented (Rose, 1989).

9.6.2.2 *Population-level efforts to reduce smoking and increase fruit and vegetable consumption and physical activity*

This thesis identified the majority of individuals in both cohorts to be members of the ‘Mainstream’ cluster. The high level of stability in membership of the ‘Mainstream’ cluster in mid-life combined with a shift in membership towards this cluster is to some extent reassuring in terms of health outcomes. For example, the shift to the ‘Mainstream’ cluster can be beneficial for health in respect of cigarette smoking, fruit and vegetable consumption, fried food consumption and physical activity, linked to both mortality and morbidity (see section 1.1).

This shift to the 'Mainstream' cluster characterised by non-smoking, corresponding with declines in the prevalence of smoking over the past 50 years, is a public health success and an example of how both 'upstream' and 'downstream' approaches can improve HRBs (RCP, 2012). Moreover, trends in fruit and vegetable intake may relate to increased availability of fresh fruit and vegetables in the post-war era (Davies, 2016)and growing awareness of the 'five a day' public health message (FSA, 2008), while increases in physical activity may relate to public health communication on the importance of physical activity (Stamatakis et al., 2007).

Despite these apparent improvements, fruit and vegetable consumption, fried food and physical activity HRB patterns in the 'Mainstream' cluster have remained less than ideal, even after 12 years had passed between the two cohorts. For example, in all models the mean frequency of fruit and vegetable consumption was found to be 3–6 days per week, still failing to meet the UK recommended guidelines of five portions of fruit and vegetables per day (WHO, 2014). Less than a third of members in this cluster achieved leisure-time physical activity 4–7 days per week, implying a lack of adherence to the current UK recommended guidelines, to take part in exercise five days per week (Davies et al., 2011).

This shift to the 'Mainstream' cluster therefore highlights success for smoking and improvements for physical activity, fruit and vegetable consumption and fried food consumption. At the same time, it is apparent that continued public health efforts are required to encourage adherence to government guidelines in relation to fruit and vegetable consumption and physical activity. Analyses using data from the Health Survey for England collated in 2012 and the Active People Survey collated in 2014 (HSCIC, 2014a) indicate that the proportion of adults eating fewer than three portions of fruit and vegetables per day (men=45%, women=41%) was higher than those eating the recommended five portions per day (men=25%, women=28%). Moreover, this analysis (HSCIC, 2014a) suggested that physical activity levels remain low, with 57% of adults in England reporting to have not participated in sport[16] in the previous 28 days.

It may be that current fruit and vegetable consumption and physical activity guidelines are unrealistic and may serve to demotivate individuals. A re-framing of fruit and vegetable consumption and physical activity public health messages from being prescriptive to collaborative (i.e. meeting the public where they are currently, rather than where they should be from a public health perspective) would make goals appear more attainable and thus

[16]These include running, golf, swimming, team sports, racket sports, gymnastics, boxing, climbing and mountaineering, winter sports, archery, gym, and fitness activities or classes.

encourage the public to improve these HRBs. This could be achieved by a closer alignment of public health messages to the population average, combined with interventions that remove barriers and enable the public to be more physically active and increase their fruit and vegetable consumption (Kelly et al, 2016).

9.6.2.3 *Population-level efforts to tackle increases in alcohol and sweet food consumption*

Increased membership of the 'Mainstream' cluster over time implies increases in sweet food consumption and moderate alcohol consumption, conceived to reflect possible period effects during the early 2000s (see section 9.5.1). These period effects are considered to relate to the wider availability of alcohol and sweet food consumption over time and changes in social norms. Consequently, 'upstream' (Short and Mollborn, 2015) policies and interventions aimed at a population level are required in order to tackle these undesirable trends in lifestyle.

Notably, a shift towards the 'Mainstream' cluster implies increases in alcohol consumption amongst women and signifies a need for intervention. At the same time, singling out women's alcohol consumption as problematic may not be helpful. For example, portrayal of women's alcohol consumption in the media may result in stereotypes that are not aligned with the recently implemented gender-neutral alcohol consumption guidelines (Patterson et al., 2016). Moreover, research into drinking typologies in the United Kingdom found that women were assimilating to existing drinking styles as opposed to developing new ones (Purshouse et al., 2017), implying that the drinking culture may need to be addressed as opposed to the targeting of women specifically. To tackle access and availability of alcohol it may be that government regulation, rather than self-regulation, of the alcohol industry is required (Knai et al., 2015b) to reduce the availability of alcohol in supermarkets and prevent the aggressive marketing of alcohol products.

A shift towards the 'Mainstream' cluster also implied general increases in the frequency of sweet food consumption, which again may be linked to changes in social norms as well as food availability. Unlike their predecessors, both birth cohorts have experienced a 'post-war consumer culture' (Leach et al., 2013: 107), and increased availability of convenience foods (Davies, 2016). In consequence, these trends may also be the result of changes in social norms for sweet food consumption. Explanations for increasing trends in sweet food include rises in the availability of processed food (POST, 2015; Swinburn et al., 2011). These may be increasingly relied upon, due to competing time demands associated with mid-life (Devine, 2005; Lachman et al., 2015; Worsley et al., 2012), disproportionately affecting women due to their increased participation in the labour market in recent decades whilst continuing to undertake domestic chores (Chou et al., 2004; Nomaguchi and Bianchi, 2004).

Informal eating practices such a snacking have risen dramatically (Harper and Hallsworth, 2016), which has been blamed for increased calorie intake and consequently a rise in obesity (Cutler et al., 2003). These practices combined with evidence that suggests that diets high in added sugars increase the risk of obesity (Neuhouser, 2010) imply that the general increases in sugar consumption found in this thesis could partially explain increases in overweight and obesity rates observed amongst participants in both cohorts (Johnson et al., 2015; Sullivan et al., 2013) but which appears to have occurred at an earlier point in the lifecourse for those born in 1970 (Johnson et al., 2015; Sullivan et al., 2013).

It has been suggested that coinciding declines in smoking prevalence and rises in obesity may be related to one another (Campbell, 2016). Research suggests that trends in smoking and obesity have been substantially influenced by period effects, although cohort does play a role (Badley et al., 2015; Jiang et al., 2013; Reither et al., 2009). Younger cohorts, such as those born in 1970, have had a greater exposure to the obesogenic environment (Johnson et al., 2015; Park et al., 2013) and the social normalisation of obesity (Curtice, 2016), whilst at the same time they have been made aware from an early age of the harms of smoking and have had less exposure to smoking as a social norm (Vedøy, 2014), in comparison to older cohorts, such as those born in 1958.

At the same time, others suggest that such campaigns have in fact led to reductions in obesity by improving diets and physical activity (Dragone et al., 2015). Whilst those who smoke tend to have a lower BMI in comparison to those who are ex-smokers or those do not smoke (Kvaavik et al., 2004; Lahti-Koski et al., 2002; Tian et al., 2016), this has not been explained by worse diets and physical activity levels amongst ex-smokers and non-smokers (Kvaavik et al., 2004; Tian et al., 2016). In consequence, some suggest that the paradox between not smoking, improved diet and physical activity yet having a higher BMI could be due to the under-reporting of calorie intake. This may occur because informal snacking is more difficult for participants to recall in comparison to standard meals (Harper and Hallsworth, 2016). Consumers may not be aware of hidden calories such as added sugars in low-fat food products (Rolls and Miller, 1997), trans-fats contained in processed foods (Allen et al., 2015) or calories in sugary alcoholic (Sayon-Orea et al., 2011) and non-alcoholic beverages (Kvaavik et al., 2005).

Amongst women members of the ‘Mainstream’ cluster, those born in 1970 had significant increases in alcohol consumption and decreases in the frequency of sweet food consumption were observed compared to those born in 1958. Such patterns may reflect a replacement of sugar intake with alcohol use amongst women in the later-born cohort (Colditz et al., 1991). Alcoholic beverages, particularly types popular amongst women (i.e. cocktails), are often calorie-dense (Sayon-Orea et al., 2011). It may be that strategies such as caloric labelling of alcoholic beverages (RSPH, 2014) or extending the soon-to-be-implemented soft drinks industry levy (Briggs et al., 2017) to include alcohol products could serve to mitigate a potential replacement of sugar intake with alcohol use by raising awareness of the sugar content in alcoholic beverages and at the same time reducing availability through price.

On the other hand, it could also be the result of social desirability bias in as much as public health campaigns have made individuals aware of recommended guidelines but this has not necessarily translated into everyday practices (Harper and Hallsworth, 2016; Stamatakis et al., 2007).

In light of the above, it appears that there is a need for ‘upstream’ population-level interventions to tackle the increasing trends in alcohol and sweet food consumption. These interventions should address both availability and social norms in alcohol and sugar consumption to improve HRBs amongst members of the ‘Mainstream’ cluster. In terms of accessibility, alcohol (Brown, 2015; Knai et al., 2015b; Martino et al., 2017; Noel et al., 2016; Petticrew et al., 2017) and food (Bateman-House et al., 2017; Knai et al., 2015a) industries have been criticised for a lack of emphasis on the most effective strategies. These strategies include pricing, access and marketing restrictions (Knai et al., 2015a; Knai et al., 2015b). Therefore, policies that restrict accessibility to health-damaging HRBs at a population level may be effective in improving lifestyles.

Such interventions relating to access and availability are considered to influence social norms but are also generated by them (Blue et al., 2014). Therefore, considering this complex interplay is crucial. Interventions at the individual and population level are unlikely to succeed without coinciding shifts in social norms (Borland, 2013d). As mentioned above (see section 9.6.2.1), interventions that seek to tackle social norms are considered to enact change by creating healthier social environments and the spread of positive HRBs from one individual to another through behavioural modelling (Perry et al., 2016) and diffusion (Dixon and Banwell, 2009). For example, De Leon et al. (2007) theorises that people can quit smoking for either social or personal reasons. The former quit because everyone else is doing so, whereas the latter quit because of a strong personal decision often related to the possibility of negative outcomes. Therefore, interventions that tap into these personal reasons and encourage individual choices to quit, such as public health campaigns highlighting possible negative outcomes of health-risk behaviours (Wakefield et al., 2010), could be accompanied by interventions that target social reasons for change through social marketing techniques (Firestone et al., 2016).

Moreover, social norms may be altered through sophisticated mass media campaigns that resonate with people's experiences of HRBs (Garrett et al., 2015; Reid et al., 2010) and community-level interventions (Harris et al., 2015; MacLellan et al., 2015; Okechukwu et al., 2014; Short and Mollborn, 2015). Interventions based on behavioural economics (i.e. making health-promoting HRBs the default option) known as 'nudge' policies have been shown to be effective (Arno and Thomas, 2016; Loewenstein et al., 2012). However, changing environments to improve HRBs requires increased public acceptance of their choices being restricted by government (Matjasko et al.), which currently may be challenging, given research that suggests a low acceptability of government interventions to improve population HRBs (Diepeveen et al., 2013; Somerville et al., 2015).

9.6.3 A joint approach to improve HRBs

Sections 9.6.1 and 9.6.2 demonstrate how upstream and downstream interventions can be used to tackle multiple negative HRBs. 'Proportionate Universalism' (Marmot, 2010: 16) may offer a solution by combining the upstream and downstream approaches outlined above to reduce inequalities in HRBs. Proportionate Universalism (Marmot, 2010) seeks to tackle social inequalities by implementing population-wide strategies that yield greater benefits for more socially disadvantaged groups, as well as some targeted policies for those with the greatest need. Such policies that focus on the social determinants of lifestyles are considered advantageous because they avoid placing blame, instead acknowledging the unequal distribution of resources that shape HRBs (Benach et al., 2013; Katikireddi et al., 2013; Maller, 2015).

9.7 Future work

This section outlines some fruitful areas of future research which have emerged from this thesis.

9.7.1 Interventions that tackle the social determinants of HRB clustering

The preceding section (see section 9.6.2.1) highlighted the importance of upstream interventions that may serve to improve the contemporaneous social circumstances in which people live their lives, in order to improve their health lifestyles in mid-adulthood.

A future area of research could be the implementation and evaluation of interventions that seek to tackle the effect of social circumstances on HRBs, through physical access, psychosocial stress and social group habitus mechanisms.

An example relating to the psychosocial stress associated with occupation would be a randomised control trial, similar to that conducted by Andersen (2011). As mentioned above (see section 9.6.2.1), they found that workplace policies improving psychosocial work environments combined with interventions to encourage physical activity improved the uptake of exercise amongst workers (Andersen, 2011).

Interventions to improve employees' working conditions – e.g. increased job control (Corbiere et al., 2009) – in order to reduce levels of psychosocial stress, could be provided alongside a lifestyle intervention addressing multiple HRBs (Ashra et al., 2015; Younge et al., 2015). This intervention would consist of three groups: the first group of workers would receive both the lifestyle and working conditions interventions; the second group would receive only the lifestyle intervention (i.e. no change in working conditions); and the third group would act as a control group, receiving neither intervention. Thus, differences in HRBs at follow-up between the first group (receiving the combined intervention) and the other two groups could be attributed to the joint effect of a holistic lifestyle intervention combined with improvements in working conditions and reductions in psychosocial stress.

Another example relates to physical access associated with material circumstances. This would consist of an evaluation of change in HRB clustering at the population level following the implementation of upstream efforts to alter HRBs through price. Evaluations of upstream interventions that address health-damaging HRBs by considering consumer price elasticity (i.e. changes in consumption with changes in price), according to differences in social circumstances, have been conducted (Smed et al., 2007). Batis et al. (2016) evaluated the efficacy of the implementation of a sugar-sweetened beverage tax in Mexico, identifying a positive effect on food purchasing amongst those in low and middle 'socio-economic status' households (measured using a validated scale based upon education, income and household assets). However, as the authors point out, this did not consider improvements in overall diet quality (Batis et al., 2016), which is important, given that price alterations for a single food group may not automatically lead to change in other food groups if their prices do not change (Smed et al., 2007). This demonstrates the limitations of only tackling physical access to negative HRBs and neglecting to improve access to positive ones.

Therefore, instead of focusing on either positive or negative HRBs, upstream interventions may be more effective if they create price barriers to health-damaging HRBs whilst simultaneously improving access to health-promoting HRBs. For example, taxation on alcohol (Vandenberg and Sharma, 2015), tobacco (Blakely et al., 2015) and foods high in sugar and fat (Ni Mhurchu et al., 2015) would be implemented in conjunction with subsidising the price of healthier products such as fruit and vegetables (Cobiac et al., 2017) and leisure-time physical activity participation (Yaniv et al., 2009).

Moreover, measuring overall lifestyle as the outcome, captured through changes in HRB clustering, may be preferable in such an evaluation. This assertion is based upon the findings of this thesis – i.e. that HRBs cluster, as well as other evidence that suggests change in one HRB leads to change in another (Prochaska et al., 2008a).

In sum, evaluations of upstream interventions that seek to tackle psychosocial and physical access mechanisms associated with occupational and material dimensions of social circumstances would be a worthwhile endeavour. Interventions focusing on psychosocial stress, associated with occupation, would combine improvements in employment conditions reducing psychosocial stress and lifestyle modification to tackle multiple HRBs. Interventions tackling price barriers to participating in health-promoting HRBs, associated with material circumstances, would alter access and promote positive lifestyle change.

9.7.2 HRB clustering and the obesity epidemic in the United Kingdom

The relationship between membership of these three distinct HRB clusters and body mass index (BMI) would provide useful insights for policymakers. The rise in obesity is considered a global public health challenge (Friel et al., 2007). As mentioned above (see section 9.6.2.3), younger generations are being exposed to the obesogenic environment for longer durations of their lifespan (Johnson et al., 2015; Park et al., 2013). Consequently, as these younger generations age, the risk of both obesity and associated health outcomes are likely to have severe long-term public health consequences in the United Kingdom (Johnson et al., 2015).

The HRB patterns identified in this thesis suggest that declines in smoking are interrelated with increases in the frequency of sweet food consumption (see chapters 6 and 8). This in turn has possible implications on obesity moving forward, given research linking sugar intake and obesity (Neuhouser, 2010; POST, 2015). Research indicates former smokers to be at greater risk of obesity in comparison to current and never smokers (Dare et al., 2015). Therefore, it may be that some members of the 'Mainstream' cluster have an increased risk of obesity compared to those in the 'Risky' and 'Moderate Smokers' cluster.

At the same time, given that disadvantaged social circumstances are found to predict 'Risky' and 'Moderate Smokers' HRB cluster membership (see chapters 7 and 8), the reverse may also be possible. Research using these two British birth cohorts suggests obesity is socially patterned in the United Kingdom (Bann et al., 2017). Individuals in more disadvantaged social circumstances are more likely to be obese than their more advantaged counterparts. Investigating social differentials across countries would a useful exercise, providing clues as to how differing levels of social inequalities across countries can serve to exacerbate and mitigate inequalities in HRB and obesity (Hoffmann et al., 2015).

The extent to which associations between HRB clustering and obesity are the consequence of age, period and cohort effects would further our understanding of the drivers of the obesity epidemic in the United Kingdom. The results of this thesis elucidated a shift towards the 'Mainstream' cluster which implies a potential period effect for sweet food consumption. This assertion could be verified by using longitudinal panel data (which would have sufficient age variation) with repeated measures of these four HRBs over time and implementing age-period-cohort models (Kerr et al., 2009). This would serve to disentangle the effects of age, period and cohort from one another to monitor current trends and thus better predict the emergence, adaptation, persistence and decline of obesity-related lifestyles in the future (Blue et al., 2014; Maller, 2015). Research conducted in the US suggests that both cohort and period effects have contributed to the obesity epidemic in recent decades (Reither et al., 2009), although more recent analyses indicate that this is largely due to age and period (An and Xiang, 2016). Some studies identify that period has had a greater influence, in comparison to cohort, on obesity (Jiang et al., 2013) and on both smoking and obesity (Badley et al., 2015) whereas others suggest cohort has a larger impact, compared to period, on smoking (Kemm, 2001; Preston and Wang, 2006; Schulze and Mons, 2005; Vedøy, 2014) and obesity (Taylor et al., 2015).

Furthermore, exposure to the obesogenic environment is worthy of attention. An interesting example is a recently conducted study by Meyer et al. (2016) who considered neighbourhood features as instruments, albeit weak, in detecting the relationship between multiple HRBs and BMI.

The inclusion of additional HRBs would also offer insights to understand how lifestyle can drive obesity. Sleep has been found to interrelate with the four HRBs investigated in this thesis (Clark et al., 2015), as has sedentary time (Watts et al., 2015). Both sleep and sedentary time have been linked to an increased risk of obesity (Chastin et al., 2015).

Harnessing new technologies is an innovative approach to detecting and monitoring HRBs (Patrick et al., 2016). Data from different technologies can be linked to better understand HRBs in the context of the obesogenic environment. For example, accelerometers are now widely available and used to measure physical activity (Shiroma et al., 2015). An innovative study conducted in the US linked data from accelerometers and GPS loggers to assess the walkability of an urban neighbourhood (Rundle et al., 2016). Furthermore, Puttelaar et al. (2016) gives examples of how mobile phone apps can record real-time food consumption data, providing a more precise assessment of everyday dietary intake. Despite their limitations and challenges, such technologies have potential for the co-production of a wealth of information by researchers and citizens, through large-scale experiments and the generation of big data, leading to quicker and greater breakthroughs (Patrick et al., 2016).

Using different approaches to come to the same conclusion, known as 'triangulation', is beneficial and improves causal inference (Lawlor et al., 2017). The triangulation of evidence generated from new technologies alongside existing evidence generated from birth cohort studies, cross-country comparisons and other research methodologies, taking into account the biases of each approach, can only serve to strengthen confidence in research findings and address questions of causality.

9.8 Final reflections on taking a person-centred approach to better understand HRBs

This thesis set out to identify the clustering of four HRBs – smoking, alcohol, diet and physical activity – to demonstrate that they are not independent from one another and, in consequence, efforts to effectively modify negative HRBs and improve health outcomes would need to consider the interrelationship between them.

To elucidate this clustering, data-driven person-centred statistical techniques were adopted, detecting HRB cluster patterns shared by subgroups of the British population. Through the extensive and systematic appraisal of the existing research evidence, this data-driven and reductionist approach was chosen over simpler methods, such as bivariate and simple regression analyses, which consider the relationship between pairs of HRBs, to maximise the efficiencies in the data and elucidate distinctive clustered patterns of HRBs.

Moreover, as mentioned in chapter 2 (see section 2.5), based on the appraisal of HRB cluster research, a person-centred rather than variable-centred approach was chosen. This was considered appropriate because, unlike other processes such as ageing or cognitive decline, lifestyle is not an inevitable process that happens to everyone. Instead, lifestyles can be adopted and dropped. Thus, a person-centred approach identifying a categorical variable, reflecting distinct groups who are distinguished from one another by different characteristics, was conceived to better reflect individuals' lifestyles compared with a linear and variable-centred process which implies a severity continuum (Stapinski et al., 2016).

At the same time, it would be naïve to suggest that person-centred approaches should always be chosen over variable-centred techniques, given that other studies using variable-centred techniques to study interrelationships between HRBs have been used to good effect (Benzies et al., 2008; van Nieuwenhuijzen et al., 2009). It is acknowledged that the person-centred models are just one way in which HRBs may be interrelated in the British population and that the triangulation of evidence, generated using different approaches, will improve causal inference (Lawlor et al., 2017).

Moreover, given the caveats associated with purely exploratory data-driven approaches, such as principal components and latent class analysis (i.e. subjectivity in model selection and thus poor replication across different populations), the inclusion of confirmatory approaches that place a greater emphasis on theory and prior knowledge is also important (Finch and Bronk, 2011) in order to elucidate the interrelationship of HRBs (Amato et al., 2016).

In sum, whilst subject to limitations, the data-driven person-centred methods used in this thesis to determine HRB clustering did provide useful insights. This work identified three HRB clusters, labelled 'Risky', 'Moderate Smokers' and 'Mainstream', which were complex, consisting of both positive and negative HRBs, as shown in the frequency of sweet food consumption and moderate alcohol use. The inclusion of social circumstances in pre-adolescence and mid-adulthood as predictors of HRB cluster membership suggested that these HRB clusters were associated with contemporaneous social circumstances which had accumulated from earlier in life. Membership of these three HRB clusters was relatively stable during mid-life, although transitions to more health-promoting clusters were observed which were not influenced by social circumstances in early mid-adulthood. Consequently, this thesis builds upon existing understandings of health lifestyles. This person-centred approach provided an alternative perspective which can contribute towards effective downstream and upstream interventions that seek to modify these four HRBs in a positive way and improve health at the individual and population level.

Appendices

Chapter 5 Appendices

Appendix 5.1

Cohort member questionnaire item wording in the NCDS at age 33 (all variables) and BCS70 at age 30 (diet) and 34 (smoking, physical activity, alcohol consumption) and cohort harmonisation.

Variable	Question NCDS	Response category NCDS	Question BCS70	Response category BCS70	Cohort harmonisation
Smoking (cigarettes smoked per day)	1) Do you smoke cigarettes at all nowadays? 2) How many cigarettes a day do you usually smoke?	Yes No	1) Now some questions about smoking. Would you say that: 2) How many cigarettes a day do you usually smoke?	a) You've never smoked cigarettes? b) You used to smoke cigarettes but don't at all now? c) You now smoke cigarettes occasionally but not every day? d) You smoke cigarettes every day?	Question 1 NCDS No=0 Question 1 BCS70 a/b/c=0 Question 2 NCDS/BCS70 response= cigarettes smoked per day (range 0–80) NOTE 1: BCS70 participants who answered a/b/c for question 1 were not asked question 2. NOTE 2: Sensitivity analysis in the BCS70 found combining 'occasional' smokers with daily smokers, rather than non-smokers, did not influence Latent Profile Analysis model estimates.

Frequency of Fruit and Vegetable Consumption	1) How often do you eat fresh fruit in summer? 2) How often do you eat salads or raw vegetables in winter?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Never	1) How often do you eat fresh fruit? 2) How often do you eat salads or raw vegetables? 3) How often do you eat cooked veg?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Occasionally g) Never	Questions 1 and 2 NCDS f=0, e=1, d=2, c=3, b=4, a=5 Questions 1 and 2 BCS70 g=0, f=1, e=1, d=2, c=3, b=4, a=5 Q1 diet score + Q2 diet score = FV diet score (1–10)
Frequency of Chips and Fried Food Consumption	1) How often do you eat chips? 2) How often do you eat fried food not including chips?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Never	1) How often do you eat food fried in vegetable oil such as olive oil or sunflower oil, not counting chips? 2) How often do you eat food fried in hard fat such as lard or butter, not counting chips? 3) How often do you eat chips?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Occasionally g) Never	Questions 1 and 2 NCDS f=0, e=1, d=2, c=3, b=4, a=5 Questions 1, 2 and 3 BCS70 g=0, f=1, e=1, d=2, c=3, b=4, a=5 Q1 diet score + Q2 diet score = CF diet score (1–10)
Frequency of Sweets, Chocolate, Biscuits/Cake Consumption	1) How often do you eat sweets, chocolates? 2) How often do you eat biscuits?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Never	1) How often do you eat sweets or chocolates? 2) How often do you eat biscuits and cakes of all kinds?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Occasionally g) Never	Questions 1 and 2 NCDS f=0, e=1, d=2, c=3, b=4, a=5 Questions 1 and 2 BCS70 g=0, f=1, e=1, d=2, c=3, b=4, a=5 Q1 diet score + Q2 diet score = SCBC diet score (1–10)

Frequency of Leisure-Time Physical Activity	1) Do you regularly take part in any activities on this card (see below) – that is at least once a month for most of the year? <u>Activities listed on card</u> •Any competitive sports •'Keep fit' or aerobics classes •Circuit training •Weight training or other repeated exercises (at home or in the gym) (listed only at age 33 NCDS) •Running or jogging •Swimming •Cycling •Going for walks •Taking part in water sports •Outdoor sports •Dancing •Any other sport or leisure activity that involves physical exercise 2) How often do you take part in any activity of this type?	Yes No a) Every day b) 4–5 days a week c) 2–3 days a week d) once a week e) 2–3 times a month f) Less often	1) Do you regularly take part in any of the activities on this card (see below), by regularly I mean at least once a month, for most of the year. <u>Activities listed on card</u> •Any competitive sports •'Keep fit' or aerobics classes •Circuit training •Running or jogging •Swimming •Cycling •Going for walks •Taking part in water sports •Outdoor sports •Dancing •Any other sport or leisure activity that involves physical exercise 2) How often do you take part in any activity of this type?	Yes No a) Every day b) 4–5 days a week c) 2–3 days a week d) once a week e) 2–3 times a month f) Less often	Question 1 NCDS/BCS70 No=0 Question 2 NCDS/BCS70 f=0, e=0, d=1, c=2, b=3, a=4 Q1 and Q2 combined = Frequency of leisure-time physical activity with 4 categories. '≤3 times a month' 'Once a week' '2–3 days a week' '4–7 days a week'. NOTE: 6 response categories collapsed into 4 due to sparseness.

Alcohol Consumption (units consumed in the previous week)	1) How often do you have an alcohol drink of any kind? 2) In the last 7 days, that is not counting today by starting from last [name present day of week], how much beer/stout/lager/ale/cider have you had? 3) In the last 7 days how many measures of spirits or liqueurs have you had? 4) In the last 7 days how many glasses of wine have you had? 5) In the last 7 days how many glasses of martini/vermouth/sherry or similar drinks have you had?	a) Most days b) 1,2,3 times a week c) 1,2,3 times a month d) Less often or only on special occasions e) Never	1) How often do you have an alcoholic drink of any kind? Would you say you had a drink … 2) In the last 7 days that is now counting today but starting from last [^Day7Ago], how much beer/stout/lager/ale/cider have you had? 3) In the last 7 days how many measures of spirits or liqueurs have you had, like Gin, Whiskey, Rum, Brandy, Vodka or Advocate? 4) In the last 7 days how many glasses of wine have you had? 5) In the last 7 days how many glasses of martini/vermouth/sherry /port or similar drinks have	a) On most days b) 2–3 days a week c) Once a week d) 2–3 times a month e) Less often or only on special occasions f) Never now a days g) Have you never had an alcohol drink?	Question 1 NCDS e=0, d=0, c=1, b=1, a=1 Question 1 BCS70 g=0, f=0, e=0, d=1, c=1, b=1, a=1 Question 2 NCDS measured in pints. Values converted to units (multiplied by 2). Question 3, 4 and 5 NCDS measured in units. Question 2, 3, 4, 5, 6 and 7 BCS70 measured in units. Alcohol consumption NCDS='no units' if Q1,Q2,Q3,Q4,Q5=0 Alcohol consumption bcs70='no units' if Q1,Q2,Q3,Q4,Q5,Q6,Q7=0 Alcohol consumption NCDS='within limits' if Q1=1 and Q2+Q3+Q4+Q5= 1 to 21 for men, 1 to 14 for women. Alcohol consumption BCS70='within limits' if Q1=1 and Q2+Q3+Q4+Q5+Q6+Q7= ≥22 for men, ≥15 for women. NOTE 1: One outlier in BCS70 (reported 280 units) coded as missing.

			you had? 6) In the last 7 days how many bottles of alcopops have you had? 7) In the last 7 days have you had any other alcohol drinks?		NOTE 2: Sensitivity analysis found that including values from Q6 and Q7 in BCS70 did not inflate alcohol consumption in this cohort (only a small number of participants drank alcopops or other drinks). NOTE 3: Sensitivity analysis found including 'never drinkers' as a separate category did not influence Latent Profile Analysis model estimates.

Appendix 5.2

Cohort member questionnaire item wording in the NCDS at age 33 and age 42 and harmonisation over time.

Variable	**Question NCDS age 33**	**Response category NCDS age 33**	**Question NCDS age 42**	**Response category NCDS age 42**	**Harmonisation**
Smoking (cigarettes smoked per day)	1) Do you smoke cigarettes at all nowadays? 2) How many cigarettes a day do you usually smoke?	Yes No	1) Now some questions about smoking. Would you say that: 2) How many cigarettes a day do you usually smoke?	a) You've never smoked cigarettes? b) You used to smoke cigarettes but don't at all now? c) You now smoke cigarettes occasionally but not every day? d) You smoke cigarettes every day?	Question 1 age 33 No=0 Question 1 age 42 a/b/c=0 Question 2 age 33/42 response= cigarettes smoked per day (range 0–80) NOTE: At age 42 participants who answered a/b/c for question 1 were not asked question 2.

Frequency of Fruit and Vegetable Consumption	1) How often do you eat fresh fruit in summer? 2) How often do you eat salads or raw vegetables in winter?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Never	1) How often do you eat fresh fruit? 2) How often do you eat salads or raw vegetables? 3) How often do you eat cooked veg?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Occasionally g) Never	Questions 1 and 2 age 33 f=0, e=1, d=2, c=3, b=4, a=5 Questions 1 and 2 age 42 g=0, f=1, e=1, d=2, c=3, b=4, a=5 Q1 diet score + Q2 diet score = FV diet score (1–10)
Frequency of Chips and Fried Food Consumption	1) How often do you eat chips? 2) How often do you eat fried food not including chips?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Never	1) How often do you eat food fried in vegetable oil such as olive oil or sunflower oil, not counting chips? 2) How often do you eat food fried in hard fat such as lard or butter, not counting chips? 3) How often do you eat chips?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Occasionally g) Never	Questions 1 and 2 age 33 f=0, e=1, d=2, c=3, b=4, a=5 Questions 1, 2 and 3 age 42 g=0, f=1, e=1, d=2, c=3, b=4, a=5 Q1 diet score + Q2 diet score = CF diet score (1–10)
Frequency of Sweets, Chocolate, Biscuits/Cake Consumption	1) How often do you eat sweets, chocolates? 2) How often do you eat biscuits?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Never	1) How often do you eat sweets or chocolates? 2) How often do you eat biscuits and cakes of all kinds?	a) More than once a day b) Once a day c) 3–6 days a week d) 1 or 2 days a week e) Less than 1 day a week f) Occasionally g) Never	Questions 1 and 2 age 33 f=0, e=1, d=2, c=3, b=4, a=5 Questions 1 and 2 age 42 g=0, f=1, e=1, d=2, c=3, b=4, a=5 Q1 diet score + Q2 diet score = SCBC diet score (1–10)

Frequency of Leisure-Time Physical Activity	1) Do you regularly take part in any activities on this card (see below) – that is at least once a month for most of the year? <u>Activities listed on card</u> •Any competitive sports •'Keep fit' or aerobics classes •Circuit training •Weight training or other repeated exercises (at home or in the gym) (listed only at age 33 NCDS) •Running or jogging •Swimming •Cycling •Going for walks •Taking part in water sports •Outdoor sports •Dancing •Any other sport or leisure activity that involves physical exercise 2) How often do you take part in any activity of this type?	Yes No a) Every day b) 4–5 days a week c) 2–3 days a week d) once a week e) 2–3 times a month f) Less often	1) Do you regularly take part in any of the activities on this card (see below), by regularly I mean at least once a month, for most of the year. <u>Activities listed on card</u> •Any competitive sports •'Keep fit' or aerobics classes •Circuit training •Running or jogging •Swimming •Cycling •Going for walks •Taking part in water sports •Outdoor sports •Dancing •Any other sport or leisure activity that involves physical exercise 2) How often do you take part in any activity of this type?	Yes No a) Every day b) 4–5 days a week c) 2–3 days a week d) once a week e) 2–3 times a month f) Less often	Question 1 age 33/42 No=0 Question 2 age 33/42 f=0, e=0, d=1, c=2, b=3, a=4 Q1 and Q2 combined = Frequency of leisure-time physical activity with 4 categories. '≤3 times a month' 'Once a week' '2–3 days a week' '4–7 days a week'. NOTE: 6 response categories collapsed into 4 due to sparseness.

Alcohol Consumption (units consumed in the previous week)	1) How often do you have an alcohol drink of any kind? 2) In the last 7 days, that is not counting today by starting from last [name present day of week], how much beer/stout/lager/ale/cider have you had? 3) In the last 7 days how many measures of spirits or liqueurs have you had? 4) In the last 7 days how many glasses of wine have you had? 5) In the last 7 days how many glasses of martini/vermouth/sherry or similar drinks have you had?	a) Most days b) 1,2,3 times a week c) 1,2,3 times a month d) Less often or only on special occasions e) Never	1) How often do you have an alcoholic drink of any kind? Would you say you had a drink … 2) In the last 7 days that is now counting today but starting from last [^Day7Ago], how much beer/stout/lager/ale/cider have you had? 3) In the last 7 days how many measures of spirits or liqueurs have you had, like Gin, Whiskey, Rum, Brandy, Vodka or Advocate? 4) In the last 7 days how many glasses of wine have you had? 5) In the last 7 days how many glasses of martini/vermouth/sherry /port or similar drinks	a) On most days b) 2–3 days a week c) Once a week d) 2–3 times a month e) Less often or only on special occasions f) Never now a days g) Have you never had an alcohol drink?	Question 1 age 33 e=0, d=0, c=1, b=1, a=1 Question 1 age 42 g=0, f=0, e=0, d=1, c=1, b=1, a=1 Question 2 age 33 measured in pints. Values converted to units (multiplied by 2). Question 3, 4 and 5 age 33 measured in units. Question 2 age 42 measured in pints. Values converted to units (multiplied by 2). Question 3, 4, 5, 6 and 7 age 42 measured in units. Alcohol consumption age 33='no units' if Q1,Q2,Q3,Q4,Q5=0 Alcohol consumption age 42='no units' if Q1,Q2,Q3,Q4,Q5,Q6,Q7=0 Alcohol consumption age 33='within limits' if Q1=1 and Q2+Q3+Q4+Q5= 1 to 21 for men, 1 to 14 for women. Alcohol consumption age 42='within limits' if Q1=1 and Q2+Q3+Q4+Q5+Q6+Q7= ≥22 for men, ≥15 for women.

			have you had? 6) In the last 7 days how many bottles of alcopops have you had? 7) In the last 7 days have you had any other alcohol drinks?		NOTE 1: Beer/stout/lager/ale/cider consumption at age 42 incorrectly recorded for some participants. Appendix C describes how this was corrected.

Appendix 5.3

Correcting for measurement error in beer consumption at age 42 in the NCDS.

At age 42 in the NCDS participants were asked their beer/stout/lager/ale/cider (shortened hereafter to ‘beer’) consumption in the previous 7 days. This information was supposed to be recorded as the number of pints. However, due to uncertainty in the interviewer instructions it is suspected for some participants their ‘beer’ consumption has been mistakenly recorded in units (2 units = 1 pint). Subsequently, there is potentially an over-estimation of ‘beer’ consumption and the data for ‘beer’ consumption may be unreliable (Elliott et al., 2007).

To circumvent this error, information on ‘beer’ consumption at age 42 was regressed onto the number of units consumed in the previous week at age 33 (incorporating four beverage categories: beer/stout/lager/ale/cider; sprits/liqueurs; wine; martini/vermouth/sherry) and the number of units consumed the previous week at age 42 of all remaining beverages (incorporating five beverage categories: sprits/liqueurs; wine; martini/vermouth/sherry/port; alcopops; other drinks) in order to obtain predicted values of ‘beer’ consumption at age 42. This regression was performed twice, the first time ‘beer’ consumption was assumed to be recorded in pints (Figure 5.3.1), the second time ‘beer’ consumption was assumed to be recorded in units (consumption multiplied by 2) (Figure 5.3.2). This was done in order to compare the outliers and determine if in either case the predicted values were closer to those observed in the data. The plotted predicted values for ‘beer’ consumption at age 42 from these two regression models were found to be similar to each other and similar to observed ‘beer’ consumption at age 42. Therefore, these predicted values provided no clear indication of which observations were coded incorrectly.

Figure 5.3.1 Predicted values of 'beer' consumption at age 42 assumed to be recorded in pints

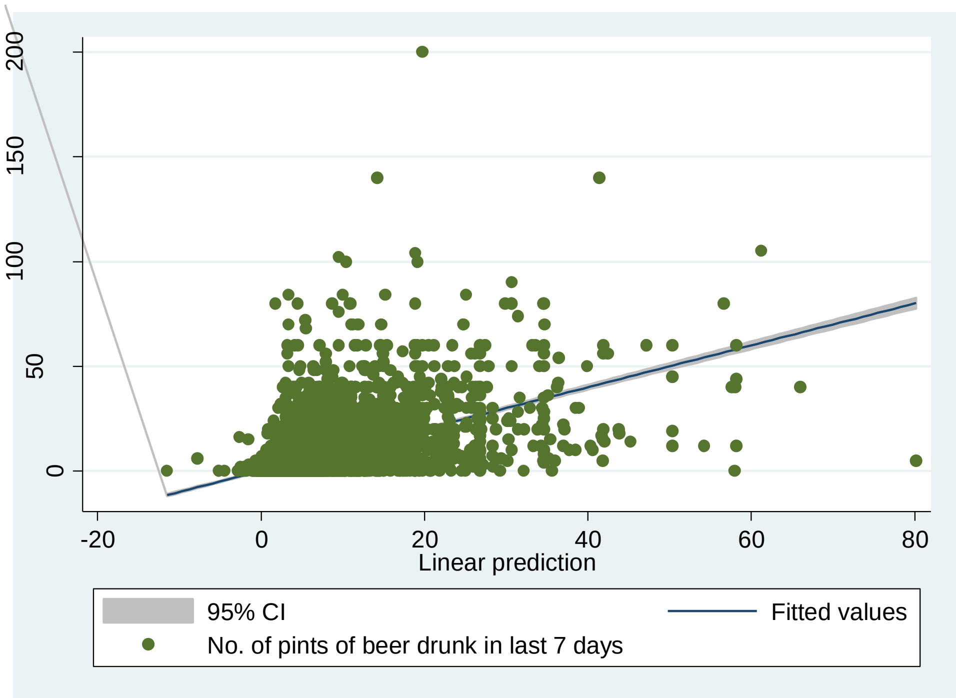
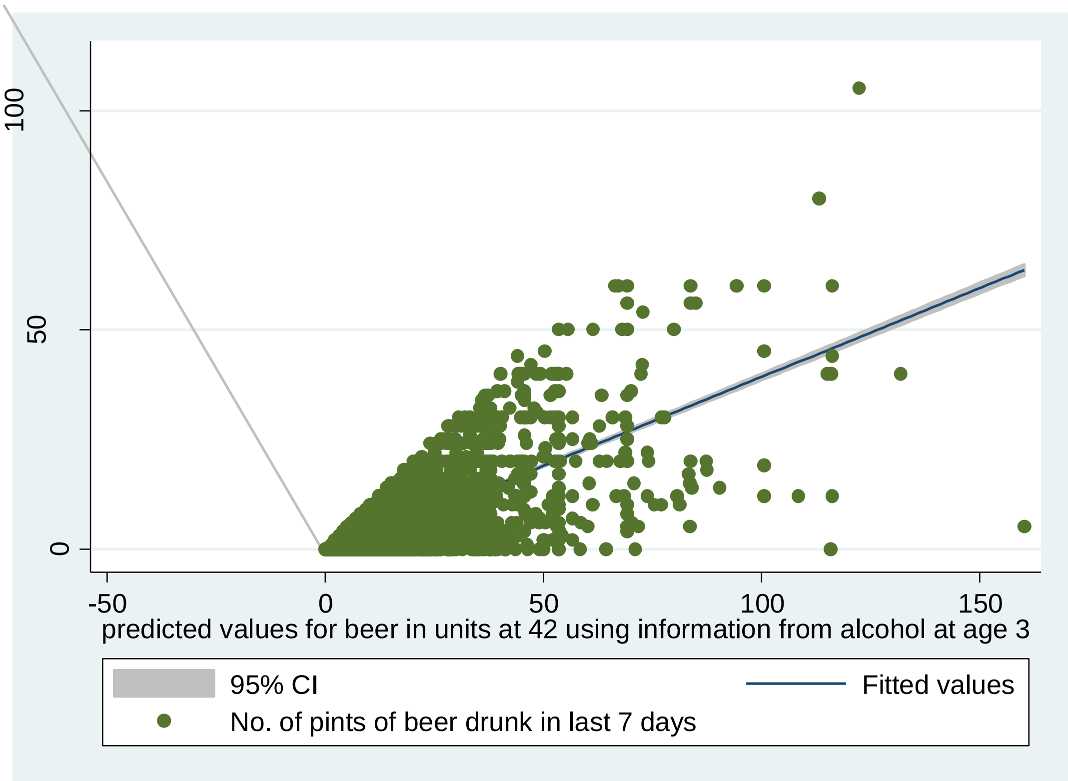


Figure 5.3.2 Predicted values of 'beer' consumption at age 42 assumed to be recorded in units

The next step was to identify participants whose observed 'beer' consumption (assumed to be in pints) at age 42 was greater than or equal to their predicted values of 'beer' consumption in units (n=1,088). These observations were temporarily removed. Following the removal of these observations, there was a notable improvement in the linearity of the distribution between observed 'beer' consumption in pints and predicted 'beer' consumption in units (Figure 5.3.3).

Figure 5.3.3 Predicted values of 'beer' consumption in units at age 42, following the removal of 1,088 participants whose observed beer' consumption in pints ≥ predicted 'beer' consumption in units



In consequence, the observed 'beer' consumption for individuals temporarily removed (n=1,088) were recoded from units to pints (divided by 2), under the assumption that their 'beer' consumption had been measured on the wrong scale.

<u>Correcting for the over-estimation of can drinkers</u>

Age 42 data regarding 'beer' consumption is also suspected to overestimate the 'beer' consumption among those who drink cans instead of pints. This is because cohort members were given incorrect calculations in order to convert cans into pints. The incorrect instructions were 1 large can = 4 * 1/2 pints and 1 small can = 2 * 1/2 pints, whereas the correct conversions were 1 large can = 2 * 1/2 pints and 1 small can = 1 * ½ pint (CLS, 2004).

Unfortunately, because can and pint 'beer' drinkers are not separate from one another in the data it is not possible to identify participants with correct and incorrect data. In response, we categorised alcohol consumption in the previous week at age 42 according to the UK government 'safe' weekly consumption guidelines, that were active at the time of data collection in the year 2000 (DOH, 1995). This categorisation make the results more applicable to public health alcohol policy (Fone et al., 2013) and may, to some extent, reduce the measurement error associated with can 'beer' drinkers in this data sweep.

Distribution of alcohol consumption before and after correction

Table 5.3.1 presents the distribution of total and beer-specific alcohol consumption categorised according to weekly recommended guidelines, before and after the above correction. This demonstrates that the above correction appears to have reduced the proportion drinking above recommended limits. This indicates that the likely over-estimation of total and beer-specific consumption at age 42 has, to some extent, been circumvented by these corrections.

Table 5.3.1: Weekly alcohol consumption at age 42 before and after correction for measurement error

Weekly alcohol consumption at age 42	Before correction n (%)		After correction n (%)	
	Total N=11,357 (100%)	Beer-specific N=9,257 (100%)	Total N=11,357 (100%)	Beer-specific N=9,257 (100%)
No units	2,106 (18.5%)	3,604 (38.9%)	2,106 (18.5%)	3,604 (38.9%)
Within limits (≤14 units women, ≤21 units men)	4,777 (42.1%)	3,790 (40.9%)	6,147 (54.1%)	4,160 (44.9%)
Above limits (≥15 units women, ≥22 units men)	4,474 (39.4%)	1,863 (20.1%)	3,104 (27.3%)	1,493 (16.1%)

Comparing corrected alcohol consumption against external datasets

The distribution of the total alcohol consumption variable at age 42 (i.e. after combining corrected 'beer' consumption with other beverage types at age 42 and categorising the variable according the 'safe' consumption guidelines) was compared with statistics from the General Lifestyles Survey collated by the Office for National Statistics (ONS) for the same period (2000) (ONS, 2006).

The proportion drinking above the weekly recommended guidelines (men > 21 units per week; women >14 units per week) in the NCDS at age 42 were similar when compared with participants aged between 40 and 45 in the General Household Survey (GHS) data (ONS, 2006) collated between 2000 and 2001 and considered to be representative of the population living in England. For men, 32% were drinking above recommended limits in the NCDS data compared to 28% in the GHS data. For women, 16% were drinking above recommended limits in the NCDS data compared to 19% in the GHS data.

The analysis was also undertaken in relation to beer consumption specifically. Overall 16% of participants in the NCDS at age 42 were drinking beer above recommended guidelines in the NCDS compared to 8% in the GHS between ages 40 and 45. Amongst men, the proportion drinking beer above the weekly recommended guidelines in the NCDS was 26% compared to 14% in the GHS data. For women, 5% drank beer above the weekly recommended guidelines in the NCDS compared to 2% in the GHS data.

The proportion drinking above the weekly recommended guidelines in the NCDS at age 42 were also found to be similar to participants aged between 40 and 45 in the Health Survey for England (HSE) data (NatCen, 2011) collated in the year 2000 and considered to be representative of the population living in Great Britain. For men, 32% were drinking above recommended limits in the NCDS data compared to 32% in the HSE data. For women, 16% were drinking above recommended limits in the NCDS data compared to 21% in the HSE data.

For beer consumption, overall 16% of participants in the NCDS drank above recommended guidelines compared to 10% in the HSE. For men, the proportion was 26% in the NCDS compared to 19% in the HSE. Amongst women, 5% drank beer above the recommended guidelines in comparison to 3% in the HSE.

This comparison of NCDS, GHS and HSE data provides some evidence that following these corrections for measurement error among 'beer' drinkers, the resulting variable does not appear to overestimate total weekly alcohol consumption at age 42. Moreover, a comparison of the proportions drinking within and above limits before and after these corrections (see table 5.1) indicates a reduction the over-estimation of total and beer-specific consumption at age 42. However, the comparison of beer-specific consumption with the external datasets does indicate that, whilst these corrections have to some extent been circumvented errors in measurement, beer consumption may remain slightly over-estimated in the NCDS at age 42.

Appendix 5.4

Office for National Statistics procedure for constructing an ordinal three-category version of the NS-SEC.

Eight class NS-SEC (nominal)	Five class NS-SEC (nominal)	Three class NS-SEC (ordinal)
1. Higher managerial, administrative and professional occupations.	1. Higher managerial, administrative and professional occupations.	1. Higher managerial, administrative and professional occupations.
1.1 Large employers and higher managerial and administrative occupations.		
1.2 Higher professional occupations.		
2. Lower managerial, administrative and professional occupations.		
3. Intermediate occupations.	2. Intermediate occupations.	2. Intermediate occupations.
4. Small employers and own account workers.	3. Small employers and own account workers.	
5. Lower supervisory and technical occupations.	4. Lower supervisory and technical occupations.	3. Routine and manual occupations.
6. Semi-routine occupations.	5. Semi-routine and routine occupations.	
7. Routine occupations.		
8. Never worked and long-term unemployed.	* Never worked and long-term unemployed.	* Never worked and long-term unemployed.

Notes: Source ONS (2010a). * Classified as missing.

Chapter 6 Appendices

Appendix 6.1

Estimated means and item response probabilities FIML of 4-cluster multi-group Latent Profile Analysis (LPA) model for men.

	NCDS Men n=5,586 (100%)				**BCS70 Men n=4,613 (100%)**			
	Cluster 1 'Risky' n=71 (1.3%)	**Cluster 2† n=358 (6.4%)**	**Cluster 3 'Moderate Smokers' n=1,346 (24.0%)**	**Cluster 4 'Mainstream' n=3,811 (68.2%)**	**Cluster 1† n=24 (0.5%)**	**Cluster 2 'Risky' n=82 (1.8%)**	**Cluster 3 'Moderate Smokers' n=1,104 (23.9%)**	**Cluster 4 'Mainstream' n=3,404 (73.8%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	43.75 (3.12)	21.44 (0.95)	16.38 (0.77)	0	45.05 (4.03)	16.62 (1.66)	15.21 (0.22)	0
Frequency of fruit and vegetable consumption	2.66 (0.44)	2.91 (0.16)	4.20 (0.26)	4.65 (0.03)	3.14 (0.43)	3.80 (0.30)	4.31 (0.07)	5.10 (0.04)
Frequency of fried food consumption	4.32 (0.27)	5.31 (0.74)	3.68 (0.14)	3.36 (0.02)	4.01 (0.39)	6.62 (0.26)	3.02 (0.04)	2.87 (0.02)
Frequency of sweet food consumption	3.75 (0.46)	3.12 (0.53)	4.44 (0.16)	4.71 (0.03)	3.76 (0.66)	5.37 (0.47)	4.35 (0.08)	4.59 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity								
≤3 times a month	0.58 (0.10)	0.57 (0.07)	0.35 (0.03)	0.28 (0.01)	0.60 (0.12)	0.45 (0.07)	0.41 (0.02)	0.26 (0.01)
Once a week	0.14 (0.07)	0.13 (0.07)	0.23 (0.01)	0.21 (0.01)	0.16 (0.08)	0.10 (0.04)	0.18 (0.01)	0.18 (0.01)
2–3 days a week	0.14 (0.06)	0.12 (0.04)	0.21 (0.02)	0.25 (0.01)	0.04 (0.05)	0.22 (0.06)	0.20 (0.01)	0.29 (0.01)
4–7 days a week	0.13 (0.05)	0.18 (0.04)	0.22 (0.02)	0.26 (0.01)	0.21 (0.09)	0.23 (0.06)	0.22 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week								
No units	0.24 (0.07)	0.15 (0.05)	0.13 (0.01)	0.13 (0.01)	0.03 (0.05)	0.29 (0.07)	0.14 (0.01)	0.12 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.26 (0.07)	0.29 (0.15)	0.56 (0.02)	0.63 (0.01)	0.17 (0.09)	0.36 (0.07)	0.44 (0.02)	0.61 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.49 (0.09)	0.57 (0.12)	0.31 (0.03)	0.24 (0.01)	0.80 (0.11)	0.35 (0.07)	0.42 (0.02)	0.28 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 3-cluster LPA at age 33 in the NCDS or age 34 in the BCS70).

Appendix 6.2

Estimated means and item response probabilities FIML of 4-cluster multi-group Latent Profile Analysis (LPA) model for women.

	NCDS Women Total N=5,787 (100%)				BCS70 Women Total N=5,033 (100%)			
	Cluster 1† n=68 (1.2%)	**Cluster 2 'Risky' n=643 (11.1%)**	**Cluster 3 'Moderate Smokers' n=1,104 (19.1%)**	**Cluster 4 'Mainstream' n=3,973 (68.7%)**	**Cluster 1† n=42 (0.8%)**	**Cluster 2 'Risky' n=190 (3.8%)**	**Cluster 3 'Moderate Smokers' n=955 (19.0%)**	**Cluster 4 'Mainstream' n=3,846 (76.4%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	37.0 (2.39)	18.65 (0.56)	13.51 (0.38)	0	6.76 (2.44)	20.61 (0.28)	12.32 (0.48)	0
Frequency of fruit and vegetable consumption	3.71 (0.50)	3.61 (0.17)	5.72 (0.10)	5.79 (0.03)	5.34 (0.59)	3.64 (0.18)	5.36 (0.21)	5.97 (0.04)
Frequency of fried food consumption	3.71 (0.49)	3.87 (0.15)	2.62 (0.06)	2.55 (0.02)	5.98 (0.31)	3.18 (0.29)	2.32 (0.07)	2.35 (0.02)
Frequency of sweet food consumption	2.78 (0.35)	4.20 (0.25)	4.29 (0.14)	4.85 (0.04)	5.32 (0.90)	3.53 (0.31)	4.48 (0.13)	4.60 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity *≤3 times a month* *Once a week* *2–3 days a week* *4–7 days a week*	0.39 (0.07) 0.26 (0.07) 0.19 (0.07) 0.17 (0.08)	0.63 (0.04) 0.15 (0.02) 0.05 (0.02) 0.17 (0.03)	0.25 (0.03) 0.22 (0.02) 0.22 (0.02) 0.32 (0.02)	0.27 (0.01) 0.24 (0.01) 0.21 (0.01) 0.28 (0.01)	0.34 (0.10) 0.08 (0.08) 0.23 (0.11) 0.36 (0.10)	0.57 (0.07) 0.09 (0.03) 0.06 (0.03) 0.28 (0.05)	0.31 (0.03) 0.14 (0.01) 0.23 (0.02) 0.32 (0.02)	0.25 (0.01) 0.17 (0.01) 0.28 (0.01) 0.31 (0.01)
Alcohol units consumed in the previous week *No units* *Within limits (≤14 units women, ≤21 units men)* *Above limits (≥15 units women, ≥22 units men)*	0.21 (0.07) 0.39 (0.08) 0.40 (0.08)	0.30 (0.03) 0.56 (0.03) 0.14 (0.02)	0.29 (0.02) 0.62 (0.02) 0.09 (0.01)	0.29 (0.01) 0.65 (0.01) 0.07 (0.01)	0.50 (0.11) 0.35 (0.12) 0.14 (0.11)	0.37 (0.06) 0.28 (0.10) 0.36 (0.10)	0.28 (0.02) 0.54 (0.02) 0.18 (0.02)	0.25 (0.01) 0.63 (0.01) 0.13 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 3-cluster LPA at age 33 in the NCDS or age 34 in the BCS70).

Appendix 6.3

Cluster assignment probabilities for most likely HRB cluster membership (row) by HRB cluster membership based on estimated model (column).

NCDS Men	**'Risky'**	**'Moderate Smokers'**	**'Mainstream'**
'Risky'	**0.984**	0.016	0.001
'Moderate Smokers'	0.155	**0.845**	0.000
'Mainstream'	0.000	0.000	**0.998**

BCS70 Men	**'Risky'**	**'Moderate Smokers'**	**'Mainstream'**
'Risky'	**0.845**	0.153	0.002
'Moderate Smokers'	0.021	**0.979**	0.000
'Mainstream'	0.001	0.001	**0.998**

NCDS Women	**'Risky'**	**'Moderate Smokers'**	**'Mainstream'**
'Risky'	**0.878**	0.122	0.005
'Moderate Smokers'	0.221	**0.779**	0.003
'Mainstream'	0.002	0.000	**0.998**

BCS70 Women	**'Risky'**	**'Moderate Smokers'**	**'Mainstream'**
'Risky'	**0.757**	0.241	0.001
'Moderate Smokers'	0.089	**0.911**	0.000
'Mainstream'	0.000	0.001	**0.999**

Descriptive statistics for individual HRBs within each HRB cluster based on most likely HRB cluster membership for men.

	NCDS Men n=5,586 (100%)			BCS70 Men n=4,613 (100%)		
	Cluster 1 'Risky' n=82 (1.5%)	**Cluster 2 'Moderate Smokers' n=1,686 (30.2%)**	**Cluster 3 'Mainstream' n=3818 (68.3%)**	**Cluster 1 'Risky' n=79 (1.7%)**	**Cluster 2 'Moderate Smokers' n=1,124 (24.4%)**	**Cluster 3 'Mainstream' n=3,410 (73.9%)**
	Mean (sd)	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**
Number of cigarettes smoked per day	43.6 (8.9)	17.3 (7.7)	0	20.4 (13.2%)	15.6 (7.3)	0
Frequency of fruit and vegetable consumption	2.5 (2.0)	3.9 (2.1)	4.6 (1.9)	3.7 (2.2)	4.3 (2.1)	5.1 (2.2)
Frequency of fried food consumption	4.6 (1.8)	4.0 (1.6)	3.4 (1.5)	6.7 (1.1)	3.0 (1.0)	2.9 (1.1)
Frequency of sweet food consumption	3.7 (2.6)	4.2 (2.4)	4.7 (2.2)	5.2 (2.9)	4.3 (2.5)	4.6 (2.3)
	n (%)	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Smoking status						
Smokers	82 (100%)	1,683 (99.8%)	72 (1.9%)	79 (100%)	1,123 (99.9%)	21 (0.6%)
Non-smokers	0	3 (0.2%)	3,746 (98.1%)	0	1 (0.1%)	3,389 (99.4%)
Frequency of leisure time physical activity						
≤3 times a month	50 (61.7%)	664 (39.5%)	1,059 (27.9%)	39 (49.4%)	459 (40.8%)	893 (26.2%)
Once a week	7 (8.6%)	350 (20.8%)	809 (21.3%)	6 (7.6%)	199 (17.7%)	620 (18.2%)
2-3 days a week	13 (16.1%)	317 (18.9%)	962 (25.3%)	16 (20.3%)	220 (19.6%)	1,001 (29.4%)
4-7 days a week	11 13.6%)	351 (20.9%)	968 (25.5%)	18 (22.8%)	246 (21.9%)	892 (26.2%)
Alcohol units consumed in the previous week						
No units	21 (25.6%)	229 (13.6%)	504 (13.2%)	20 (25.3%)	151 (13.5%)	398 (11.7%)
Within limits (≤14 units women, ≤21 units men)	19 (23.2%)	846 (50.2%)	2,415 (63.3%)	25 (31.7%)	486 (43.6%)	2,067 (60.8%)
Above limits (≥15 units women, ≥22 units men)	42 (51.2%)	610 (36.2%)	897 (23.5%)	34 (43.0%)	479 (42.9%)	937 (27.5%)

Descriptive statistics for individual HRBs within each HRB cluster based on most likely HRB cluster membership for women.

	NCDS Women Total N=5,787 (100%)			**BCS70 Women Total N=5,033 (100%)**		
	Cluster 1 'Risky' n=515 (8.9%)	**Cluster 2 'Moderate Smokers' n=1,292 (22.3%)**	**Cluster 3 'Mainstream' n=3,980 (68.8%)**	**Cluster 1 'Risky' n=183 (3.6%)**	**Cluster 2 'Moderate Smokers' n=984 (19.6%)**	**Cluster 3 'Mainstream' n=3,866 (76.8%)**
	Mean (sd)	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**	**Mean (sd)**
Number of cigarettes smoked per day	22.2 (8.3)	13.9 (6.8)	0	21.1 (7.2)	12.3 (5.6)	0
Frequency of fruit and vegetable consumption	3.0 (1.7)	5.7 (1.8)	5.8 (1.9)	3.4 (1.9)	5.4 (2.2)	6.0 (2.2)
Frequency of fried food consumption	4.1 (1.6)	2.7 (1.3)	2.6 (1.4)	3.6 (1.2)	2.3 (0.9)	2.4 (1.0)
Frequency of sweet food consumption	3.7 (2.4)	4.4 (2.4)	4.8 (2.3)	3.5 (2.3)	4.5 (2.5)	4.6 (2.2)
	n (%)	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Smoking status						
Smokers	513 (99.6%)	1,292 (100%)	72 (1.8%)	183 (100%)	984 (100%)	23 (0.6%)
Non-smokers	2 (0.4%)	0	3,908 (98.2%)	0	0	3,843 (99.4%)
Frequency of leisure time physical activity						
≤3 times a month	353 (68.8%)	346 (26.9%)	1,076 (27.2%)	107 (58.8%)	304 (30.9%)	950 (24.6%)
Once a week	77 (15.0%)	277 (21.5%)	960 (24.3%)	15 (8.2%)	141 (14.3%)	641 (16.6%)
2-3 days a week	25 (4.9%)	256 (19.9%)	829 (21.0%)	8 (4.4%)	224 (22.8%)	1,061 (27.5%)
4-7 days a week	58 (11.3%)	407 (31.7%)	1,086 (27.5%)	52 (28.6%)	315 (32.0%)	1,207 (31.3%)
Alcohol units consumed in the previous week						
No units	129 (25.1%)	394 (30.5%)	1,147 (28.8%)	81 (44.5%)	266 (27.1%)	952 (24.7%)
Within limits (≤14 units women, ≤21 units men)	277 (53.8%)	791 (61.2%)	2,572 (64.7%)	39 (21.4%)	534 (54.3%)	2,413 (62.5%)
Above limits (≥15 units women, ≥22 units men)	109 (21.2%)	107 (8.3%)	258 (6.5%)	62 (34.1%)	183 (18.6%)	494 (12.8%)

Chapter 7 Appendices

Appendix 7.1

Estimates from NCDS EFA models for pre-adolescent material and cultural dimensions of SEP.

Pre-adolescent Indicator Variables	NCDS Men				NCDS Women			
	Material Sample 1	**Cultural Sample 1**	*Uniqueness*	*Eigenvalue*	**Material Sample 1**	**Cultural Sample 1**	*Uniqueness*	*Eigenvalue*
Housing tenure	0.69	-	*0.52*	*1.92*	*0.54*	-	*0.70*	*1.56*
Overcrowding	0.68	-	*0.53*		*0.64*	-	*0.59*	
Free school meals	0.77	-	*0.40*		*0.67*	-	*0.56*	
Benefits received	0.60	-	*0.63*		*0.64*	-	*0.60*	
Mothers education	-	0.65	*0.58*	*1.16*	-	*0.63*	*0.60*	*1.12*
Fathers education	-	0.69	*0.53*		-	*0.69*	*0.53*	
Parental interest in education	-	0.52	*0.73*		-	*0.50*	*0.75*	

Estimates from NCDS CFA models for pre-adolescent material and cultural dimensions of SEP.

Pre-adolescent Indicator Variables	NCDS Men		NCDS Women	
	Material Sample 2 Estimate(S.E)	**Cultural Sample 2 Estimate(S.E)**	**Material Sample 2 Estimate(S.E)**	**Cultural Sample 2 Estimate(S.E)**
Housing tenure	0.78 (0.03)	-	0.73 (0.03)	-
Overcrowding	0.62 (0.03)	-	0.63 (0.03)	-
Free school meals	0.66 (0.04)	-	0.69 (0.03)	-
Benefits received	0.59 (0.04)	-	0.65 (0.04)	-
Mothers education	-	0.70 (0.03)	-	0.67 (0.03)
Fathers education	-	0.69 (0.03)	-	0.73 (0.03)
Parental interest in education	-	0.61 (0.03)	-	0.61 (0.03)
Pearson r*	0.86		0.88	

Note: Standardised factor loadings. S.E= standard error. All loadings are statistically significant (p<0.001).
*Correlation between dimensions of SEP.

Appendix 7.2

Estimates from NCDS EFA models for mid-adulthood material, occupational and cultural dimensions of SEP.

Mid-adulthood Indicator Variables	**NCDS Men**					**NCDS Women**				
	Material Sample 1	**Occupational Sample 1**	**Cultural Sample 1**	*Uniqueness*	*Eigenvalue*	**Material Sample 1**	**Occupational Sample 1**	**Cultural Sample 1**	*Uniqueness*	*Eigenvalue*
Housing tenure	0.87	-	-	*0.24*	*1.98*	0.83	-	-	*0.32*	*2.04*
Overcrowding	0.61	-	-	*0.63*		0.40	-	-	*0.84*	
Car ownership	0.50	-	-	*0.75*		0.64	-	-	*0.59*	
Benefits received	0.68	-	-	*0.54*		0.73	-	-	*0.46*	
Household income	-0.38	-	-	*0.85*		-0.50	-	-	*0.75*	
NS-SEC	-	-0.62	-	*0.62*	*1.66*	-	-0.61	-	*0.62*	*1.59*
Access to employer pension scheme	-	0.30	-	*0.91*		-	0.64	-	*0.59*	
Chance to buy shares	-	0.40	-	*0.84*		-	0.44	-	*0.81*	
Access to company car	-	0.70	-	*0.51*		-	0.47	-	*0.78*	
Private medical insurance	-	0.73	-	*0.47*		-	0.62	-	*0.61*	
Highest held qualification	-	-	0.78	*0.39*	*1.80*	-	-	0.84	*0.29*	*2.01*
Age left full-time education	-	-	0.73	*0.46*		-	-	0.77	*0.41*	
Highest household Cambridge scale	-	-	0.65	*0.60*		-	-	0.67	*0.59*	

Estimates from NCDS CFA models for mid-adulthood material, occupational and cultural dimensions of SEP.

Mid-adulthood Indicator Variables	**NCDS Men**			**NCDS Women**		
	Material Sample 2 **Estimate (S.E)**	**Cultural Sample 2** **Estimate (S.E)**	**Occupational Sample 2** **Estimate (S.E)**	**Material Sample 2** **Estimate (S.E)**	**Cultural Sample 2** **Estimate (S.E)**	**Occupational Sample 2** **Estimate (S.E)**
Housing tenure	0.90 (0.02)	-	-	0.87 (0.02)	-	-
Overcrowding	0.60 (0.03)	-	-	0.53 (0.03)	-	-
Car ownership	0.36 (0.03)	-	-	0.56 (0.03)	-	-
Benefits received	0.59 (0.04)	-	-	0.74 (0.02)	-	-
Household income	-0.58 (0.02)	-	-	-0.64 (0.02)	-	-
NS-SEC	-	-	0.96 (0.02)	-	-	0.91 (0.02)
Access to employer pension scheme	-	-	-0.26 (0.03)	-	-	-0.53 (0.03)
Chance to buy shares	-	-	-0.25 (0.03)	-	-	-0.04 (0.04)†
Access to company car	-	-	-0.48 (0.03)	-	-	-0.43 (0.04)
Private medical insurance	-	-	-0.57 (0.03)	-	-	-0.32 (0.04)
Highest held qualification	-	-0.90 (0.01)	-	-	-0.95 (0.01)	-
Age left full-time education	-	-0.75 (0.01)	-	-	-0.78 (0.01)	-
Highest household Cambridge scale	-	-0.63 (0.01)	-	-	-0.63 (0.01)	-
Pearson r* OCC MAT CULT	OCC MAT CULT - 0.87 - 0.68 0.74 -			OCC MAT CULT - 0.70 - 0.89 0.68 -		

Note: Standardised factor loadings. S.E= standard error. All loadings statistically significant (p<0.001), except †= p>0.05.
*Correlation between dimensions of SEP. OCC=occupational dimension of SEP, MAT=material dimension of SEP, CULT=cultural dimension of SEP.

Appendix 7.3

Estimates from BCS70 EFA models for pre-adolescent material and cultural dimensions of SEP.

Pre-adolescent Indicator Variables	NCDS Men				NCDS Women			
	Material Sample 1	**Cultural Sample 1**	*Uniqueness*	*Eigenvalue*	**Material Sample 1**	**Cultural Sample 1**	*Uniqueness*	*Eigenvalue*
Housing tenure	0.74	-	*0.46*	*2.57*	0.71	-	*0.50*	*2.65*
Overcrowding	0.41	-	*0.83*		0.36	-	*0.87*	
Free school meals	0.88	-	*0.22*		0.87	-	*0.30*	
Benefits received	0.78	-	*0.39*		0.84	-	*0.30*	
Household income	-0.68	-	*0.53*		-0.74	-	*0.45*	
Mothers education	-	0.74	*0.45*	*1.23*	-	*0.66*	*0.57*	*1.14*
Fathers education	-	0.72	*0.48*		-	*0.70*	*0.52*	
Parental interest in education	-	0.40	*0.84*		-	*0.47*	*0.78*	

Estimates from BCS70 CFA models for pre-adolescent material and cultural dimensions of SEP.

Pre-adolescent Indicator Variables	BCS70 Men		BCS70 Women	
	Material Sample 2 Estimate(S.E)	**Cultural Sample 2 Estimate(S.E)**	**Material Sample 2 Estimate(S.E)**	**Cultural Sample 2 Estimate(S.E)**
Housing tenure	0.76 (0.02)	-	0.71 (0.02)	-
Overcrowding	0.47 (0.03)	-	0.50 (0.03)	-
Free school meals	0.87 (0.02)	-	0.85 (0.02)	-
Benefits received	0.85 (0.02)	-	0.88 (0.02)	-
Household income	-0.70 (0.02)	-	-0.76 (0.02)	-
Mothers education	-	0.73 (0.03)	-	0.73 (0.03)
Fathers education	-	0.85 (0.03)	-	0.80 (0.03)
Parental interest in education	-	0.48 (0.03)	-	0.49 (0.03)
Pearson r*	0.70		0.73	

Note: Standardised factor loadings. S.E= standard error. All loadings are statistically significant ($p<0.001$). *Correlation between dimensions of SEP.

Appendix 7.4

Estimates from BCS70 EFA models for mid-adulthood material, occupational and cultural dimensions of SEP.

Mid-adulthood Indicator Variables	**NCDS Men**					**NCDS Women**				
	Material Sample 1	**Occupational Sample 1**	**Cultural Sample 1**	*Uniqueness*	*Eigenvalue*	**Material Sample 1**	**Occupational Sample 1**	**Cultural Sample 1**	*Uniqueness*	*Eigenvalue*
Housing tenure	0.80	-	-	*0.36*	*1.76*	0.82	-	-	*0.32*	*1.88*
Overcrowding	0.56	-	-	*0.68*		0.54	-	-	*0.70*	
Car ownership	0.33	-	-	*0.89*		0.38	-	-	*0.86*	
Benefits received	0.66	-	-	*0.56*		0.72	-	-	*0.48*	
Household income	-0.50	-	-	*0.74*		-0.50	-	-	*0.75*	
NS-SEC	-	0.48	-	*0.77*	*1.24*	-	*0.52*	-	*0.73*	*1.35*
Access to employer pension scheme	-	0.73	-	*0.47*		-	*0.74*	-	*0.45*	
Has employer pension scheme	-	0.70	-	*0.52*		-	*0.73*	-	*0.47*	
Highest held qualification	-	-	0.78	*0.39*	*1.45*	-	-	0.77	*0.41*	*1.37*
Age left full-time education	-	-	0.69	*0.53*		-	-	0.64	*0.58*	
Highest household Cambridge scale	-	-	0.61	*0.63*		-	-	0.60	*0.64*	

Estimates from BCS70 CFA models for mid-adulthood material, occupational and cultural dimensions of SEP.

Mid-adulthood Indicator Variables	BCS70 Men			BCS70 Women		
	Material Sample 2 Estimate (S.E)	Cultural Sample 2 Estimate (S.E)	Occupational Sample 2 Estimate (S.E)	Material Sample 2 Estimate (S.E)	Cultural Sample 2 Estimate (S.E)	Occupational Sample 2 Estimate (S.E)
Housing tenure	0.79 (0.02)	-	-	0.82 (0.02)	-	-
Overcrowding	0.63 (0.03)	-	-	0.53 (0.03)	-	-
Car ownership	0.16 (0.03)	-	-	0.33 (0.03)	-	-
Benefits received	0.53 (0.04)	-	-	0.78 (0.02)	-	-
Household income	-0.62 (0.02)	-	-	-0.59 (0.02)	-	-
NS-SEC	-	-	0.91 (0.02)	-	-	0.93 (0.02)
Access to employer pension scheme	-	-	0.39 (0.03)	-	-	0.45 (0.03)
Has employer pension scheme	-	-	0.46 (0.03)	-	-	0.58 (0.03)
Highest held qualification	-	-0.93 (0.01)	-	-	-0.95 (0.01)	-
Age left full-time education	-	-0.66 (0.01)	-	-	-0.62 (0.01)	-
Highest household Cambridge scale	-	-0.58 (0.01)	-	-	-0.58 (0.01)	-
Pearson r* OCC MAT CULT	OCC MAT CULT - 0.67 - 0.86 0.63 -			OCC MAT CULT - 0.78 - 0.80 0.62 -		

Note: Standardised factor loadings. Est= estimate, S.E= standard error. All loadings are statistically significant (p<0.001).
*Correlation between dimensions of SEP. OCC=occupational dimension of SEP, MAT=material dimension of SEP, CULT=cultural dimension of SEP.

Appendix 7.5

Distributions for the dimensions of SEP constructs in pre-adolescence and mid-adulthood in the NCDS and BCS70.

SEP construct*	NCDS Men		NCDS Women		BCS70 Men		BCS70 Women	
	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
Pre-adolescent material	0	0.52	0	0.49	0	0.58	0	0.48
Pre-adolescent cultural	0	0.41	0	0.45	0	0.59	0	0.51
Mid-adulthood material	0	0.11	0	0.21	0	0.13	0	0.15
Mid-adulthood occupational	0	0.52	0	0.72	0	0.42	0	0.40
Mid-adulthood cultural	0	0.83	0	0.92	0	0.65	0	0.64

Note: *Product of CFA analysis using WLSMV estimator in Mplus Version 7.

Appendix 7.6

Tested paths between pre-adolescent dimensions of SEP and mid-adulthood dimensions of SEP in the NCDS and BCS70.

Dimensions of SEP	NCDS Men	NCDS Women	BCS70 Men	BCS70 Women
Pre-adolescent material and mid-adulthood material	0.54 (0.51, 0.58)	0.59 (0.56, 0.63)	0.51 (0.47, 0.55)	0.52 (0.49, 0.55)
Pre-adolescent material and mid-adulthood cultural	0.58 (0.55, 0.60)	0.56 (0.53, 0.58)	0.46 (0.43, 0.48)	0.44 (0.41, 0.46)
Pre-adolescent material and mid-adulthood occupational	0.50 (0.45, 0.54)	0.44 (0.40, 0.49)	0.48 (0.43, 0.53)	0.40 (0.36, 0.45)
Pre-adolescent cultural and mid-adulthood material	0.58 (0.54, 0.62)	0.59 (0.55, 0.62)	0.51 (0.47, 0.56)	0.50 (0.46, 0.50)
Pre-adolescent cultural and mid-adulthood cultural	0.76 (0.74, 0.79)	0.65 (0.63, 0.67)	0.59 (0.57, 0.62)	0.57 (0.54, 0.60)
Pre-adolescent cultural and mid-adulthood occupational	0.64 (0.60, 0.68)	0.56 (0.52, 0.61)	0.59 (0.53, 0.64)	0.54 (0.49, 0.59)

Note: One unit increase in pre-adolescent and mid-adulthood dimensions of SEP = more disadvantaged. Effect estimates are standardised linear regression coefficients and 95% confidence intervals, all p values ≤ 0.001.

Appendix 7.7

Missing data patterns for pre-adolescent and mid-adulthood SEP indicator variables in the NCDS.

NCDS Pre-adolescent Indicator Variables	**Missing in Risky cluster n(%)**	**Missing in Moderate Smokers cluster n(%)**	**Missing in Mainstream cluster n(%)**	**p value***
Housing tenure	95 (15.9%)	475 (16.0%)	1,154 (14.8%)	0.29
Overcrowding	95 (15.9%)	477 (16.0%)	1,153 (14.7%)	0.25
Free school meals	100 (16.8%)	508 (17.1%)	1,231 (15.8%)	0.26
Benefits received	185 (30.9%)	939 (31.5%)	2,279 (29.2%)	0.06
Mothers education	58 (9.7%)	322 (10.8%)	885 (11.4%)	0.39
Fathers education	118 (19.8%)	602 (20.2%)	1,503 (19.3%)	0.54
Parental interest in education	162 (27.1%)	714 (24.0%)	1,667 (21.4%)	<0.001
NCDS Mid-adulthood Indicator Variables	**Missing in Risky cluster n(%)**	**Missing in Moderate Smokers cluster n(%)**	**Missing in Mainstream cluster n(%)**	**p value***
Housing tenure	60 (10.1%)	279 (9.4%)	597 (7.7%)	0.004
Overcrowding	11 (1.8%)	55 (1.9%)	164 (2.1%)	0.67
Car ownership	8 (1.3%)	41 (1.4%)	62 (0.8%)	0.02
Benefits received	4 (0.7%)	13 (0.4%)	31 (0.4%)	0.61
NS-SEC	69 (11.6%)	160 (5.4%)	436 (5.6%)	<0.001
Access to employer pension scheme	278 (46.6%)	1,138 (38.2%)	2,555 (32.8%)	<0.001
Chance to buy shares	278 (46.6%)	1,138 (38.2%)	2,555 (32.8%)	<0.001
Access to company car	278 (46.6%)	1,138 (38.2%)	2,555 (32.8%)	<0.001
Private medical insurance	278 (46.6%)	1,138 (38.2%)	2,555 (32.8%)	<0.001
Highest held qualification	15 (2.5%)	85 (2.9%)	137 (1.8%)	0.001
Age left full-time education	125 (20.9%)	522 (17.5%)	1,138 (14.6%)	<0.001
Highest household Cambridge scale	68 (11.4%)	157 (5.3%)	422 (5.4%)	<0.001
Household income	112 (18.8%)	476 (16.0%)	1,095 (14.0%)	0.001

Note: *p value refers to chi-square test of differences in proportion of missing values for each predictor across the three outcome categories.

Appendix 7.8

Missing data patterns for pre-adolescent and mid-adulthood SEP indicator variables in the BCS70.

BCS70 Pre-adolescent Indicator Variables	**Missing in Risky cluster n(%)**	**Missing in Moderate Smokers cluster n(%)**	**Missing in Mainstream cluster n(%)**	**p value***
Housing tenure	43 (16.4%)	291 (13.8%)	914 (12.6%)	0.08
Overcrowding	44 (16.8%)	306 (14.5%)	930 (12.8%)	0.03
Free school meals	46 (17.6%)	287 (13.6%)	898 (12.3%)	0.02
Benefits received	76 (29.01%)	599 (28.4%)	2,100 (28.9%)	0.92
Mothers education	67 (25.6%)	491 (23.3%)	1,539 (21.2%)	0.04
Fathers education	82 (31.3%)	599 (28.4%)	1,741 (23.9%)	<0.001
Parental interest in education	93 (35.5%)	645 (30.6%)	2,002 (27.5%)	0.001
Household income	55 (21.0%)	423 (20.1%)	1,408 (19.4%)	0.64
BCS70 Mid-adulthood Indicator Variables	**Missing in Risky cluster n(%)**	**Missing in Moderate Smokers cluster n(%)**	**Missing in Mainstream cluster n(%)**	**p value***
Housing tenure	4 (1.5%)	8 (0.4%)	34 (0.5%)	0.04
Overcrowding	4 (1.5%)	10 (0.5%)	48 (0.7%)	0.12
Car ownership	8 (3.1%)	183 (8.7%)	481 (6.6%)	<0.001
Benefits received	1 (0.4%)	2 (0.1%)	11 (0.2%)	0.50
NS-SEC	57 (21.8%)	196 (9.3%)	485 (6.7%)	<0.001
Access to employer pension scheme	111 (42.4%)	661 (31.4%)	1,893 (26.02%)	<0.001
Has joined employer pension scheme	111 (42.4%)	652 (30.9%)	1,868 (25.7%)	<0.001
Highest held qualification	2 (0.8%)	1 (0.05%)	19 (0.3%)	0.04
Age left full-time education	1 (0.4%)	4 (0.2%)	17 (0.2%)	0.81
Highest household Cambridge scale	57 (21.8%)	200 (9.5%)	492 (6.8%)	<0.001
Household income	85 (32.4%)	432 (20.5%)	976 (13.4%)	<0.001

Note: *p value refers to chi-square test of differences in proportion of missing values for each predictor across the three outcome categories.

Appendix 7.9

Covariates included in path models to predict missing data in the NCDS and BCS70.

NCDS	**BCS70**
Birthweight (age 0)	Single mother at birth (age 0)
Reading score (age 16)	Parity (age 0)
Rutter score (age 16)	Breastfeeding not attempted (age 0)
Number of jobs (age 23)	Mothers age at delivery (age 0)
Ever unemployed (age 23)	
Employed (age 23)	
Lives with parents (age 23)	
Number of rooms in household (age 23)	

Note: NCDS sources are (Atherton et al., 2008), Hawkes and Plewis (2006). BCS70 source is Mostafa and Wiggins (2014).

Appendix 7.10

Estimates from NCDS EFA models for pre-adolescent and mid-adulthood uni-dimensional SEP constructs.

NCDS Pre-adolescent Indicator Variables	**Men (sample 1)**	**Women (sample 1)**
Housing tenure	0.73	0.70
Overcrowding	0.62	0.65
Free school meals	0.71	0.57
Benefits received	0.64	0.46
Mothers education	0.58	0.62
Fathers education	0.70	0.74
Parental interest in education	0.58	0.59
1 Factor Eigenvalue	*3.43*	*3.24*
NCDS Mid-adulthood Indicator Variables	**Men (sample 1)**	**Women (sample 1)**
Housing tenure	-0.68	-0.73
Overcrowding	-0.44	-0.45
Car ownership	-0.26	-0.46
Benefits received	-0.51	-0.62
Household income	0.47	0.51
NS-SEC	-0.88	-0.82
Access to employer pension scheme	0.29	0.51
Chance to buy shares	0.21	0.11
Access to company car	0.45	0.26
Private medical insurance	0.48	0.27
Highest held qualification	0.74	0.83
Age left full-time education	0.59	0.66
Highest household Cambridge scale	0.76	0.76
Eigenvalue	*5.20*	*4.95*

Note: Total sample split and EFA models ran in one half. All factor loadings statistically significant ($p<0.05$).

Appendix 7.11

Estimates from BCS70 EFA models for pre-adolescent and mid-adulthood uni-dimensional SEP constructs.

BCS70 Pre-adolescent Indicator Variables	**Men (sample 1)**	**Women (sample 1)**
Housing tenure	0.77	0.72
Overcrowding	0.46	0.42
Free school meals	0.81	0.85
Benefits received	0.79	0.84
Household income	-0.66	-0.72
Mothers education	0.64	0.55
Fathers education	0.70	0.67
Parental interest in education	0.42	0.45
Eigenvalue	*3.71*	*3.77*
BCS70 Mid-adulthood Indicator Variables	**Men (sample 1)**	**Women (sample 1)**
Housing tenure	0.66	0.74
Overcrowding	0.47	0.48
Car ownership	0.08	0.28
Benefits received	0.50	0.73
Household income	-0.48	-0.51
NS-SEC	0.80	0.81
Access to employer pension scheme	0.43	0.48
Has an employer pension scheme	0.49	0.56
Highest held qualification	-0.79	-0.75
Age left full-time education	-0.58	-0.48
Highest household Cambridge scale	-0.81	-0.75
Eigenvalue	*4.36*	*4.61*

Note: Total sample split and EFA models ran in one half. All factor loadings statistically significant ($p<0.05$).

Appendix 7.12

Distributions for SEP in pre-adolescence and mid-adulthood in the NCDS.

Table 7.12.1: NCDS pre-adolescent and mid-adulthood SEP construct descriptive statistics

SEP construct*	NCDS Men		NCDS Women	
	Mean	Variance	Mean	Variance
Pre-adolescent	0	0.50	0	0.46
Mid-adulthood	0	0.18	0	0.24

*Product of CFA analysis using WLSMV estimator in Mplus Version 7.

Table 7.12.2: NCDS pre-adolescent and mid-adulthood SEP estimated construct score descriptive statistics

SEP estimated construct score*	NCDS Men			NCDS Women		
	Mean	Variance	Range	Mean	Variance	Range
Pre-adolescent	0.009	0.35	-1.57, 1.76	0.010	0.31	-2.01, 1.87
Mid-adulthood	0.005	0.10	-0.91, 0.95	0.009	0.19	-1.79, 1.26

*Estimated using regression methods in Mplus Version 7, descriptive analyses undertaken in Stata Version 14.

Figure 7.12.1: NCDS Men estimated pre-adolescent SEP estimated construct score histogram

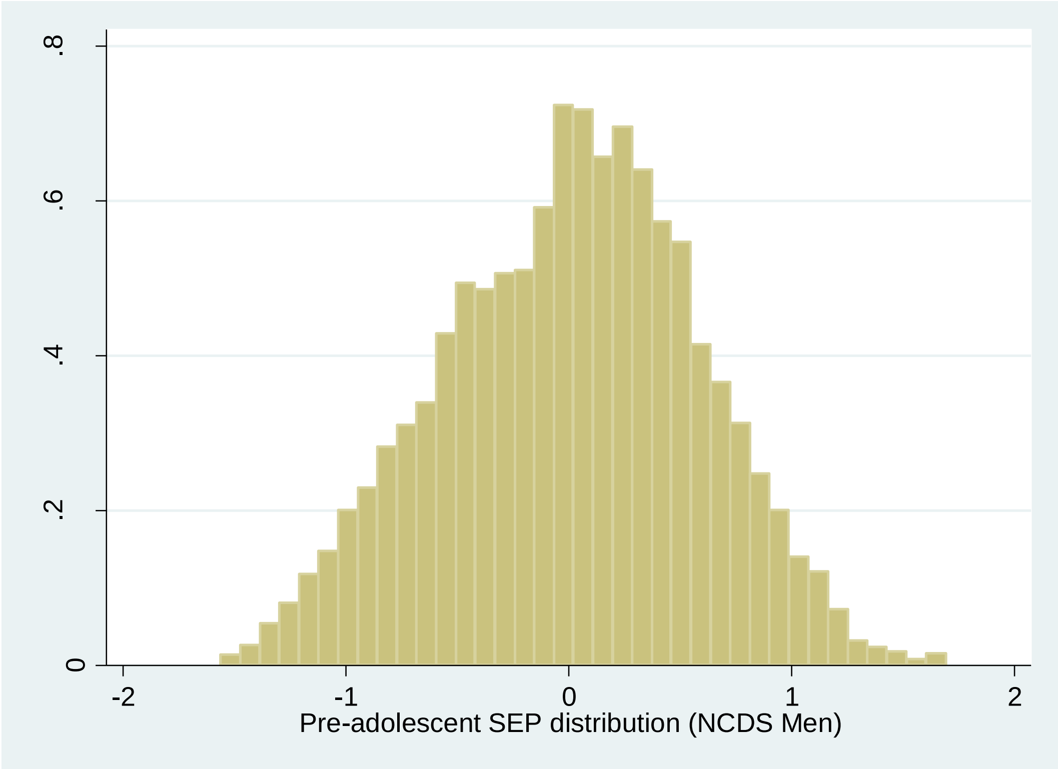


Figure 7.12.2: NCDS Men estimated mid-adulthood SEP estimated construct score histogram

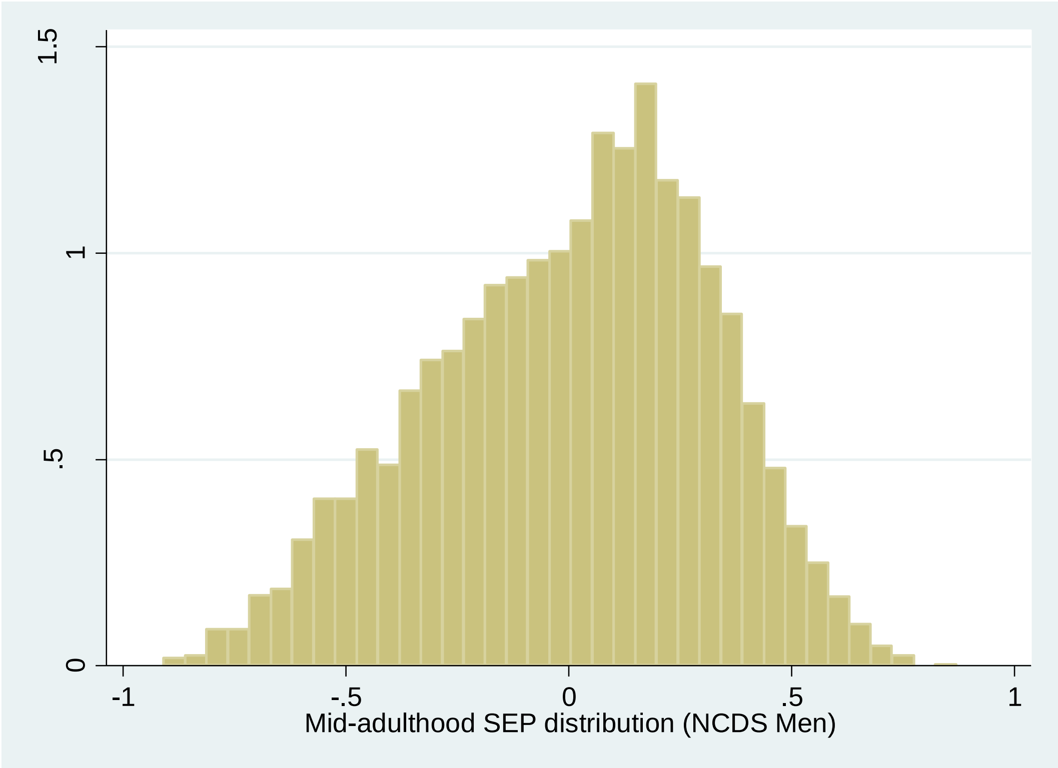


Figure 7.12.3: NCDS Women estimated pre-adolescent SEP estimated construct score histogram

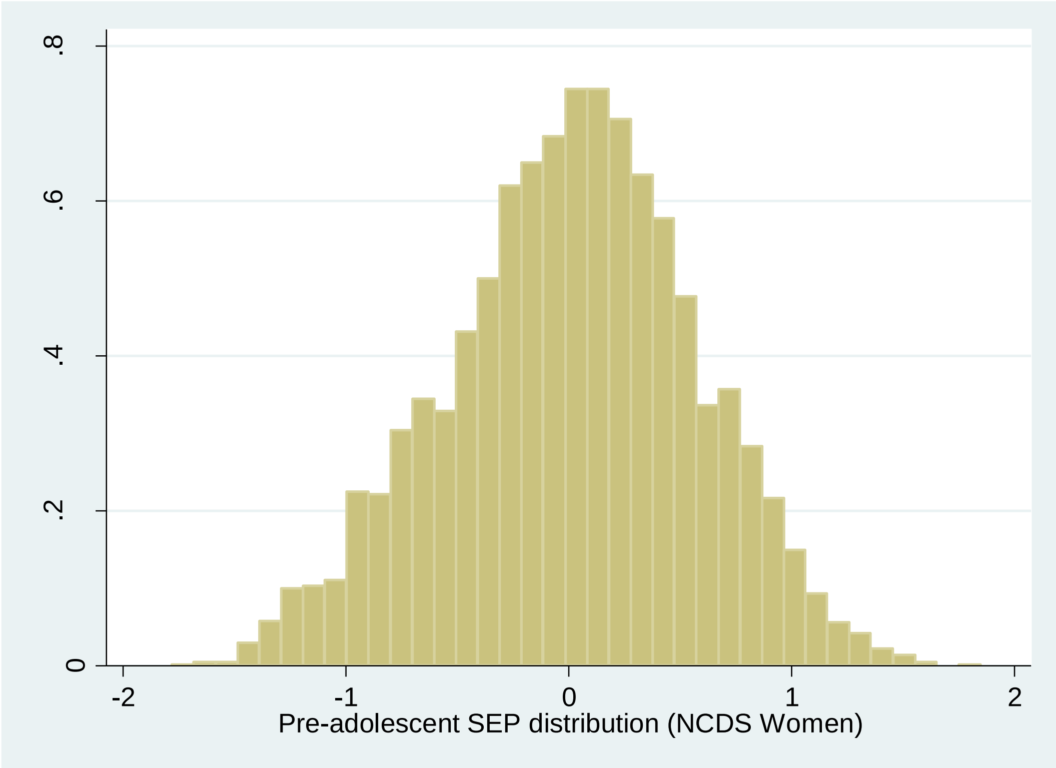
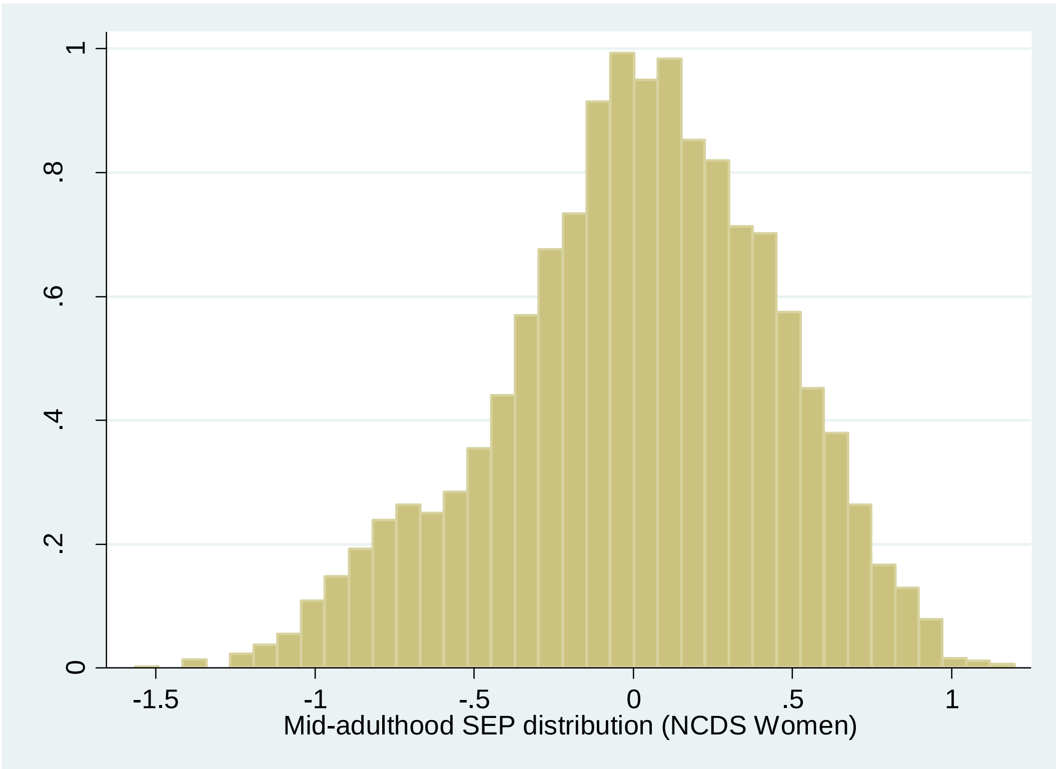


Figure 7.12.4: NCDS Women estimated mid-adulthood SEP estimated construct score histogram



Appendix 7.13

Distributions for SEP in pre-adolescence and mid-adulthood in the BCS70.

Table 7.13.1: BCS70 pre-adolescent and mid-adulthood SEP construct descriptive statistics

SEP construct*	BCS70 Men		BCS70 Women	
	Mean	Variance	Mean	Variance
Pre-adolescent	0	0.63	0	0.77
Mid-adulthood	0	0.19	0	0.24

*Product of CFA analysis using WLSMV estimator in Mplus Version 7.

Table 7.13.2: BCS70 pre-adolescent and mid-adulthood SEP estimated construct score descriptive statistics

SEP estimated construct score*	BCS70 Men			BCS70 Women		
	Mean	Variance	Range	Mean	Variance	Range
Pre-adolescent	0.007	0.44	-1.85, 2.35	0.02	0.55	-2.13, 2.41
Mid-adulthood	0.005	0.11	-0.95, 0.94	0.02	0.15	-1.14, 1.15

*Estimated using regression methods in Mplus Version 7, descriptive analyses undertaken in Stata Version 14.

Figure 7.13.1: BCS70 Men estimated pre-adolescent SEP estimated construct score histogram

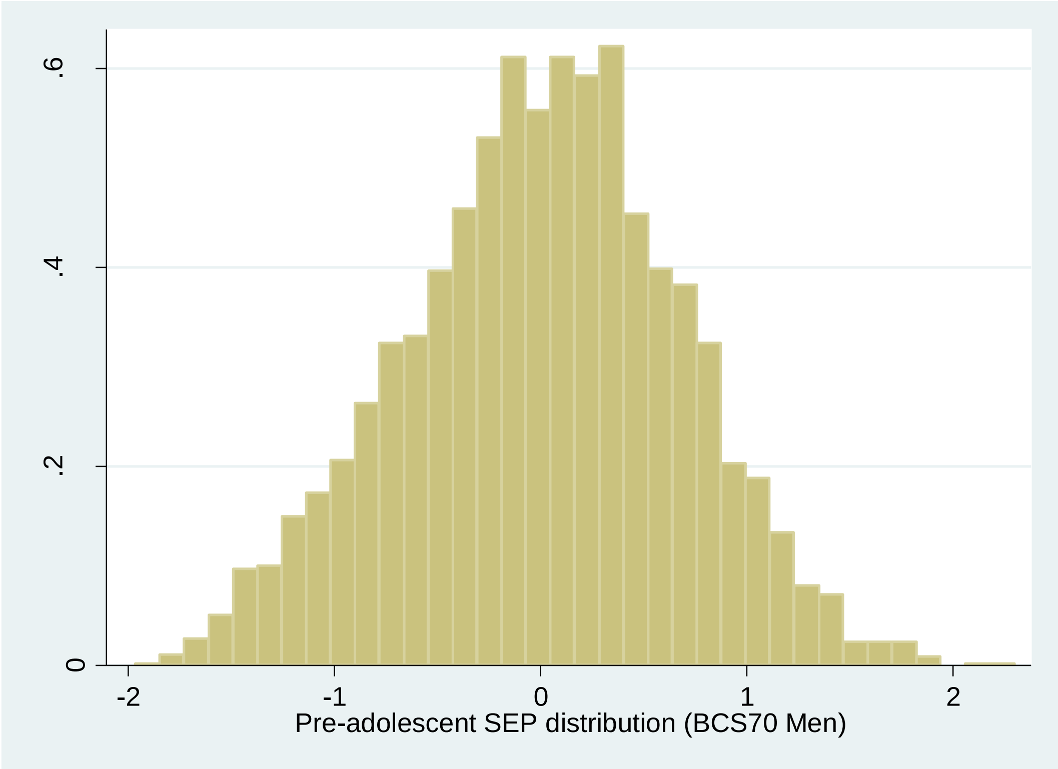


Figure 7.13.2: BCS70 Men estimated mid-adulthood SEP estimated construct score histogram

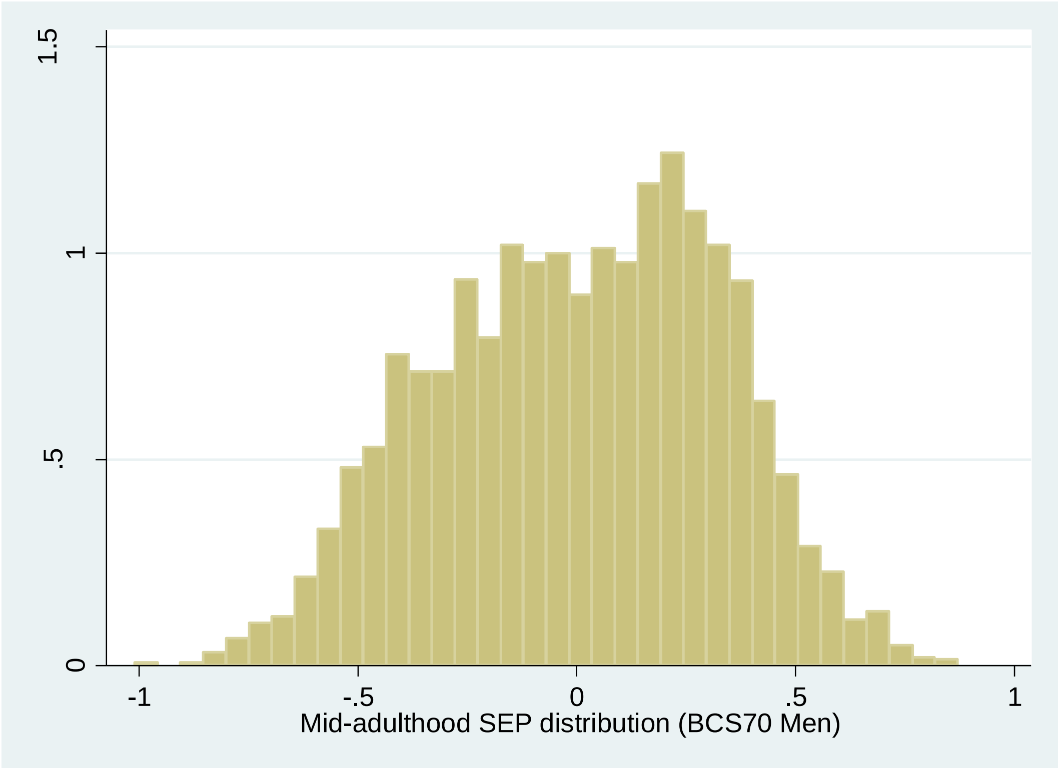


Figure 7.13.3: BCS70 Women estimated pre-adolescent SEP estimated construct score histogram

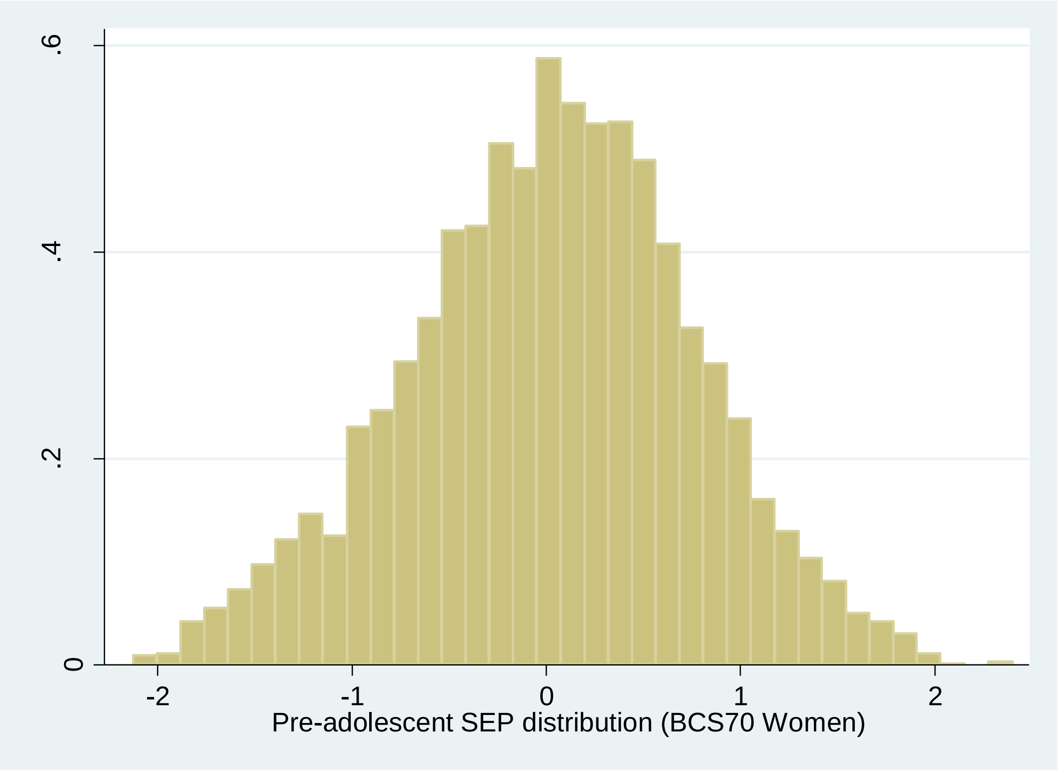
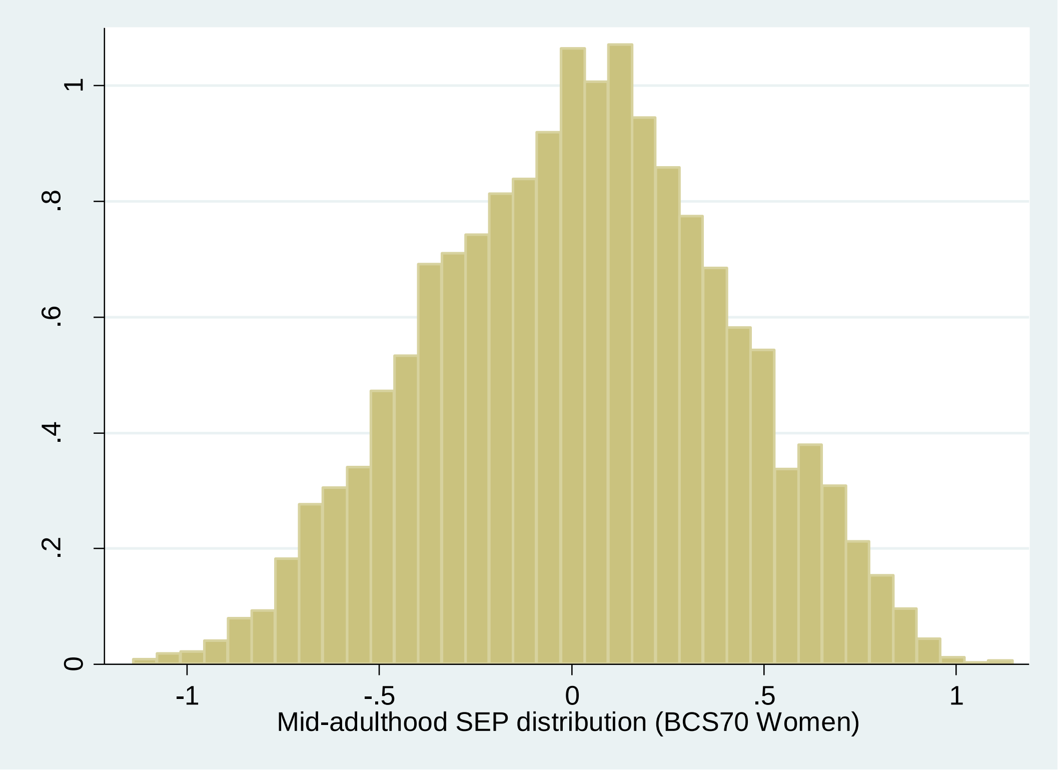


Figure 7.13.4: BCS70 Women estimated mid-adulthood SEP estimated construct score histogram



Appendix 7.14

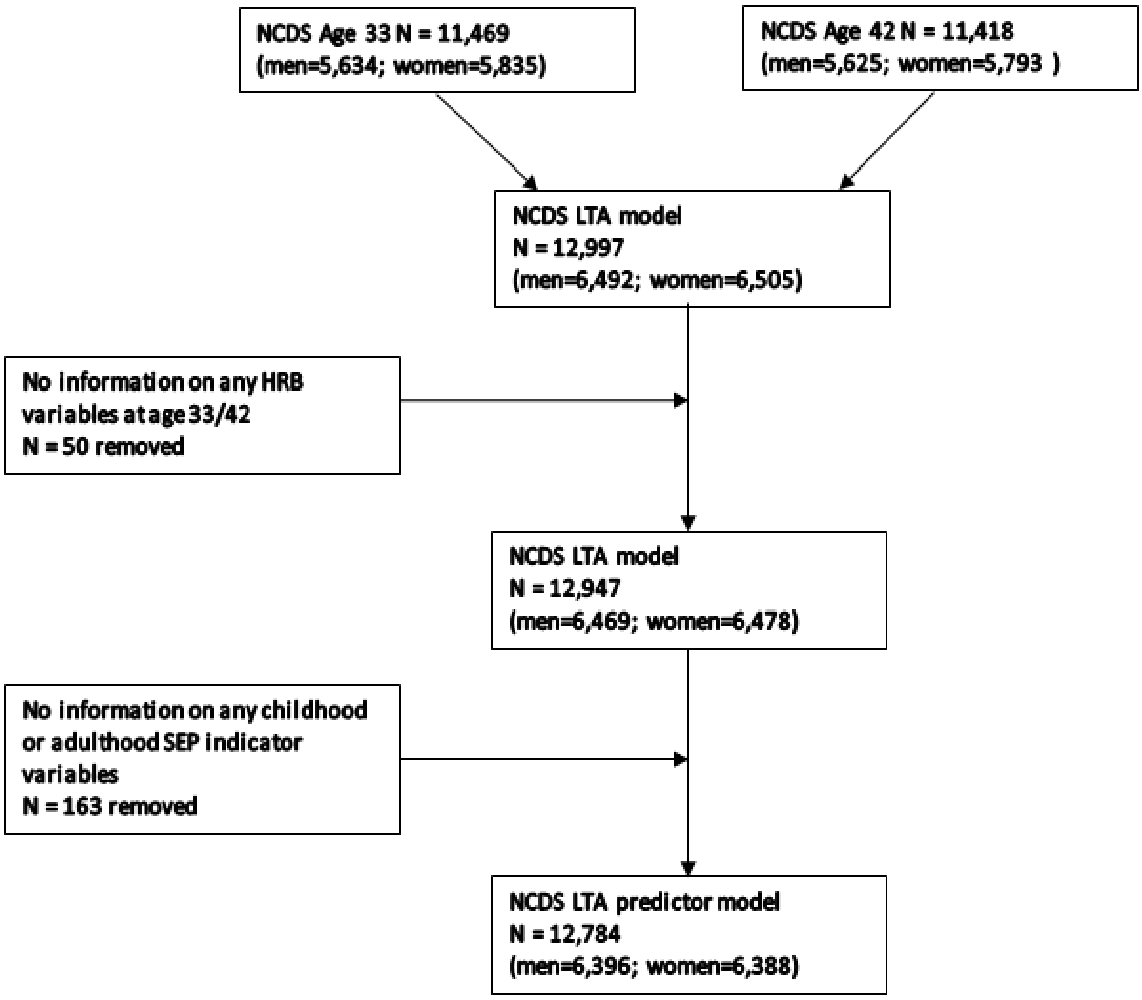
Correlations between individual HRBs and SEP constructs in pre-adolescence and mid-adulthood.

HRB variables	**SEP†**	**Men NCDS**	**Men BCS70**	**Women NCDS**	**Women BCS70**
		Pearson r	Pearson r	Pearson r	Pearson r
Number of cigarettes smoked per day	Pre-adolescent	0.23*	0.20*	0.31*	0.24*
(range 0-80)	Mid-adulthood	0.30*	0.29*	0.33*	0.31*
Frequency of fruit and vegetable consumption	Pre-adolescent	-0.16*	-0.17*	-0.25*	-0.20*
(never-everyday)	Mid-adulthood	-0.19*	-0.20*	-0.28*	-0.22*
Frequency of fried food consumption	Pre-adolescent	0.25*	0.08*	0.28*	0.06*
(never-everyday)	Mid-adulthood	0.30*	0.11*	0.30*	0.09*
Frequency of sweet food consumption	Pre-adolescent	-0.04**	-0.04**	0.01	0.01
(never-everyday)	Mid-adulthood	-0.02	-0.05*	-0.01	0.01
		Spearman rho	Spearman rho	Spearman rho	Spearman rho
Smoking status	Pre-adolescent	0.20*	0.19*	0.29*	0.24*
(non-smoker, smoker)	Mid-adulthood	0.30*	0.29*	0.30*	0.30*
Frequency of leisure-time physical activity	Pre-adolescent	-0.04**	-0.07*	-0.17*	-0.17*
(never-everyday)	Mid-adulthood	-0.05*	-0.04**	-0.18*	-0.20*
Alcohol units consumed in the previous week	Pre-adolescent	-0.07*	-0.07*	-0.08*	-0.04**
(no units, within limits, above limits)	Mid-adulthood	-0.09*	-0.10*	-0.07*	-0.03***

†Higher score = more disadvantaged SEP, *p≤0.001, **p≤0.01, *** p≤0.05.

Chapter 8 Appendices

Appendix 8.1
LTA analytical sample flowchart.



Note: In the final analytical sample N=12,784 (100%).
Participants with some data at age 33 and age 42 n=9,890 (77.4%).
Participants with no data at age 33 n=1,350 (10.6%).
Participants with no data at age 42 n=1,544 (12.1%).

Appendix 8.2
Missing data patterns for SEP indicator variables at age 11 and age 33 in the NCDS LTA analytical sample.

Variables (out of 20)	Frequency (N=12,784)	%
0	2,602	20.6
1	2,180	17.1
2	1,098	8.6
3	483	3.8
4	1,200	9.4
5	1,177	9.2
6	792	6.2
7	456	3.6
8	634	5.0
9	291	2.3
10	110	0.9
11	58	0.5
12	155	1.2
13	713	5.6
14	474	3.7
15	181	1.4
16	74	0.6
17	57	0.5
18	36	0.3
19	13	0.1

Appendix 8.3

Distribution for SEP in pre-adolescence and mid-adulthood in the NCDS LTA analytical sample.

Table 8.3.1: NCDS pre-adolescent and mid-adulthood SEP construct descriptive statistics

SEP construct*	NCDS Men		NCDS Women	
	Mean	Variance	Mean	Variance
Pre-adolescent	0	0.50	0	0.47
Mid-adulthood	0	0.39	0	0.39

*Product of CFA analysis using WLSMV estimator in Mplus Version 7.

Table 8.3.2: NCDS pre-adolescent and mid-adulthood SEP estimated construct (i.e. factor score) descriptive statistics

SEP estimated construct*	NCDS Men			NCDS Women		
	Mean	Variance	Range	Mean	Variance	Range
Pre-adolescent	0.01	0.76	-1.59, 1. 79	0.01	0.75	-1.85, 1.90
Mid-adulthood	0.02	0.83	-1.97, 2.10	0.02	0.84	-2.91, 2.12

*Estimated using regression methods in Mplus Version 7, descriptive analyses undertaken in Stata Version 14.

Figure 8.3.1: NCDS Men estimated mid-adulthood SEP estimated construct (i.e. factor score) histogram

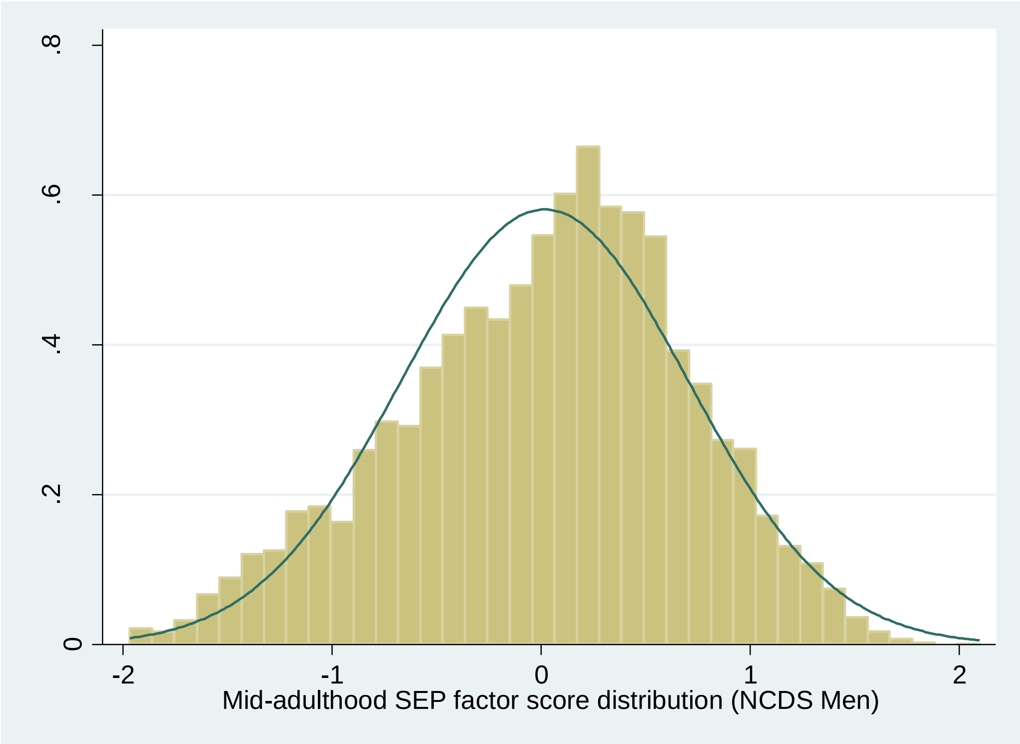
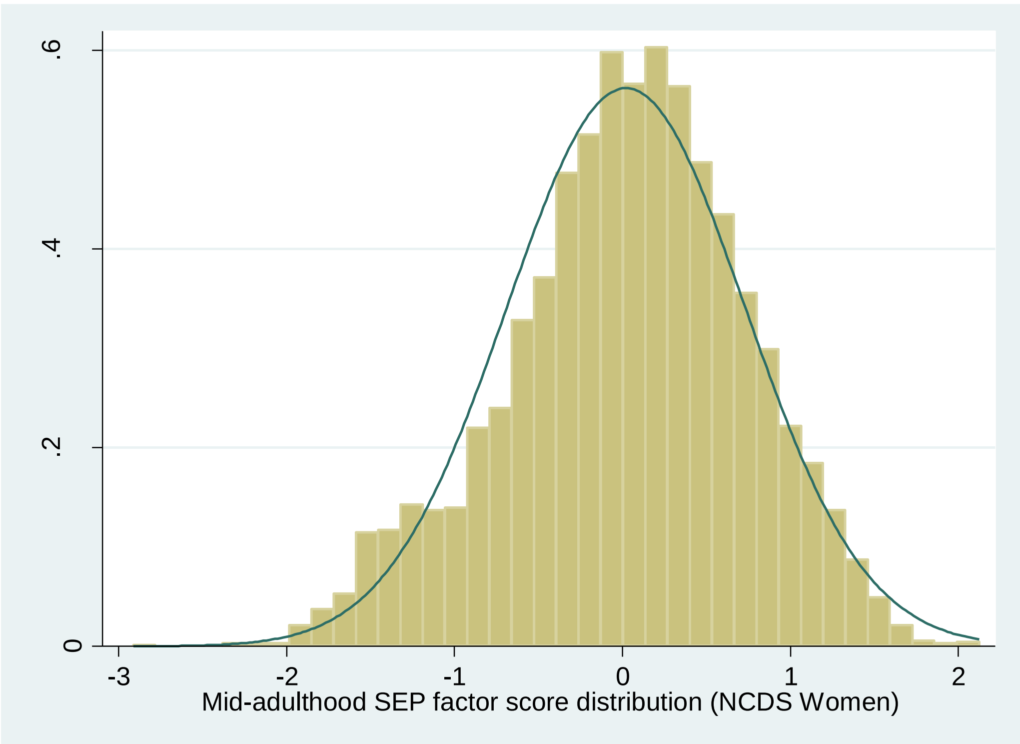


Figure 8.3.2: NCDS Women estimated mid-adulthood SEP estimated construct (i.e. factor score) histogram



Appendix 8.4
Estimates from exploratory 4, 5 and 6 cluster LPA and LTA models.

4-cluster LTA and LPA models
The 4-cluster LTA in men suggested cluster separation according to the number of cigarettes smoked per day. A cluster of very heavy smokers emerged (mean=35) whose alcohol consumption was also the highest, although their fried food consumption was lower than the 'Risky' cluster. This was considered theoretically reasonable, possibly reflecting reduced appetite due to heavier smoking and alcohol consumption. The transition probabilities were consistent with those identified in the 3-cluster LTA; additionally, they suggest a 10% probability of movement from the 'Risky' cluster at age 33 to this emerging cluster at age 42.

The 4-clustered patterns of the LPAs at ages 33 and 42 appeared similar, suggesting measurement invariance could be reasonably assumed. The patterns appeared to be more symmetrical in terms of smoking when compared to the 3-cluster LPAs. At both ages there was a cluster characterised by a particularly high number of cigarettes smoked per day (age 33 mean=44; age 42 mean=35). The prevalence of the cluster characterised by higher fried food consumption was lower at age 42 compared to age 33 which was consistent with the LTA results.

However, the clustered patterns from the LPAs were not consistent with the LTA in relation to alcohol consumption. Unlike the LTA, the LPAs indicated that the alcohol consumption was lower in the very heavy smokers cluster compared to the cluster whose members smoke around 20 cigarettes per day.

Amongst women, the 4-cluster LTA model also separated individuals according to their fried food and alcohol consumption but not the number of cigarettes smoked per day, which was found to be similar in the 'Risky' and emerging cluster. The probability of transitioning from the emerging cluster to the 'Risky' cluster was high (>70%).

The clustered patterns from the 4-cluster LPAs at ages 33 and 42 were relatively similar to one another in terms of smoking and diet, although physical activity patterns across the clusters did differ across time points. At age 33 the emerging cluster, characterised by heavy smoking, had higher levels of physical activity than the 'Risky' cluster whereas at age 42 these clusters had similar levels of physical activity.

The cluster patterns found in the LPA models were not consistent with those identified in the LTA model. For example, cluster separation for the LPAs was according to heavy smoking, as well as alcohol consumption at age 33 and fried food at age 42. However, there was consistency across the LPAs and LTAs in terms of cluster prevalence. In all models the ‘Risky’ cluster prevalence (characterised by higher consumption of fried food) was smaller at age 42.

Given that a notable characteristic of this emergent cluster in the LTAs and LPAs for both men and women was higher consumption of fried foods, this cluster was hereafter labelled ‘fried food’.

Table 8.4.1: NCDS Men 4 cluster LTA

NCDS Men Total N=6,396	Cluster 1†	Cluster 2 'Risky'	Cluster 3 'Moderate Smokers'	Cluster 4 'Mainstream'
Latent status prevalence	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Time 1 (Age 33)	159 (2.5%)	776 (12.1%)	1,161 (18.1%)	4,300 (67.2%)
Time 2 (Age 42)	219 (3.4%)	90 (1.4%)	1,502 (23.5%)	4,585 (71.7%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	35.3 (1.83)	19.42 (0.76)	16.02 (0.28)	0
Frequency of fruit and vegetable consumption	2.99 (0.18)	3.31 (0.17)	4.39 (0.07)	4.97 (0.03)
Frequency of fried food consumption	3.59 (0.15)	5.29 (0.23)	3.00 (0.07)	3.03 (0.02)
Frequency of sweet food consumption	3.44 (0.25)	3.84 (0.24)	4.31 (0.08)	4.64 (0.03)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity				
≤3 times a month	0.63 (0.05)	0.45 (0.03)	0.40 (0.01)	0.29 (0.01)
Once a week	0.13 (0.03)	0.19 (0.02)	0.19 (0.01)	0.21 (0.01)
2–3 days a week	0.11 (0.02)	0.17 (0.02)	0.19 (0.01)	0.25 (0.01)
4–7 days a week	0.13 (0.03)	0.19 (0.02)	0.23 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week				
No units	0.18 (0.03)	0.14 (0.02)	0.16 (0.01)	0.12 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.26 (0.04)	0.43 (0.05)	0.46 (0.02)	0.59 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.55 (0.05)	0.44 (0.04)	0.38 (0.02)	0.29 (0.01)
Transition probabilities from age 33 (rows) to age 42 (columns)				
Cluster 1	**0.87**	0[a]	0[a]	0.13
Cluster 2 'Risky'	0.10	**0.12**	0.61	0.18
Cluster 3 'Moderate Smokers'	0[a]	0[a]	**0.73**	0.28
Cluster 4 'Mainstream'	0[a]	0[a]	0.04	**0.95**

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 3-cluster LTA). Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.4.2: NCDS Women 4 cluster LTA

NCDS Women Total N=6,388	Cluster 1 'Risky'	Cluster 2†	Cluster 3 'Moderate Smokers'	Cluster 4 'Mainstream'
Latent status prevalence	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Time 1 (Age 33)	311 (4.5%)	644 (10.1%)	1,090 (17.0%)	4,342 (68.0%)
Time 2 (Age 42)	790 (11.7%)	69 (1.1%)	954 (14.9%)	4,575 (71.6%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	19.56 (0.70)	19.41 (0.63)	13.28 (0.30)	0
Frequency of fruit and vegetable consumption	3.70 (0.15)	3.76 (0.12)	5.96 (0.09)	6.07 (0.03)
Frequency of fried food consumption	2.52 (0.07)	4.56 (0.12)	2.33 (0.05)	2.41 (0.02)
Frequency of sweet food consumption	3.68 (0.19)	4.40 (0.28)	4.26 (0.13)	4.67 (0.03)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity				
≤3 times a month	0.62 (0.03)	0.54 (0.03)	0.28 (0.03)	0.29 (0.01)
Once a week	0.09 (0.01)	0.19 (0.02)	0.19 (0.01)	0.21 (0.01)
2–3 days a week	0.08 (0.01)	0.09 (0.02)	0.22 (0.01)	0.22 (0.01)
4–7 days a week	0.21 (0.02)	0.19 (0.02)	0.31 (0.01)	0.28 (0.01)
Alcohol units consumed in the previous week				
No units	0.32 (0.03)	0.29 (0.02)	0.27 (0.02)	0.25 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.37 (0.02)	0.59 (0.03)	0.60 (0.02)	0.64 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.31 (0.03)	0.11 (0.03)	0.13 (0.01)	0.11 (0.01)
Transition probabilities from age 33 (rows) to age 42 (columns)				
Cluster 1 'Risky'	**.94**	0[a]	0[a]	0.06
Cluster 2	0.74	**0.10**	0.03	0.13
Cluster 3 'Moderate Smokers'	0[a]	0[a]	**0.72**	0.28
Cluster 4 'Mainstream'	<0.01	<0.01	0.03	**0.96**

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 3-cluster LTA). Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.4.3: NCDS Men 4 cluster LPA age 33

NCDS Men Total N=5,586	Cluster 1 'Risky' n=71 (1.3%)	Cluster 2† n=358 (6.4%)	Cluster 3 'Moderate Smokers' n=1,346 (24.0%)	Cluster 4 'Mainstream' n=3,811 (68.2%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	43.75 (3.12)	21.44 (0.95)	16.38 (0.77)	0
Frequency of fruit and vegetable consumption	2.66 (0.44)	2.91 (0.16)	4.20 (0.26)	4.65 (0.03)
Frequency of fried food consumption	4.32 (0.27)	5.31 (0.74)	3.68 (0.14)	3.36 (0.02)
Frequency of sweet food consumption	3.75 (0.46)	3.12 (0.53)	4.44 (0.16)	4.71 (0.03)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity				
≤3 times a month	0.58 (0.10)	0.57 (0.07)	0.35 (0.03)	0.28 (0.01)
Once a week	0.14 (0.07)	0.13 (0.07)	0.23 (0.01)	0.21 (0.01)
2–3 days a week	0.14 (0.06)	0.12 (0.04)	0.21 (0.02)	0.25 (0.01)
4–7 days a week	0.13 (0.05)	0.18 (0.04)	0.22 (0.02)	0.26 (0.01)
Alcohol units consumed in the previous week				
No units	0.24 (0.07)	0.15 (0.05)	0.13 (0.01)	0.13 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.26 (0.07)	0.29 (0.15)	0.56 (0.02)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.49 (0.09)	0.57 (0.12)	0.31 (0.03)	0.24 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 3-cluster LPA at age 33).

Table 8.4.4: NCDS Men 4 cluster LPA age 42

NCDS Men Total N=5,529	**Cluster 1† n=161 (2.9%)**	**Cluster 2 'Risky' n=80 (1.4%)**	**Cluster 3 'Moderate Smokers' n=1,246 (22.5%)**	**Cluster 4 'Mainstream' n=4,042 (73.1%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	34.42 (1.70)	20.10 (2.21)	16.55 (0.31)	0
Frequency of fruit and vegetable consumption	3.09 (0.20)	3.28 (0.33)	4.33 (0.07)	5.28 (0.03)
Frequency of fried food consumption	3.34 (0.16)	6.20 (0.30)	2.78 (0.04)	2.70 (0.02)
Frequency of sweet food consumption	3.56 (0.30)	3.67 (0.37)	4.20 (0.08)	4.51 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity				
≤3 times a month	0.64 (0.05)	0.50 (0.07)	0.45 (0.02)	0.30 (0.01)
Once a week	0.13 (0.04)	0.14 (0.05)	0.17 (0.01)	0.21 (0.01)
2–3 days a week	0.10 (0.03)	0.08 (0.05)	0.17 (0.01)	0.24 (0.01)
4–7 days a week	0.13 (0.04)	0.28 (0.06)	0.21 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week				
No units	0.16 (0.04)	0.14 (0.05)	0.17 (0.01)	0.11 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.28 (0.05)	0.17 (0.06)	0.40 (0.02)	0.54 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.56 (0.06)	0.69 (0.07)	0.43 (0.02)	0.34 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 3-cluster LPA at age 42).

Table 8.4.5: NCDS Women 4 cluster LPA age 33

NCDS Women Total N=5,787	Cluster 1† n=68 (1.2%)	Cluster 2 'Risky' n=643 (11.1%)	Cluster 3 'Moderate Smokers' n=1,104 (19.1%)	Cluster 4 'Mainstream' n=3,973 (68.7%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	37.0 (2.39)	18.65 (0.56)	13.51 (0.38)	0
Frequency of fruit and vegetable consumption	3.71 (0.50)	3.61 (0.17)	5.72 (0.10)	5.79 (0.03)
Frequency of fried food consumption	3.71 (0.49)	3.87 (0.15)	2.62 (0.06)	2.55 (0.02)
Frequency of sweet food consumption	2.78 (0.35)	4.20 (0.25)	4.29 (0.14)	4.85 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity				
≤3 times a month	0.39 (0.07)	0.63 (0.04)	0.25 (0.03)	0.27 (0.01)
Once a week	0.26 (0.07)	0.15 (0.02)	0.22 (0.02)	0.24 (0.01)
2–3 days a week	0.19 (0.07)	0.05 (0.02)	0.22 (0.02)	0.21 (0.01)
4–7 days a week	0.17 (0.08)	0.17 (0.03)	0.32 (0.02)	0.28 (0.01)
Alcohol units consumed in the previous week				
No units	0.21 (0.07)	0.30 (0.03)	0.29 (0.02)	0.29 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.39 (0.08)	0.56 (0.03)	0.62 (0.02)	0.65 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.40 (0.08)	0.14 (0.02)	0.09 (0.01)	0.07 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 3-cluster LPA age 33).

Table 8.4.6: NCDS Women 4 cluster LPA age 42

NCDS Women Total N=5,683	Cluster 1† n=77 (1.4%)	Cluster 2 'Risky' n=86 (1.5%)	Cluster 3 'Moderate Smokers' n=1,376 (24.2%)	Cluster 4 'Mainstream' n=4,144 (72.9%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	36.19 (1.94)	18.02 (2.06)	14.93 (0.23)	0
Frequency of fruit and vegetable consumption	3.27 (0.30)	3.31 (0.37)	5.14 (0.08)	6.35 (0.03)
Frequency of fried food consumption	2.62 (0.19)	4.90 (0.54)	2.26 (0.04)	2.25 (0.02)
Frequency of sweet food consumption	4.28 (0.45)	5.15 (0.61)	3.92 (0.07)	4.47 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity				
≤3 times a month	0.72 (0.07)	0.70 (0.07)	0.44 (0.02)	0.31 (0.01)
Once a week	0.06 (0.03)	0.06 (0.04)	0.13 (0.01)	0.18 (0.01)
2–3 days a week	0.09 (0.04)	0.08 (0.04)	0.16 (0.01)	0.23 (0.01)
4–7 days a week	0.16 (0.05)	0.16 (0.05)	0.27 (0.01)	0.28 (0.01)
Alcohol units consumed in the previous week				
No units	0.41 (0.07)	0.51 (0.10)	0.28 (0.01)	0.22 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.30 (0.06)	0.24 (0.08)	0.48 (0.01)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.28 (0.06)	0.25 (0.07)	0.24 (0.01)	0.15 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 3-cluster LPA at age 42).

5 cluster LTA and LPA models

Amongst men, the 5-cluster LTA model further separated members of the ‘Moderate Smokers’ cluster according to the number of cigarettes smoked per day, physical activity levels, alcohol consumption and sweet food consumption. Members of the emergent cluster smoked fewer cigarettes per day, consumed less units of alcohol, were more physically active and had a higher consumption of sweet foods than the ‘Moderate Smokers’ cluster. The ‘fried food’ cluster that emerged in the 4-cluster LTA model remained stable, most of whom appear to transition to the ‘Moderate Smokers’ cluster at the second time point.

In the LPA models at each time point, the emergent cluster again separated according to smoking, physical activity and sweet food. However, alcohol consumption patterns were in the opposite direction, i.e. those who were more physically active drank more alcohol. The emergent cluster members had higher alcohol consumption and lower frequency of sweet food consumption. LPA clustered patterns were quite consistent at both time points although differences in physical activity levels between members of the ‘Moderate Smokers’ and the emerging cluster were less distinct at age 42 compared to age 33. The ‘fried food’ cluster that emerged in 4-cluster models remained stable.

The same cluster pattern identified for men was found amongst women in the 5-cluster LTA model. The emergent cluster smoke fewer cigarettes per day, consumed less alcohol units, were more physically active and had a higher consumption frequency of sweet foods, compared to members of the ‘Moderate Smokers’ cluster. The ‘fried food’ cluster that emerged in the 4-cluster LTA model remained stable, most of them appearing to transition to the ‘Moderate Smokers’ cluster.

For women the emergent cluster was not distinguished according to sweet food consumption or alcohol consumption, which were relatively similar for ‘Moderate Smokers’ and the emergent cluster in the LPAs at both ages. However, physical activity levels and smoking did separate these clusters from one another, i.e. the emergent cluster had higher levels of physical activity and members smoke fewer cigarettes per day. The ‘fried food’ cluster at age 33 in the LPA model was no longer aligned with that of the ‘fried food’ cluster at age 33 in the LTA model. The ‘fried food’ cluster pattern that emerged in 4-cluster models was instead characterised by heavier alcohol consumption at age 33 and the cluster prevalence at age 33 was smaller and similar to that at age 42. Overall across the time points the LPA clustered patterns appeared similar although differences in sweet food consumption and alcohol consumption between the 'Risky' and the emergent cluster were more apparent at age 33 compared to age 42.

Given that a notable characteristic of this emergent cluster in the LTAs and LPAs for both men and women was physical activity and smoking, this cluster was hereafter labelled 'active smokers'.

Table 8.4.7: NCDS Men 5 cluster LTA

NCDS Men Total N=6,396	Cluster 1 'Risky'	Cluster 2 'Moderate Smokers'	Cluster 3†	Cluster 4 'Fried Food'	Cluster 6 'Mainstream'
Latent status prevalence	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Time 1 (Age 33)	130 (2.0%)	310 (4.9%)	885 (13.8%)	770 (12.0%)	4,301 (67.2%)
Time 2 (Age 42)	130 (2.0%)	874 (13.7%)	739 (11.6%)	69 (1.1%)	4,585 (71.7%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	38.89 (18.90)	18.68 (11.0)	15.01 (0.90)	19.46 (0.74)	0
Frequency of fruit and vegetable consumption	3.13 (0.26)	3.54 (0.20)	4.91 (0.17)	3.19 (0.11)	4.97 (0.03)
Frequency of fried food consumption	3.69 (0.18)	2.98 (0.06)	3.10 (0.07)	5.23 (0.23)	3.03 (0.02)
Frequency of sweet food consumption	3.61 (0.29)	3.25 (0.30)	5.16 (0.31)	3.43 (0.28)	4.64 (0.03)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity					
≤3 times a month	0.61 (0.06)	0.49 (0.05)	0.36 (0.04)	0.45 (0.03)	0.29 (0.01)
Once a week	0.15 (0.04)	0.17 (0.02)	0.19 (0.02)	0.20 (0.02)	0.21 (0.01)
2–3 days a week	0.12 (0.03)	0.14 (0.04)	0.21 (0.03)	0.17 (0.02)	0.25 (0.01)
4–7 days a week	0.13 (0.03)	0.19 (0.02)	0.24 (0.03)	0.18 (0.02)	0.26 (0.01)
Alcohol units consumed in the previous week					
No units	0.21 (0.04)	0.13 (0.04)	0.18 (0.03)	0.12 (0.02)	0.12 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.30 (0.05)	0.30 (0.03)	0.56 (0.02)	0.43 (0.04)	0.59 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.49 (0.06)	0.56 (0.04)	0.27 (0.03)	0.45 (0.03)	0.29 (0.01)
Transition probabilities from age 33 (rows) to age 42 (columns)					
Cluster 1 'Risky'	**0.76**	0.10	0[a]	0[a]	0.15
Cluster 2 'Moderate Smokers'	0[a]	**0.89**	0[a]	0[a]	0.11
Cluster 3	0[a]	0[a]	**0.69**	0[a]	0.31
Cluster 4 'Fried Food'	0.03	0.68	0[a]	**0.09**	0.20
Cluster 6 'Mainstream'	<0.01	0.02	0.03	0[a]	**0.95**

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 4-cluster LTA).Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Supersrcipt a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.4.8: NCDS Women 5 cluster LTA

NCDS Women Total N=6,388	Cluster 1 'Risky'	Cluster 2†	Cluster 3 'Moderate Smokers'	Cluster 4 'Fried Food'	Cluster 5 'Mainstream'
Latent status prevalence	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Time 1 (Age 33)	86 (1.3%)	905 (14.2%)	382 (6.0%)	672 (10.5%)	4,343 (68.0%)
Time 2 (Age 42)	78 (1.2%)	748 (11.7%)	913 (14.3%)	75 (1.2%)	4,575 (71.6%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	37.01 (18.5)	12.23 (0.91)	17.54 (0.42)	18.31 (0.54)	0
Frequency of fruit and vegetable consumption	3.65 (0.3)	6.10 (0.09)	4.13 (0.24)	3.90 (0.16)	6.07 (0.03)
Frequency of fried food consumption	3.20 (0.2)	2.33 (0.06)	2.40 (0.08)	4.40 (0.12)	2.41 (0.02)
Frequency of sweet food consumption	3.46 (0.3)	4.22 (0.15)	3.76 (0.19)	4.64 (0.34)	4.67 (0.03)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity					
≤3 times a month	0.55 (0.05)	0.25 (0.02)	0.58 (0.04)	0.53 (0.03)	0.29 (0.01)
Once a week	0.17 (0.04)	0.21 (0.02)	0.09 (0.01)	0.18 (0.02)	0.21 (0.01)
2–3 days a week	0.14 (0.03)	0.24 (0.02)	0.09 (0.02)	0.08 (0.02)	0.22 (0.01)
4–7 days a week	0.14 (0.04)	0.31 (0.02)	0.25 (0.02)	0.21 (0.02)	0.28 (0.01)
Alcohol units consumed in the previous week					
No units	0.30 (0.06)	0.25 (0.03)	0.33 (0.03)	0.32 (0.03)	0.25 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.37 (0.05)	0.62 (0.02)	0.40 (0.03)	0.59 (0.03)	0.64 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.32 (0.06)	0.13 (0.02)	0.27 (0.04)	0.09 (0.02)	0.11 (0.01)
Transition probabilities from age 33 (rows) to age 42 (columns)					
Cluster 1 'Risky'	**0.63**	0[a]	0.28	0.02	0.06
Cluster 2	0[a]	**0.69**	0.01	0[a]	0.30
Cluster 3 'Moderate Smokers'	0.01	0[a]	**0.88**	0[a]	0.10
Cluster 4 'Fried Food'	0.03	0[a]	0.74	**0.10**	0.08
Cluster 5 'Mainstream'	0[a]	0.03	<0.01	<0.01	**0.96**

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 4-cluster LTA). Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.4.9: NCDS Men 5 cluster LPA age 33

NCDS Men age 33 (n=5,586)	Cluster 1 'Risky' n=71 (1.3%)	Cluster 2† n=748 (13.4%)	Cluster 3 'Moderate Smokers' n=635 (11.4%)	Cluster 4 'Fried Food' n=321 (5.8%)	Cluster 5 'Mainstream' n=3,811 (68%)
Item response probabilities (IRP) and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	43.8 (2.9)	15.4 (1.1)	17.9 (0.6)	21.1 (0.9)	0
Frequency of fruit and vegetable consumption	2.7 (0.4)	4.7 (0.5)	3.4 (0.2)	3.1 (0.3)	4.6 (0.03)
Frequency of fried food consumption	4.3 (0.3)	3.7 (0.2)	3.7 (0.2)	5.4 (0.5)	3.4 (0.02)
Frequency of sweet food consumption	3.6 (0.5)	3.7 (0.2)	5.4 (0.7)	2.8 (0.2)	4.7 (0.03)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity					
≤3 times a month	0.57 (0.09)	0.20 (0.09)	0.55 (0.06)	0.57 (0.08)	0.28 (0.01)
Once a week	0.15 (0.07)	0.28 (0.03)	0.16 (0.05)	0.13 (0.06)	0.21 (0.01)
2–3 days a week	0.15 (0.06)	0.28 (0.04)	0.12 (0.04)	0.12 (0.04)	0.25 (0.01)
4–7 days a week	0.13 (0.05)	0.25 (0.06)	0.18 (0.04)	0.18 (0.04)	0.26 (0.01)
Alcohol units consumed in the previous week					
No units	0.24 (0.07)	0.09 (0.02)	0.19 (0.05)	0.15 (0.05)	0.13 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.25 (0.07)	0.56 (0.04)	0.55 (0.04)	0.26 (0.11)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.51 (0.09)	0.36 (0.04)	0.26 (0.05)	0.59 (0.08)	0.24 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 4-cluster LPA at age 33).

Table 8.4.10: NCDS Men 5 cluster LPA age 42

NCDS Men age 42 (n=5,529)	Cluster 1 'Risky' n=154 (2.8%)	Cluster 2† n=955 (17.3%)	Cluster 3 'Moderate Smokers' n=301 (5.4%)	Cluster 4 'Fried Food' n=78 (1.4%)	Cluster 5 'Mainstream' n=4,042 (73%)
Item response probabilities (IRP) and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	34.7 (1.8)	16.3 (0.4)	17.7 (1.3)	20.5 (2.7)	0
Frequency of fruit and vegetable consumption	3.1 (0.2)	4.6 (0.02)	3.5 (0.4)	3.2 (0.4)	5.3 (0.03)
Frequency of fried food consumption	3.3 (0.2)	2.8 (0.08)	2.8 (0.3)	6.2 (0.3)	2.7 (0.02)
Frequency of sweet food consumption	3.4 (0.3)	3.5 (0.2)	6.0 (0.7)	3.6 (0.4)	4.5 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity					
≤3 times a month	0.64 (0.06)	0.41 (0.02)	0.56 (0.06)	0.50 (0.07)	0.30 (0.01)
Once a week	0.13 (0.04)	0.19 (0.02)	0.12 (0.04)	0.13 (0.05)	0.21 (0.01)
2–3 days a week	0.10 (0.03)	0.19 (0.02)	0.09 (0.02)	0.08 (0.04)	0.24 (0.01)
4–7 days a week	0.13 (0.04)	0.21 (0.02)	0.23 (0.05)	0.29 (0.06)	0.26 (0.01)
Alcohol units consumed in the previous week					
No units	0.15 (0.05)	0.12 (0.02)	0.34 (0.08)	0.13 (0.07)	0.11 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.28 (0.05)	0.39 (0.02)	0.44 (0.07)	0.17 (0.06)	0.54 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.57 (0.06)	0.49 (0.04)	0.22 (0.07)	0.70 (0.08)	0.34 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 4-cluster LPA at age 42).

Table 8.4.11: NCDS Women 5 cluster LPA age 33

NCDS Women age 33 (n=5,787)	Cluster 1 'Risky' n=56 (0.9%)	Cluster 2† n=1,035 (17.9%)	Cluster 3 'Moderate Smokers' n=642 (11.0%)	Cluster 4 'Fried Food' n=81 (1.4%)	Cluster 5 'Mainstream' n=3,973 (69%)
Item estimated means and standard errors	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	38.1 (2.5)	13.4 (0.5)	18.0 (0.5)	22.8 (2.8)	0
Frequency of fruit and vegetable consumption	3.9 (0.7)	5.8 (0.1)	3.8 (0.2)	3.2 (1.2)	4.6 (0.03)
Frequency of fried food consumption	3.3 (0.5)	2.6 (0.1)	3.4 (0.2)	6.1 (0.6)	3.4 (0.02)
Frequency of sweet food consumption	2.8 (0.4)	4.3 (0.2)	4.3 (0.3)	3.2 (0.5)	4.7 (0.03)
Item response probabilities and standard errors	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity					
≤3 times a month	0.39 (0.08)	0.22 (0.03)	0.64 (0.05)	0.53 (0.09)	0.27 (0.01)
Once a week	0.27 (0.09)	0.22 (0.02)	0.15 (0.03)	0.14 (0.07)	0.24 (0.01)
2–3 days a week	0.19 (0.08)	0.23 (0.02)	0.04 (0.03)	0.12 (0.07)	0.21 (0.01)
4–7 days a week	0.14 (0.07)	0.33 (0.02)	0.16 (0.04)	0.21 (0.07)	0.28 (0.01)
Alcohol units consumed in the previous week					
No units	0.23 (0.08)	0.28 (0.02)	0.31 (0.03)	0.20 (0.09)	0.28 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.39 (0.08)	0.63 (0.04)	0.56 (0.03)	0.53 (0.10)	0.65 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.38 (0.09)	0.09 (0.04)	0.12 (0.02)	0.27 (0.09)	0.07 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 4-cluster LPA at age 33).

Table 8.4.12: NCDS Women 5 cluster LPA age 42

NCDS Women age 42 (n=5,683)	Cluster 1 'Risky' n=154 (2.8%)	Cluster 2† n=815 (14.3%)	Cluster 3 'Moderate Smokers' n=598 (10.5%)	Cluster 'Fried Food' n=68 (1.2%)	Cluster 5 'Mainstream' n=4,143 (72.9%)
Item estimated means and standard errors	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	38.4 (1.7)	13.6 (0.7)	17.3 (0.5)	17.3 (2.1)	0
Frequency of fruit and vegetable consumption	3.4 (0.2)	5.9 (0.02)	3.9 (0.2)	3.6 (0.4)	6.4 (0.03)
Frequency of fried food consumption	2.6 (0.2)	2.2 (0.06)	2.4 (0.09)	5.2 (0.5)	2.3 (0.02)
Frequency of sweet food consumption	4.3 (0.5)	3.8 (0.2)	4.0 (0.3)	5.6 (0.7)	4.5 (0.04)
Item response probabilities and standard errors	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity					
≤3 times a month	0.65 (0.07)	0.26 (0.03)	0.71 (0.09)	0.65 (0.08)	0.31 (0.01)
Once a week	0.07 (0.04)	0.18 (0.02)	0.05 (0.04)	0.08 (0.05)	0.18 (0.01)
2–3 days a week	0.09 (0.05)	0.24 (0.02)	0.03 (0.05)	0.10 (0.05)	0.23 (0.01)
4–7 days a week	0.19 (0.06)	0.32 (0.03)	0.21 (0.03)	0.18 (0.06)	0.28 (0.01)
Alcohol units consumed in the previous week					
No units	0.43 (0.08)	0.24 (0.03)	0.34 (0.05)	0.54 (0.10)	0.22 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.31 (0.07)	0.55 (0.03)	0.38 (0.04)	0.25 (0.08)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.26 (0.07)	0.22 (0.03)	0.29 (0.04)	0.22 (0.08)	0.15 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 4-cluster LPA at age 42).

6-cluster LTA and LPA models

Amongst men, the 6-cluster LTA model suggested that the remaining members of the previously labelled 'Moderate Smokers' cluster separated according to sweet food and alcohol consumption, whilst smoking and physical activity remained similar. One of the clusters had very high alcohol consumption (higher than the 'Risky' cluster) the other had the highest consumption of sweet foods. The emergent 'active smokers' cluster remained stable in the 6-cluster LTA model.

The 6-cluster LPA models at both time points did not reflect the estimates of the LTA model. The emergent cluster prevalence in these LPA models was a very small cluster. Due to their size, the cluster interpretability was poor and the behavioural patterns in these emergent clusters did not match across the two time points.

Similar to men, the 6-cluster LTA model for women appeared to separate the 'Moderate Smokers' cluster according to alcohol and sweet food consumption. Yet unlike men, these clusters remained similar in terms of their smoking and physical activity levels. The 'active smokers' cluster that emerged in the 5-cluster model was stable in the 6-cluster LTA model.

Given the separation of the 'Moderate Smokers' cluster for men and women in the 6-cluster LTA models the 'Moderate Smokers' cluster label was replaced by 'inactive smokers and alcohol' and 'inactive smokers and sweet food'.

Cluster emergence amongst women in the 6-cluster LPA models followed a similar pattern to that of women in the LTA model. At age 42 members of the 'Moderate Smokers' cluster separated according to their sweet food and alcohol consumption. At age 33 a sweet food consumption cluster also emerged although the 'fried food' cluster continued to have higher alcohol consumption than the 'Moderate Smokers' cluster. This meant that the clustered patterns for alcohol consumption across the two time points differed. At age 42 alcohol consumption was higher in the emergent cluster (characterised by high alcohol consumption) compared to the 'Risky' cluster, this was not the case at age 33. Similar to the 5-cluster LPA models, the prevalence of the 'Fried food' cluster at age 33 differed to that found in the LTA, and was small at both time points. The 'active smokers' cluster that emerged in the 5-cluster models was stable in the 6-cluster LPA models.

For men and women, cluster transition probabilities were small. The largest probability of movement was for members of the 'fried food' cluster who transitioned to the cluster characterised by heavy alcohol consumption, labelled 'inactive smokers and alcohol'.

Table 8.4.13: NCDS Men 6 cluster LTA

NCDS Men Total N=6,396	Cluster 1 'Risky'	Cluster 2* 'Active Smokers'	Cluster 3† 'Inactive Smokers and Sweet Food'	Cluster 4* 'Inactive Smokers and Alcohol'	Cluster 5* 'Inactive Smokers and Fried Food'	Cluster 6 'Mainstream'
Latent status prevalence	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Time 1 (Age 33)	121 (1.9%)	638 (10.0%)	381 (6.0%)	292 (4.6%)	663 (10.4%)	4,300 (67.2%)
Time 2 (Age 42)	115 (1.8%)	525 (8.2%)	348 (5.4%)	760 (11.9%)	63 (1.0%)	4,585 (71.7%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	40.01 (1.49)	12.71 (1.08)	18.29 (1.28)	19.66 (0.77)	20.22 (0.80)	0
Frequency of fruit and vegetable consumption	3.20 (0.24)	5.19 (0.18)	3.71 (0.23)	3.59 (0.21)	3.26 (0.13)	4.69 (0.03)
Frequency of fried food consumption	3.73 (0.19)	3.07 (0.08)	3.39 (0.14)	2.94 (0.06)	5.38 (0.21)	3.03 (0.02)
Frequency of sweet food consumption	3.44 (0.27)	4.03 (0.31)	6.55 (0.32)	3.15 (0.13)	3.41 (0.17)	4.64 (0.03)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity						
≤3 times a month	0.59 (0.05)	0.28 (0.03)	0.53 (0.04)	0.50 (0.04)	0.44 (0.03)	0.29 (0.01)
Once a week	0.16 (0.04)	0.22 (0.02)	0.14 (0.03)	0.17 (0.02)	0.20 (0.03)	0.21 (0.01)
2–3 days a week	0.13 (0.03)	0.26 (0.02)	0.11 (0.02)	0.14 (0.02)	0.17 (0.02)	0.25 (0.01)
4–7 days a week	0.12 (0.03)	0.24 (0.02)	0.23 (0.03)	0.20 (0.02)	0.18 (0.02)	0.26 (0.01)
Alcohol units consumed in the previous week						
No units	0.19 (0.04)	0.09 (0.03)	0.29 (0.04)	0.14 (0.03)	0.12 (0.02)	0.12 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.30 (0.04)	0.56 (0.04)	0.51 (0.05)	0.30 (0.04)	0.41 (0.04)	0.59 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.51 (0.06)	0.35 (0.05)	0.20 (0.04)	0.56 (0.04)	0.47 (0.04)	0.29 (0.01)
Transition probabilities from age 33 (rows) to age 42 (columns)						
Cluster 1 'Risky'	**0.68**	0[a]	0[a]	0.15	0[a]	0.17
Cluster 2 'Active Smokers'	0[a]	**0.64**	0[a]	0[a]	0[a]	0.36
Cluster 3 'Inactive Smokers and Sweet Food'	0.01	0[a]	**0.78**	0[a]	0[a]	0.22
Cluster 4 'Inactive Smokers and Alcohol'	0[a]	0[a]	0[a]	**0.92**	0[a]	0.08
Cluster 5 'Inactive Smokers and Fried Food'	0.04	0[a]	0.02	0.66	**0.09**	0.19
Cluster 6 'Mainstream'	<0.01	0.03	<0.01	<0.01	0[a]	**0.95**

Note: Cluster prevalence based on estimated model. *=re-labelled. †=Additional cluster (not identified in 5-cluster LTA). Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.4.14: NCDS Women 6 cluster LTA

NCDS Women Total N=6,388	Cluster 1 'Risky'	Cluster 2* 'Active Smokers'	Cluster 3† 'Inactive Smokers and Sweet Food'	Cluster 4* 'Inactive Smokers and Alcohol'	Cluster 5* 'Inactive Smokers an Fried Food'	Cluster 6 'Mainstream'
Latent status prevalence	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**	**n (%)**
Time 1 (Age 33)	83 (1.2%)	724 (11%)	511 (8.0%)	236 (3.7%)	490 (7.7%)	4343 (68.0%)
Time 2 (Age 42)	79 (1.3%)	591 (9%)	601 (9.4%)	494 (7.7%)	48 (0.7%)	4575 (71.6%)
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	37.21 (1.69)	10.83 (1.43)	17.34 (0.68)	18.03 (0.51)	18.48 (0.72)	0
Frequency of fruit and vegetable consumption	3.55 (0.27)	6.12 (0.15)	4.63 (0.22)	4.20 (0.31)	3.99 (0.18)	6.07 (0.03)
Frequency of fried food consumption	3.18 (0.23)	2.37 (0.07)	2.49 (0.11)	2.52 (0.08)	4.67 (0.18)	2.41 (0.02)
Frequency of sweet food consumption	3.58 (0.34)	4.15 (0.14)	5.01 (0.17)	2.95 (0.29)	3.91 (0.26)	4.67 (0.03)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity						
≤3 times a month	0.57 (0.05)	0.23 (0.03)	0.51 (0.04)	0.52 (0.06)	0.54 (0.04)	0.29 (0.01)
Once a week	0.17 (0.04)	0.22 (0.02)	0.12 (0.02)	0.13 (0.03)	0.17 (0.03)	0.21 (0.01)
2–3 days a week	0.13 (0.03)	0.25 (0.02)	0.11 (0.02)	0.01 (0.02)	0.09 (0.02)	0.22 (0.01)
4–7 days a week	0.13 (0.04)	0.30 (0.02)	0.27 (0.02)	0.26 (0.03)	0.20 (0.03)	0.28 (0.01)
Alcohol units consumed in the previous week						
No units	0.31 (0.06)	0.19 (0.04)	0.61 (0.01)	0	0.26 (0.05)	0.25 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.40 (0.05)	0.66 (0.03)	0.39 (0.07)	0.49 (0.06)	0.63 (0.04)	0.64 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.30 (0.05)	0.15 (0.02)	<0.01 (0.03)	0.51 (0.06)	0.11 (0.03)	0.11 (0.01)
Transition probabilities from age 33 (rows) to age 42 (columns)						
Cluster 1 'Risky'	**0.63**	0[a]	0.07	0.24	0[a]	0.07
Cluster 2 'Active Smokers'	0[a]	**0.65**	0[a]	0.02	0[a]	0.33
Cluster 3 'Inactive Smokers and Sweet Food'	0.01	0[a]	**0.82**	0.01	0[a]	0.15
Cluster 4 'Inactive Smokers and Alcohol'	0[a]	0[a]	0[a]	**0.93**	0[a]	0.07
Cluster 5 'Inactive Smokers and Fried Food'	0.04	0[a]	0.30	0.44	**0.09**	0.13
Cluster 6 'Mainstream'	0[a]	0.03	<0.01	<0.01	<0.01	**0.96**

Note: Cluster prevalence based on estimated model. *=re-labelled. †=Additional cluster (not identified in 5-cluster LTA). Transitions probabilities in bold correspond to staying in the same HRB cluster. Measurement invariance assumed over time (i.e. item means and response probabilities restricted to be equal across time). Transition probabilities sum to 1.0 (with rounding error) across rows. Superscript a = transitions not estimated in model but instead fixed at 0 in Mplus Version 7.

Table 8.4.15: NCDS Men 6 cluster LPA age 33

NCDS Men age 33 (n=5,586)	**Cluster 1 'Risky' n=121 (2.2%)**	**Cluster 'Active smokers' n=685 (12.3%)**	**Cluster 3 'Moderate Smokers' n=704 (12.6%)**	**Cluster 4 'Fried Food' n=249 (4.5%)**	**Cluster 5† n=16 (0.3%)**	**Cluster 6 'Mainstream' n=3,811 (68.2%)**
Item response probabilities (IRP) and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	34.41 (1.62)	14.49 (1.32)	17.77 (0.55)	21.25 (0.98)	59.49 (2.81)	0
Frequency of fruit and vegetable consumption	3.20 (0.36)	4.79 (0.47)	3.42 (0.17)	3.10 (0.31)	2.48 (0.56)	4.65 (0.03)
Frequency of fried food consumption	4.16 (0.20)	3.67 (0.13)	3.78 (0.17)	5.64 (0.50)	4.61 (0.59)	3.37 (0.02)
Frequency of sweet food consumption	3.71 (0.39)	3.66 (0.17)	5.21 (0.66)	2.68 (0.25)	3.57 (0.70)	4.71 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity *≤3 times a month* *Once a week* *2–3 days a week* *4–7 days a week*	0.43 (0.08) 0.27 (0.07) 0.15 (0.05) 0.15 (0.05)	0.19 (0.07) 0.27 (0.04) 0.28 (0.04) 0.26 (0.06)	0.54 (0.07) 0.16 (0.06) 0.12 (0.04) 0.17 (0.04)	0.58 (0.06) 0.11 (0.05) 0.12 (0.04) 0.18 (0.05)	0.75 (0.08) 0.00 (0.55) 0.12 (0.29) 0.13 (0.26)	0.28 (0.01) 0.21 (0.01) 0.25 (0.01) 0.26 (0.01)
Alcohol units consumed in the previous week *No units* *Within limits (≤14 units women, ≤21 units men)* *Above limits (≥15 units women, ≥22 units men)*	0.15 (0.06) 0.29 (0.07) 0.56 (0.08)	0.09 (0.02) 0.56 (0.04) 0.36 (0.04)	0.18 (0.05) 0.55 (0.04) 0.26 (0.04)	0.15 (0.05) 0.24 (0.09) 0.62 (0.08)	0.45 (0.14) 0.20 (0.11) 0.36 (0.14)	0.13 (0.01) 0.63 (0.01) 0.24 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 5-cluster LPA at age 33).

Table 8.4.16: NCDS Men 6 cluster LPA age 42

NCDS Men age 42 (n=5,529)	Cluster 1 'Risky' n=151 (2.7%)	Cluster 'Active Smokers' n=968 (17.8%)	Cluster 3 'Moderate Smokers' n=281 (5.1%)	Cluster 4 'Fried Food' n=82 (1.5%)	Cluster 5† n=10 (0.2%)	Cluster 6 'Mainstream' n=4,038 (73.0%)
Item response probabilities (IRP) and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	34.90 (1.63)	16.21 (0.30)	17.88 (1.04)	21.74 (1.94)	2.89 (7.48)	0
Frequency of fruit and vegetable consumption	3.16 (0.21)	4.58 (0.03)	3.39 (0.20)	3.04 (0.26)	7.07 (2.53)	5.27 (0.03)
Frequency of fried food consumption	3.28 (0.15)	2.80 (0.05)	2.71 (0.17)	6.03 (0.27)	6.94 (1.88)	2.70 (0.02)
Frequency of sweet food consumption	3.40 (0.33)	3.57 (0.15)	6.39 (0.10)	3.51 (0.36)	6.65 (2.27)	4.51 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity						
≤3 times a month	0.64 (0.06)	0.41 (0.02)	0.56 (0.06)	0.51 (0.08)	0.38 (1.28)	0.30 (0.01)
Once a week	0.14 (0.04)	0.19 (0.02)	0.12 (0.04)	0.13 (0.05)	0.10 (0.11)	0.21 (0.01)
2–3 days a week	0.10 (0.03)	0.19 (0.02)	0.09 (0.03)	0.09 (0.05)	0.00 (1.94)	0.24 (0.01)
4–7 days a week	0.12 (0.04)	0.21 (0.02)	0.23 (0.05)	0.28 (0.06)	0.52 (0.70)	0.26 (0.01)
Alcohol units consumed in the previous week						
No units	0.15 (0.04)	0.12 (0.02)	0.35 (0.06)	0.11 (0.07)	0.59 (0.32)	0.11 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.28 (0.05)	0.39 (0.02)	0.43 (0.06)	0.19 (0.06)	0.21 (0.32)	0.54 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.57 (0.06)	0.48 (0.03)	0.23 (0.05)	0.70 (0.08)	0.20 (0.18)	0.34 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 5-cluster LPA at age 42).

Table 8.4.17: NCDS Women 6 cluster LPA age 33

NCDS Women age 33 (n=5,787)	Cluster 1 'Risky' n=55 (0.9%)	Cluster 2 'Active Smokers' n=952 (16.5%)	Cluster 3 'Moderate Smokers' n=467 (8.1%)	Cluster 4 'Fried Food and alcohol' n=76 (1.3%)	Cluster 5† 'sweet food' n=264 (4.6%)	Cluster 6 'Mainstream' n=3,973 (68.7%)
Item response probabilities (IRP) and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	38.40 (3.90)	12.97 (1.10)	17.86 (0.88)	23.00 (2.8)	18.31 (0.84)	0
Frequency of fruit and vegetable consumption	3.89 (1.30)	5.83 (0.17)	4.00 (0.65)	3.27 (3.12)	3.81 (1.37)	5.79 (0.03)
Frequency of fried food consumption	3.36 (0.76)	2.62 (0.10)	3.28 (0.26)	6.12 (0.73)	3.63 (2.18)	2.55 (0.02)
Frequency of sweet food consumption	2.83 (0.72)	4.22 (0.47)	3.38 (0.28)	3.18 (1.43)	6.21 (0.72)	4.85 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity						
≤3 times a month	0.39 (0.08)	0.21 (0.03)	0.64 (0.13)	0.53 (0.36)	0.57 (0.40)	0.27 (0.01)
Once a week	0.27 (0.10)	0.23 (0.02)	0.17 (0.04)	0.14 (0.15)	0.12 (0.15)	0.24 (0.01)
2–3 days a week	0.19 (0.08)	0.24 (0.04)	0.05 (0.07)	0.12 (0.07)	0.05 (0.06)	0.21 (0.01)
4–7 days a week	0.14 (0.07)	0.33 (0.03)	0.14 (0.06)	0.22 (0.21)	0.26 (0.22)	0.28 (0.01)
Alcohol units consumed in the previous week						
No units	0.24 (0.10)	0.28 (0.05)	0.20 (0.07)	0.17 (0.62)	0.52 (0.16)	0.29 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.39 (0.10)	0.63 (0.03)	0.60 (0.06)	0.55 (0.24)	0.48 (0.16)	0.65 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.37 (0.09)	0.09 (0.03)	0.20 (0.05)	0.27 (0.41)	0.00 (0.00)	0.07 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 5-cluster LPA at age 33).

Table 8.4.18: NCDS Women 6 cluster LPA age 42

NCDS Women age 42 (n=5,683)	**Cluster 1 'Risky' n=57 (1.0%)**	**Cluster 2 'Active Smokers' n=773 (13.6%)**	**Cluster 3 'Moderate Smokers and alcohol' n=423 (7.4%)**	**Cluster 'Fried Food' n=52 (0.9%)**	**Cluster 5† 'sweet food' n=237 (4.2%)**	**Cluster 6 'Mainstream' n=4,141 (72.9%)**
Item response probabilities (IRP) and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	38.78 (1.61)	13.36 (0.04)	17.76 (0.69)	15.92 (0.35)	16.87 (0.81)	0
Frequency of fruit and vegetable consumption	3.42 (0.34)	5.98 (0.13)	3.89 (0.21)	3.86 (0.61)	4.08 (0.30)	6.35 (0.03)
Frequency of fried food consumption	2.63 (0.20)	2.18 (0.05)	2.69 (0.15)	5.52 (0.51)	2.03 (0.19)	2.25 (0.02)
Frequency of sweet food consumption	4.27 (0.48)	3.77 (0.17)	3.03 (0.19)	4.83 (0.83)	6.30 (0.71)	4.47 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity						
≤3 times a month	0.65 (0.07)	0.28 (0.04)	0.65 (0.07)	0.66 (0.11)	0.68 (0.09)	0.31 (0.01)
Once a week	0.07 (0.04)	0.17 (0.02)	0.09 (0.03)	0.08 (0.06)	0.05 (0.03)	0.18 (0.01)
2–3 days a week	0.09 (0.05)	0.24 (0.03)	0.06 (0.03)	0.11 (0.06)	0.02 (0.04)	0.23 (0.01)
4–7 days a week	0.19 (0.06)	0.31 (0.03)	0.20 (0.04)	0.16 (0.08)	0.26 (0.06)	0.28 (0.01)
Alcohol units consumed in the previous week						
No units	0.44 (0.08)	0.25 (0.03)	0.18 (0.09)	0.58 (0.12)	0.59 (0.06)	0.22 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.31 (0.07)	0.56 (0.03)	0.38 (0.05)	0.25 (0.09)	0.35 (0.05)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.25 (0.07)	0.20 (0.03)	0.44 (0.08)	0.17 (0.11)	0.06 (0.04)	0.15 (0.01)

Note: Cluster prevalence based on estimated model. †=Additional cluster (not identified in 5-cluster LPA at age 42).

Appendix 8.5
Estimates from 3-cluster LPA models at ages 33 and 42

The LPA models show that at age 33 and 42 the cluster patterns for women are similar suggesting the nature of the HRB clusters is the same at both time points. For men, the cluster patterns are similar except the mean number of cigarettes smoked per day at age 33 for members of the 'Risky' cluster is substantially higher than those at age 42 (age 33 mean=41 cigarettes per day, age 42 mean=22 cigarettes per day). This is reflected in the larger cluster prevalence of the 'Risky' cluster at age 33 in the 'fixed' LTA (see tables 8.8 and 8.9 in section 8.3.3.2) when compared to the prevalence of the 'Risky' cluster at age 33 in the LPA. Imposing measurement invariance in the 'fixed' LTA has reduced the mean number of cigarettes smoked per day in the 'Risky' cluster and therefore a large proportion of individuals assigned to the 'Moderate Smokers' cluster at age 33 in the LPA models are assigned to the 'Risky' cluster in the 'fixed' LTA models who subsequently have a high probability of transitioning to the 'Moderate Smokers' cluster at age 42.

The clustered patterns of the 'fixed' 3 cluster LTAs are similar to those of the LPAs at age 33 and are considered to have good interpretability. The latent transitions make theoretical sense and are in line with the literature which suggests that HRBs are relatively stable during mid-age and that as people age their HRBs tend to improve. This is demonstrated by the relatively high probability (>70%) of remaining in the 'Moderate Smokers' and 'Mainstream' cluster for men and women as well as the 'Risky' cluster for women. However, amongst those who do move they tend to transition to clusters characterised by HRBs that are more beneficial for health.

Table 8.5.1 Estimates from 3-cluster LPA model age 33 in the NCDS (men)

NCDS Men Total N=5,586	**Cluster 1 'Risky'** **n=96 (1.7%)**	**Cluster 2 'Moderate Smokers'** **n=1,679 (30.1%)**	**Cluster 3 'Mainstream'** **n=3,811 (68.2%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	40.84 (3.67)	17.22 (0.31)	0
Frequency of fruit and vegetable consumption	2.61 (0.37)	3.95 (0.05)	4.64 (0.03)
Frequency of fried food consumption	4.73 (0.45)	3.99 (0.05)	3.36 (0.02)
Frequency of sweet food consumption	3.58 (0.45)	4.18 (0.06)	4.71 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.61 (0.07)	0.39 (0.01)	0.28 (0.01)
Once a week	0.12 (0.05)	0.21 (0.01)	0.21 (0.01)
2–3 days a week	0.14 (0.04)	0.19 (0.01)	0.25 (0.01)
4–7 days a week	0.13 (0.04)	0.21 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week			
No units	0.26 (0.08)	0.14 (0.01)	0.13 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.23 (0.06)	0.50 (0.01)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.51 (0.08)	0.36 (0.01)	0.24 (0.01)

Note: Cluster prevalence based on estimated model.

Table 8.5.2 Estimates from 3-cluster LPA model age 42 in the NCDS (men)

NCDS Men Total N=5,529	**Cluster 1 'Risky'** **n=93 (1.7%)**	**Cluster 2 'Moderate Smokers'** **n=1,395 (25.2%)**	**Cluster 3 'Mainstream'** **n=4,041 (73.1%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	22.01 (3.32)	18.43 (0.31)	0
Frequency of fruit and vegetable consumption	3.26 (0.32)	4.20 (0.06)	5.28 (0.03)
Frequency of fried food consumption	6.05 (0.42)	2.82 (0.04)	2.70 (0.02)
Frequency of sweet food consumption	3.75 (0.36)	4.12 (0.07)	4.51 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.52 (0.07)	0.47 (0.01)	0.30 (0.01)
Once a week	0.12 (0.05)	0.17 (0.01)	0.21 (0.01)
2–3 days a week	0.08 (0.05)	0.16 (0.01)	0.24 (0.01)
4–7 days a week	0.27 (0.06)	0.21 (0.01)	0.26 (0.01)
Alcohol units consumed in the previous week			
No units	0.13 (0.05)	0.17 (0.01)	0.11 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.19 (0.07)	0.39 (0.01)	0.54 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.68 (0.07)	0.44 (0.01)	0.34 (0.01)

Note: Cluster prevalence based on estimated model.

Table 8.5.3 Estimates from 3-cluster LPA model age 33 in the NCDS (women)

NCDS Women Total N=5,787	**Cluster 1 'Risky'** **n=561 (9.7%)**	**Cluster 2 'Moderate Smokers'** **n=1,254 (21.7%)**	**Cluster 3 'Mainstream'** **n=3,972 (68.6%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	20.96 (0.96)	14.07 (0.31)	0
Frequency of fruit and vegetable consumption	3.39 (0.15)	5.57 (0.14)	5.79 (0.03)
Frequency of fried food consumption	4.02 (0.15)	2.69 (0.07)	2.55 (0.02)
Frequency of sweet food consumption	3.76 (0.24)	4.40 (0.10)	4.85 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.62 (0.03)	0.29 (0.03)	0.27 (0.01)
Once a week	0.17 (0.02)	0.21 (0.01)	0.24 (0.01)
2–3 days a week	0.07 (0.02)	0.20 (0.02)	0.21 (0.01)
4–7 days a week	0.15 (0.02)	0.31 (0.02)	0.28 (0.01)
Alcohol units consumed in the previous week			
No units	0.27 (0.03)	0.30 (0.02)	0.29 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.54 (0.03)	0.61 (0.02)	0.65 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.19 (0.03)	0.09 (0.01)	0.07 (0.01)

Note: Cluster prevalence based on estimated model.

Table 8.5.4 Estimates from 3-cluster LPA model age 42 in the NCDS (women)

NCDS Women Total N=5,683	**Cluster 1 'Risky'** **n=252 (4.4%)**	**Cluster 2 'Moderate Smokers'** **n=1,284 (22.6%)**	**Cluster 3 'Mainstream'** **n=4,146 (73.0%)**
Item estimated means and standard errors (S.E)	**Mean (S.E)**	**Mean (S.E)**	**Mean (S.E)**
Number of cigarettes smoked per day	23.33 (1.74)	14.77 (0.06)	0
Frequency of fruit and vegetable consumption	3.09 (0.16)	5.31 (0.19)	6.35 (0.03)
Frequency of fried food consumption	3.51 (0.54)	2.20 (0.05)	2.26 (0.02)
Frequency of sweet food consumption	4.35 (0.61)	3.94 (0.11)	4.47 (0.04)
Item response probabilities (IRP) and standard errors (S.E)	**IRP (S.E)**	**IRP (S.E)**	**IRP (S.E)**
Frequency of leisure-time physical activity			
≤3 times a month	0.75 (0.04)	0.41 (0.04)	0.31 (0.01)
Once a week	0.07 (0.03)	0.13 (0.01)	0.18 (0.01)
2–3 days a week	0.04 (0.02)	0.17 (0.02)	0.23 (0.01)
4–7 days a week	0.15 (0.03)	0.29 (0.02)	0.28 (0.01)
Alcohol units consumed in the previous week			
No units	0.39 (0.07)	0.28 (0.02)	0.22 (0.01)
Within limits (≤14 units women, ≤21 units men)	0.30 (0.04)	0.49 (0.02)	0.63 (0.01)
Above limits (≥15 units women, ≥22 units men)	0.30 (0.07)	0.23 (0.02)	0.15 (0.01)

Note: Cluster prevalence based on estimated model.

Appendix 8.6

Analysis of differences in individual HRBs for those who moved HRB cluster between ages 33 and 42 in comparison to those who remained in the same HRB cluster over time.

Tables 8.6.1 to 8.6.3 show how individual HRBs differed for those that moved HRB clusters between ages 33 and 42 compared to those who did not move. Not surprisingly, differences for individual HRBs at age 42 were found when comparing those that stayed in the same HRB cluster over time compared to those that moved. For example, for men and women the proportion undertaking physical activity ≤3 times per week was significantly higher at age 42 for those who moved from the 'Risky' cluster to either the 'Moderate Smokers' or 'Mainstream' cluster in comparison to those who stayed in the 'Risky' cluster between ages 33 and 42 (men=58.0% vs 33.3%; women=73.4% vs 39.5%, p≤0.05).

Interestingly, tables 8.6.1 to 8.6.3 also show differences for some individual HRBs at age 33 when comparing those that stayed in the same HRB cluster over time with those that moved. This suggests that differentials exist for some individual HRBs amongst participants assigned to the same HRB cluster at age 33 and that these differences may to some extent predict whether they remain in the same HRB cluster over time.

For example, amongst men assigned to the 'Risky' cluster at age 33, those who moved from the 'Risky' cluster smoked significantly fewer cigarettes than those who stayed in the 'Risky' cluster (23.6 vs 20.5, p≤0.05). Moreover, the proportion drinking above recommended limits was significantly higher for men who stayed in the 'Risky' cluster compared to those who moved (50.0% vs 39.4%, p≤0.05). A similar result was found for alcohol consumption amongst women (18.2% vs 11.6%), although this was not statistically significant (p>0.05). For men and women, the frequency of fried food consumption at age 33 in the 'Risky' cluster was higher amongst movers than stayers (men=5.4 vs 5.8; women 4.2 vs 4.6, p≤0.05).These results imply that patterns of smoking, alcohol consumption and fried food consumption may to some extent differentiate those who persist with 'Risky' HRBs during mid-life from those that move to a HRB cluster characterised by more positive HRBs.

Unlike movers vs stayers in the 'Risky' cluster, the number of cigarettes smoked per day was the only significant difference between those who moved from the 'Moderate Smokers' to the 'Mainstream' cluster compared to those who remained in the 'Moderate Smokers' cluster at age 33 (men=16.5 vs 14.6; women 14.0 vs 12.1, p≤0.05). This suggests smoking differentiates individuals at age 33 in the 'Moderate Smokers' whose HRB cluster membership changes in a positive direction from those whose HRBs remain the same.

In contrast, those who moved from the 'Mainstream' to the 'Moderate Smokers' cluster differed not only in relation to smoking at age 33 but also in terms of their frequency of fried food consumption (men=3.5 vs 3.7; women=2.5 vs 2.8, p≤0.05) and the proportion consuming alcohol consumption above recommended limits (men=23.2% vs 32.1%, women=6.3% vs 12.9%, p≤0.05). This suggests that differences at age 33 in relation to alcohol consumption and fried food consumption may to some extent predict movement from the 'Mainstream' to the 'Moderate Smokers' cluster.

Table 8.6.1 Estimates comparing those who stayed in the 'Risky' cluster and those that moved.

HRB cluster indicator variables	**Men NCDS Stable in Risky N=590 (85.6%)**		**Men NCDS Risky to Moderate Smokers or Mainstream N=99 (14.4%)**		**Women NCDS Stable in Risky N=560 (86.7%)**		**Women NCDS Risky to Moderate Smokers or Mainstream N=86 (13.3%)**	
	Age 33	Age 42	Age 33	Age 42	Age 33	Age 42	Age 33	Age 42
Number of cigarettes smoked per day	23.6 (10.5)*	24.5 (9.6)*	20.5 (9.0)*	2.1 (5.5)*	21.8 (8.3)	21.8 (7.5)*	22.0 (7.2)	9.4 (8.7)*
Frequency of fruit and vegetable consumption	2.8 (1.9)	2.8 (1.7)*	2.8 (1.8)	4.2 (2.3)*	3.3 (1.8)	3.1 (1.6)*	3.4 (1.8)	5.6 (2.0)*
Frequency of fried food consumption	5.4 (1.4)*	3.6 (1.5)*	5.8 (1.2)*	3.0 (1.2)*	4.2 (1.5)*	3.1 (1.2)*	4.6 (1.4)*	2.2 (1.1)*
Frequency of sweet food consumption	3.4 (2.3)	3.3 (2.2)*	3.5 (2.3)	3.8 (2.3)*	4.1 (2.5)	4.0 (2.5)	4.1 (2.3)	3.9 (2.1)
Smoking status								
Smoker	504 (99.4%)	449 (100%)*	99 (100%)	82 (82.2%)*	493 (99.6%)	447 (100%)*	86 (100%)	53 (61.6%)*
Non-smoker	3 (0.6%)	0	0	17 (17.2%)	2 (0.4%)	0	0	33 (38.4%)
Frequency of leisure-time physical activity								
≤3 times a month	257 (50.8%)	261 (58.0%)*	51 (51.5%)	33 (33.3%)*	292 (59.2%)	328 (73.4%)*	58 (67.2%)	34 (39.5%)*
Once a week	96 (19.0%)	63 (14.0%)	18 (18.2%)	25 (25.3%)	84 (17.0%)	34 (7.6%)	14 (16.3%)	12 (14.0%)
2–3 days a week	17 (14.0%)	46 (10.2%)	16 (16.2%)	19 (19.2%)	42 (8.5%)	21 (4.7%)	2 (2.3%)	12 (14.0%)
4–7 days a week	82 (16.2%)	80 (17.8%)	14 (14.1%)	22 (22.2%)	75 (15.2%)	64 (14.3%)	12 (14.0%)	28 (32.6%)
Alcohol units consumed in the previous week								
No units	75 (14.8%)*	68 (15.2%)*	9 (9.1%)*	16 (16.2%)*	141 (28.5%)	137 (30.8%)	32 (37.2%)	37 (43.0%)
Within limits (≤14 units women, ≤21 units men)	179 (35.3%)	123 (27.5%)	51 (51.6%)	40 (40.4%)	264 (53.3%)	166 (37.3%)	44 (51.2%)	30 (34.9%)
Above limits (≥15 units women, ≥22 units men)	253 (50.0%)	257 (57.4%)	39 (39.4%)	43 (43.4%)	90 (18.2%)	142 (31.9%)	10 (11.6%)	19 (22.1%)

Note: Participants assigned to their most likely HRB cluster, mean (sd) or n (%), *p≤0.05 (test for difference in HRBs across the three transition groups at each age using t-test or chi-squared).

Table 8.6.2 Estimates comparing those who stayed in the 'Moderate Smokers' cluster and those that moved.

HRB cluster indicator variables	**Men NCDS Stable in Moderate Smokers N=1,125 (82.4%)**		**Men NCDS Moderate Smokers to Mainstream N=240 (17.6%)**		**Women NCDS Stable in Moderate Smokers N=1,116 (81.2%)**		**Women NCDS Moderate Smokers to Mainstream N=259 (18.8%)**	
	Age 33	Age 42	Age 33	Age 42	Age 33	Age 42	Age 33	Age 42
Number of cigarettes smoked per day	16.5 (7.9)*	16.6 (7.3)*	14.6 (8.8)*	0*	14.0 (6.7)*	14.2 (6.2)*	12.1 (7.0)*	0*
Frequency of fruit and vegetable consumption	4.4 (1.9)	4.8 (2.0)*	4.5 (2.0)	5.4 (2.1)*	5.7 (1.8)	5.8 (1.9)*	5.5 (1.9)	6.4 (2.2)*
Frequency of fried food consumption	3.3 (1.2)	2.7 (1.0)	3.3 (1.2)	2.7 (1.0)	2.5 (1.3)	2.1 (0.9)	2.6 (1.3)	2.2 (0.9)
Frequency of sweet food consumption	4.5 (2.4)	4.5 (2.4)	4.6 (2.1)	4.3 (2.1)	4.3 (2.4)	4.0 (2.4)	4.3 (2.4)	4.3 (2.3)
Smoking status								
Smoker	921 (99.9%)	913 (100%)*	240 (100%)	4 (1.7%)*	969 (100%)	960 (100%)*	259 (100%)	0*
Non-smoker	1 (0.1%)	0	0	236 (98.3%)	0	0	0	259 (100%)
Frequency of leisure-time physical activity								
≤3 times a month	327 (35.5%)	378 (41.4%)	80 (33.5%)	83 (34.6%)	264 (27.4%)	334 (34.8%)	85 (33.1%)	91 (35.1%)
Once a week	195 (21.2%)	156 (17.1%)	48 (20.1%)	45 (18.8%)	200 (20.7%)	143 (14.9%)	56 (21.8%)	43 (16.6%)
2–3 days a week	189 (20.5%)	166 (18.2%)	54 (22.6%)	42 (17.5%)	189 (20.6%)	189 (19.7%)	48 (18.7%)	45 (17.4%)
4–7 days a week	209 (22.7%)	213 (23.3%)	57 (23.9%)	70 (29.2%)	312 (32.3%)	294 (30.6%)	68 (26.5%)	80 (30.9%)
Alcohol units consumed in the previous week								
No units	124 (13.5%)	169 (18.6%)*	42 (17.5%)	34 (14.2%)*	281 (29.0%)	260 (27.2%)	69 (26.6%)	62 (23.9%)
Within limits (≤14 units women, ≤21 units men)	507 (55.0%)	372 (40.9%)	128 (53.3%)	120 (50.0%)	592 (61.1%)	506 (52.9%)	169 (65.3%)	153 (59.1%)
Above limits (≥15 units women, ≥22 units men)	291 (31.6%)	368 (40.5%)	70 (29.2%)	86 (35.8%)	96 (9.9%)	191 (20.0%)	21 (8.1%)	44 (17.0%)

Note: Participants assigned to their most likely HRB cluster, mean (sd) or n (%), *p≤0.05 (test for difference in HRBs across the three transition groups at each age using t-test or chi-squared).

Table 8.6.3 Estimates comparing those who stayed in the 'Mainstream' cluster and those that moved.

HRB cluster indicator variables	**Men NCDS Stable in Mainstream N=4,205 (96.8%)**		**Men NCDS Mainstream to Moderate Smokers N=137 (3.2%)**		**Women NCDS Stable in Mainstream N=4,235 (97.0%)**		**Women NCDS Mainstream to Moderate Smokers N=132 (3.0%)**	
	Age 33	Age 42	Age 33	Age 42	Age 33	Age 42	Age 33	Age 42
Number of cigarettes smoked per day	0	0*	0	14.3 (8.4)*	0	0*	0	11.4 (6.5)*
Frequency of fruit and vegetable consumption	4.7 (1.9)*	5.3 (2.0)*	4.3 (2.0)*	4.4 (2.1)*	5.8 (1.8)	6.4 (2.1)*	5.8 (1.9)	5.4 (2.1)*
Frequency of fried food consumption	3.4 (1.5)*	2.7 (1.0)	3.7 (1.4)*	2.7 (1.1)	2.5 (1.4)*	2.3 (0.9)	2.8 (1.4)*	2.4 (1.0)
Frequency of sweet food consumption	4.7 (2.2)	4.5 (2.2)	4.7 (2.3)	4.3 (2.5)	4.9 (2.3)	4.5 (2.2)	4.6 (2.3)	4.3 (2.2)
Smoking status								
Smoker	89 (2.4%)*	153 (4.2%)*	7 (5.1%)*	137 (100%)*	82 (2.1%)*	116 (3.0%)*	10 (7.6%)*	132 (100%)*
Non-smoker	3,579 (97.6%)	3,537 (95.8%)	130 (94.9%)	0	3,757 (97.9%)	3,683 (97.0%)	122 (92.4%)	0
Frequency of leisure-time physical activity								
≤3 times a month	1,018 (27.8%)	1,083 (29.4%)*	40 (29.4%)	57 (41.6%)*	1,037 (27.2%)	1,145 (30.2%)*	39 (29.6%)	50 (37.9%)*
Once a week	774 (21.1%)	765 (20.7%)	35 (25.7%)	26 (19.0%)	936 (24.5%)	692 (18.2%)	24 (18.2%)	18 (13.6%)
2–3 days a week	933 (25.5%)	887 (24.1%)	29 (21.3%)	33 (24.1%)	805 (21.1%)	896 (23.6%)	24 (18.2%)	21 (15.9%)
4–7 days a week	936 (25.6%)	953 (25.8%)	32 (23.5%)	21 (15.3%)	1,039 (27.2%)	1,064 28.0%)	45 (34.1%)	43 (32.6%)
Alcohol units consumed in the previous week								
No units	487 (13.2%)*	412 (11.2%)	17 (12.4%)*	13 (9.5%)	1,099 (28.6%)*	821 (21.6%)*	48 (36.4%)*	36 (27.5%)*
Within limits (≤14 units women, ≤21 units men)	2,339 (63.6%)	2,023 (54.9%)	76 (55.5%)	68 (49.6%)	2,504 (65.2%)	2,401 (63.2%)	67 (50.8%)	60 (45.8%)
Above limits (≥15 units women, ≥22 units men)	852 (23.2%)	1,250 (33.9%)	44 (32.1%)	56 (40.9%)	240 (6.3%)	576 (15.2%)	17 (12.9%)	35 (26.7%)

Note: Participants assigned to their most likely HRB cluster, mean (sd) or n (%), *p≤0.05 (test for difference in HRBs across the two transition groups at each age using t-test or chi-squared).

Thesis publication

MAWDITT, C., SACKER, A., BRITTON, A., KELLY, Y. & CABLE, N. 2016. The clustering of health-related behaviours in a British population sample: Testing for cohort differences. *Preventive Medicine,* 88**,** 95-107*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2016.03.003.

References

ABEL, T. 2008. Cultural capital and social inequality in health. *Journal of Epidemiology and Community Health,* 62**,** e13*,* DOI: 10.1136/jech.2007.066159.

ABEL, T. & FROHLICH, K. L. 2012. Capitals and capabilities: Linking structure and agency to reduce health inequalities. *Social Science & Medicine,* 74**,** 236-244*,* DOI: http://dx.doi.org/10.1016/j.socscimed.2011.10.028.

ABEL, T., FUHR, D. C., BISEGGER, C., RAU, S. A. & EUROPEAN KIDSCREEN, G. 2011. Money is not enough: Exploring the impact of social and cultural resources on youth health. *Scandinavian Journal of Public Health,* 39**,** 57-61*,* DOI: 10.1177/1403494810378924.

ALAGEEL, S., WRIGHT, A. J. & GULLIFORD, M. C. 2016. Changes in cardiovascular disease risk and behavioural risk factors before the introduction of a health check programme in England. *Preventive Medicine,* 91**,** 158-163*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2016.08.025.

ALLEN, A. M., SCHEUERMANN, T. S., NOLLEN, N., HATSUKAMI, D. & AHLUWALIA, J. S. 2016. Gender Differences in Smoking Behavior and Dependence Motives Among Daily and Nondaily Smokers. *Nicotine & Tobacco Research,* 18**,** 1408-1413*,* DOI: 10.1093/ntr/ntv138.

ALLEN, K., PEARSON-STUTTARD, J., HOOTON, W., DIGGLE, P., CAPEWELL, S. & O'FLAHERTY, M. 2015. Potential of trans fats policies to reduce socioeconomic inequalities in mortality from coronary heart disease in England: cost effectiveness modelling study.

ALLENDER, S., FOSTER, C. & BOXER, A. 2008a. Occupational and non-occupational physical activity and the social determinants of physical activity: results from the Health Survey for England. *Journal of physical activity & health,* 5**,** 104-116.

ALLENDER, S., HUTCHINSON, L. & FOSTER, C. 2008b. Life-change events and participation in physical activity: a systematic review. *Health promotion international,* 23**,** 160-172.

ALLENDER, S., PETO, V., SCARBOROUGH, P., BOXER, A. & RAYNER, M. 2006. Diet, Physical Activity and Obesity Statistics. Oxford: University of Oxford, http://dro.deakin.edu.au/eserv/DU:30020497/allender-dietphysicalactivity-2006.pdf.

AMATO, K., PARK, E. & NIGG, C. R. 2016. Prioritizing multiple health behavior change research topics: expert opinions in behavior change science. *Translational Behavioral Medicine,* 6**,** 220-227*,* DOI: 10.1007/s13142-015-0381-5.

AMATO, P. R. 1994. Life-span adjustment of children to their parents' divorce. *The future of children***,** 143-164.

AN, R. & XIANG, X. 2016. Age-period-cohort analyses of obesity prevalence in US adults. *Public Health,* 141**,** 163-169*,* DOI: 10.1016/j.puhe.2016.09.021.

AN, R., XIANG, X., YANG, Y. & YAN, H. 2016. Mapping the Prevalence of Physical Inactivity in U.S. States, 1984-2015. *PLOS ONE,* 11**,** e0168175*,* DOI: 10.1371/journal.pone.0168175.

ANDERS, J. & DORSETT, R. 2017. What young English people do once they reach school-leaving age: A cross-cohort comparison for the last 30 years. *2017,* 8**,** 29*,* DOI: 10.14301/llcs.v8i1.399.

ANDERSEN, L. L. 2011. Influence of psychosocial work environment on adherence to workplace exercise. *Journal of Occupational and Environmental Medicine,* 53**,** 182-184.

ANDERSSON, M. A. & MARALANI, V. 2015. Early-life characteristics and educational disparities in smoking. *Soc Sci Med,* 144**,** 138-47*,* DOI: 10.1016/j.socscimed.2015.09.027.
ANYAEGBU, G. 2010. Using the OECD equivalence scale in taxes and benefits analysis. *The Labour gazette,* 4**,** 49-54.
ARNO, A. & THOMAS, S. 2016. The efficacy of nudge theory strategies in influencing adult dietary behaviour: a systematic review and meta-analysis. *BMC Public Health,* 16**,** 1-11*,* DOI: 10.1186/s12889-016-3272-x.
ARTAUD, F., SABIA, S., DUGRAVOT, A., KIVIMAKI, M., SINGH-MANOUX, A. & ELBAZ, A. 2016. Trajectories of Unhealthy Behaviors in Midlife and Risk of Disability at Older Ages in the Whitehall II Cohort Study. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences,* 71**,** 1500-1506*,* DOI: 10.1093/gerona/glw060.
ASHRA, N. B., SPONG, R., CARTER, P., DAVIES, M. J., DUNKLEY, A., GILLIES, C., GREAVES, C., KHUNTI, K., SUTTON, S., YATES, T., YOUSSEF, D. & GRAY, L. J. 2015. A systematic review and meta-analysis assessing the effectiveness of pragmatic lifestyle interventions for the prevention of type 2 diabetes mellitus in routine practice. London: Public Health England, https://www.gov.uk/government/publications/diabetes-prevention-programmes-evidence-review.
ASPAROUHOV, T. & MUTHÉN, B. 2014. Auxiliary Variables in Mixture Modeling: Three-Step Approaches Using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal,* 21**,** 329-341*,* DOI: 10.1080/10705511.2014.915181.
ATHERTON, K., FULLER, E., SHEPHERD, P., STRACHAN, D. & POWER, C. 2008. Loss and representativeness in a biomedical survey at age 45 years: 1958 British birth cohort. *Journal of epidemiology and community health,* 62**,** 216-223.
BACKETT, K. C. & DAVISON, C. 1995. Lifecourse and lifestyle: The social and cultural location of health behaviours. *Social Science & Medicine,* 40**,** 629-638*,* DOI: http://dx.doi.org/10.1016/0277-9536(95)80007-7.
BADLEY, E. M., CANIZARES, M., PERRUCCIO, A. V., HOGG-JOHNSON, S. & GIGNAC, M. A. 2015. Benefits Gained, Benefits Lost: Comparing Baby Boomers to Other Generations in a Longitudinal Cohort Study of Self-Rated Health. *Milbank Quarterly,* 93**,** 40-72.
BAMBRA, C., HILLIER, F., CAIRNS, J. M., KASIM, A., MOORE, H. & SUMMERBELL, C. 2015. How effective are interventions at reducing socioeconomic inequalities in obesity among children and adults? Two systematic reviews. *Public Health Res,* 3*,* DOI: 10.3310/phr03010.
BANDURA, A. 1991. Social cognitive theory of self-regulation. *Organizational behavior and human decision processes,* 50**,** 248-287.
BANN, D., HAMER, M., PARSONS, S., PLOUBIDIS, G. B. & SULLIVAN, A. 2016. Does an elite education benefit health? Findings from the 1970 British Cohort Study. *International Journal of Epidemiology,* DOI: 10.1093/ije/dyw045.
BANN, D., JOHNSON, W., LI, L., KUH, D. & HARDY, R. 2017. Socioeconomic Inequalities in Body Mass Index across Adulthood: Coordinated Analyses of Individual Participant Data from Three British Birth Cohort Studies Initiated in 1946, 1958 and 1970. *PLOS Medicine,* 14**,** e1002214*,* DOI: 10.1371/journal.pmed.1002214.
BARTLEY, M. 2004. Models of Aetiological Pathways, IV: The Life-Course Approach. *Health Inequality: An introduction to theories, concepts and methods.* Cambridge: Polity Press.
BARTLEY, M. 2016a. Explanatory Models I: Behavioural and 'Cultural' Explanations. *In:* BARTLEY, M. (ed.) *Health inequality: an introduction to concepts, theories and methods.* 2nd ed. Cambridge: Polity Press.

BARTLEY, M. 2016b. Gender and Inequality in Health. *In:* BARTLEY, M. (ed.) *Health inequality: an introduction to concepts, theories and methods.* Cambridge: Polity Press.

BARTLEY, M., SACKER, A., FIRTH, D. & FITZPATRICK, R. 1999. Understanding social variation in cardiovascular risk factors in women and men: the advantage of theoretically based measures. *Social science & medicine,* 49**,** 831-845.

BATEMAN-HOUSE, A., BAYER, R., COLGROVE, J., FAIRCHILD, A. L. & MCMAHON, C. E. 2017. Free to Consume? Anti-Paternalism and the Politics of New York City's Soda Cap Saga. *Public Health Ethics,* DOI: 10.1093/phe/phw046.

BATIS, C., RIVERA, J. A., POPKIN, B. M. & TAILLIE, L. S. 2016. First-Year Evaluation of Mexico's Tax on Nonessential Energy-Dense Foods: An Observational Study. *PLOS Medicine,* 13**,** e1002057*,* DOI: 10.1371/journal.pmed.1002057.

BAULD, L. 2011. The impact of smokefree legislation in England: evidence review. *Department of Health*.

BEAUCHAMP, A., PEETERS, A., WOLFE, R., TURRELL, G., HARRISS, L. R., GILES, G. G., ENGLISH, D. R., MCNEIL, J., MAGLIANO, D., HARRAP, S., LIEW, D., HUNT, D. & TONKIN, A. 2010. Inequalities in cardiovascular disease mortality: the role of behavioural, physiological and social risk factors. *Journal of Epidemiology and Community Health,* 64**,** 542-548*,* DOI: 10.1136/jech.2009.094516.

BECKER, S. & ZIMMERMANN-STENZEL, M. 2008. Physical activity, obesity, and educational attainment in 50- to 70-year-old adults. *Journal of Public Health,* 17**,** 145*,* DOI: 10.1007/s10389-008-0222-9.

BÉCUE-BERTAUT, M., KERN, J., HERNÁNDEZ-MALDONADO, M.-L., JURESA, V. & VULETIC, S. 2008. Health-risk behaviour in Croatia. *Public health,* 122**,** 140-150.

BEENACKERS, M. A., KAMPHUIS, C., GISKES, K., BRUG, J., KUNST, A. E., BURDORF, A. & VAN LENTHE, F. J. 2012. Socioeconomic inequalities in occupational, leisure-time, and transport related physical activity among European adults: A systematic review. *Int J Behav Nutr Phys Act,* 9**,** 116.

BELLIS, M. A., HUGHES, K., NICHOLLS, J., SHERON, N., GILMORE, I. & JONES, L. 2016. The alcohol harm paradox: using a national survey to explore how alcohol may disproportionately impact health in deprived individuals. *BMC Public Health,* 16**,** 1-10*,* DOI: 10.1186/s12889-016-2766-x.

BEN-SHLOMO, Y. & KUH, D. 2002. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *International Journal of Epidemiology,* 31**,** 285-293*,* DOI: 10.1093/ije/31.2.285.

BENACH, J., MALMUSI, D., YASUI, Y. & MARTINEZ, J. M. 2013. A new typology of policies to tackle health inequalities and scenarios of impact based on Rose's population approach. *J Epidemiol Community Health,* 67**,** 286-91*,* DOI: 10.1136/jech-2011-200363.

BENTLER, P. M. 1990. Comparative fit indexes in structural models. *Psychological bulletin,* 107**,** 238.

BENZIES, K. M., WÅNGBY, M. & BERGMAN, L. R. 2008. Stability and change in health-related behaviors of midlife Swedish women. *Health Care for Women International,* 29**,** 997-1018.

BERRIGAN, D., DODD, K., TROIANO, R. P., KREBS-SMITH, S. M. & BARBASH, R. B. 2003. Patterns of health behavior in U.S. adults. *Preventive Medicine,* 36**,** 615-623*,* DOI: http://dx.doi.org/10.1016/S0091-7435(02)00067-1.

BERSTAD, P., BOTTERI, E., LARSEN, I. K., LØBERG, M., KALAGER, M., HOLME, Ø., BRETTHAUER, M. & HOFF, G. 2016. Lifestyle changes at middle age and mortality: a population-based prospective cohort study. *Journal of Epidemiology and Community Health,* DOI: 10.1136/jech-2015-206760.

BHIMJIYANI, A., KNUCHEL-TAKANO, A., SELVARAJAH, C., WEBBER, L., JACCARD, A., BROWN, M., LANDON, J., RETAT, L. & HUNT, D. 2016. Short and sweet: Why the government should introduce a sugary drinks tax. London: Cancer Research UK, http://www.cancerresearchuk.org/sites/default/files/short_and_sweet_exec_sum_live.pdf.

BIEN, T. H. & BURGE, R. 1990. Smoking and drinking: a review of the literature. *Substance Use & Misuse,* 25**,** 1429-1454.

BIRCH, S. 2010. I dreamed a dream: England reduces health inequalities and wins the world cup. *Health economics,* 19**,** 881-885.

BLAKELY, T., COBIAC, L. J., CLEGHORN, C. L., PEARSON, A. L., VAN DER DEEN, F. S., KVIZHINADZE, G., NGHIEM, N., MCLEOD, M. & WILSON, N. 2015. Health, Health Inequality, and Cost Impacts of Annual Increases in Tobacco Tax: Multistate Life Table Modeling in New Zealand. *PLOS Medicine,* 12**,** e1001856*,* DOI: 10.1371/journal.pmed.1001856.

BLAXTER, M. 1990. Behaviour and Health. *Health and Lifestyles.* London: Routledge.

BLUE, S., SHOVE, E., CARMONA, C. & KELLY, M. P. 2014. Theories of practice and public health: understanding (un) healthy practices. *Critical Public Health***,** 1-15.

BONDY, S. & REHM, J. 1998. The interplay of drinking patterns and other determinants of health. *Drug and alcohol review,* 17**,** 399-412.

BORLAND, R. 2013a. Characteristics of hard-to-maintain behaviours. *Understanding Hard to Maintain Behaviour Change.* John Wiley & Sons, Ltd. DOI: 10.1002/9781118572894.ch2.

BORLAND, R. 2013b. Conceptual influences on change. *Understanding Hard to Maintain Behaviour Change.* John Wiley & Sons, Ltd. DOI: 10.1002/9781118572894.ch5.

BORLAND, R. 2013c. Environmental influences: the context of change. *Understanding Hard to Maintain Behaviour Change.* John Wiley & Sons, Ltd. DOI: 10.1002/9781118572894.ch4.

BORLAND, R. 2013d. Interventions for behaviour change. *Understanding Hard to Maintain Behaviour Change.* John Wiley & Sons, Ltd. DOI: 10.1002/9781118572894.ch7.

BORLAND, R. 2013e. An overview of the theory. *Understanding Hard to Maintain Behaviour Change.* John Wiley & Sons, Ltd. DOI: 10.1002/9781118572894.ch1.

BORODULIN, K., ZIMMER, C., SIPPOLA, R., MÄKINEN, T. E., LAATIKAINEN, T. & PRÄTTÄLÄ, R. 2012. Health behaviours as mediating pathways between socioeconomic position and body mass index. *International journal of behavioral medicine,* 19**,** 14-22.

BOSQUE-PROUS, M., ESPELT, A., BORRELL, C., BARTROLI, M., GUITART, A. M., VILLALBÍ, J. R. & BRUGAL, M. T. 2015. *Gender differences in hazardous drinking among middle-aged in Europe: the role of social context and women's empowerment,* DOI: 10.1093/eurpub/cku234.

BREEN, R. & GOLDTHORPE, J. H. 2001. Class, Mobility and Merit The Experience of Two British Birth Cohorts. *European Sociological Review,* 17**,** 81-101*,* DOI: 10.1093/esr/17.2.81.

BRIGGS, A. D. M., MYTTON, O. T., KEHLBACHER, A., TIFFIN, R., ELHUSSEIN, A., RAYNER, M., JEBB, S. A., BLAKELY, T. & SCARBOROUGH, P. 2017. Health impact assessment of the UK soft drinks industry levy: a comparative risk assessment modelling study. *The Lancet Public Health,* 2**,** e15-e22*,* DOI: http://dx.doi.org/10.1016/S2468-2667(16)30037-8.

BRITTON, A. & BELL, S. 2015. Reasons Why People Change Their Alcohol Consumption in Later Life: Findings from the Whitehall II Cohort Study. *PLoS ONE,* 10**,** e0119421*,* DOI: 10.1371/journal.pone.0119421.

BRITTON, A., BEN-SHLOMO, Y., BENZEVAL, M., KUH, D. & BELL, S. 2015. Life course trajectories of alcohol consumption in the United Kingdom using longitudinal data from nine cohort studies. *BMC medicine,* 13**,** 47.

BRITTON, A., MARMOT, M. G. & SHIPLEY, M. J. 2010. How does variability in alcohol consumption over time affect the relationship with mortality and coronary heart disease? *Addiction,* 105**,** 639-645*,* DOI: 10.1111/j.1360-0443.2009.02832.x.

BROWN, J., WEST, R., BEARD, E., BRENNAN, A., DRUMMOND, C., GILLESPIE, D., HICKMAN, M., HOLMES, J., KANER, E. & MICHIE, S. 2016. Are recent attempts to quit smoking associated with reduced drinking in England? A cross-sectional population survey. *BMC Public Health,* 16**,** 1-7*,* DOI: 10.1186/s12889-016-3223-6.

BROWN, K. 2015. The Public Health Responsibility Deal: why alcohol industry partnerships are bad for health? *Addiction,* 110**,** 1227-1228*,* DOI: 10.1111/add.12974.

BROWN, R. & OGDEN, J. 2004. Children's eating attitudes and behaviour: a study of the modelling and control theories of parental influence. *Health Education Research,* 19**,** 261-271*,* DOI: 10.1093/her/cyg040.

BRUHN, J. G. & REBACH, H. M. 2014. Socioeconomic Status and Caregiving. *The Sociology of Caregiving.* Dordrecht: Springer Netherlands. DOI: 10.1007/978-94-017-8857-1_7.

BRUNNER, E. & MARMOT, M. 2006. Social organization, stress and health. *In:* MARMOT, M. & WILKINSON, R. (eds.) *Social determinants of health.* 2nd ed. Oxford: Oxford University Press.

BRUNNER, E. J., CHANDOLA, T. & MARMOT, M. G. 2007. Prospective Effect of Job Strain on General and Central Obesity in the Whitehall II Study. *American Journal of Epidemiology,* 165**,** 828-837*,* DOI: 10.1093/aje/kwk058.

BUCK, D. & FROSINI, F. 2012. Clustering of unhealthy behaviours over time: Implications for policy and practice.

BUKODI, E. & GOLDTHORPE, J. 2011. Social class returns to higher education: chances of access to the professional and managerial salariat for men in three British birth cohorts. *Longitudinal and Life Course Studies,* 2**,** 185-201.

BUKODI, E. & GOLDTHORPE, J. H. 2012. Decomposing 'Social Origins': The Effects of Parents' Class, Status, and Education on the Educational Attainment of Their Children. *European sociological review***,** jcs079.

BULL, E. R., DOMBROWSKI, S. U., MCCLEARY, N. & JOHNSTON, M. 2014. Are interventions for low-income groups effective in changing healthy eating, physical activity and smoking behaviours? A systematic review and meta-analysis. *BMJ Open,* 4*,* DOI: 10.1136/bmjopen-2014-006046.

BURTON-JEANGROS, C., CULLATI, S., SACKER, A. & BLANE, D. 2015. A life course perspective on health trajectories and transitions. Springer.

BURTON, N. W., KHAN, A. & BROWN, W. J. 2012. How, where and with whom? Physical activity context preferences of three adult groups at risk of inactivity. *British Journal of Sports Medicine,* 46**,** 1125-1131*,* DOI: 10.1136/bjsports-2011-090554.

BURTON, R., HENN, C., LAVOIE, D., WOLFF, A., MARSDEN, J. & SHERON, N. 2016. The Public Health Burden of Alcohol and the Effectiveness and Cost-Effectiveness of Alcohol Control Policies An evidence review. London: Public Health England, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/574427/Alcohol_public_health_burden_evidence_review.pdf.

BUTLER, N., DOWLING, S., OSBORN, A., 2016. 1970 British Cohort Study: Five-Year Follow-Up, 1975. 5th Edition ed. London: UK Data Service, DOI: http://doi.org/10.5255/UKDA-SN-2699-4.

BYRNE, D. W., ROLANDO, L. A., ALIYU, M. H., MCGOWN, P. W., CONNOR, L. R., AWALT, B. M., HOLMES, M. C., WANG, L. & YARBROUGH, M. I. 2016. Modifiable Healthy Lifestyle

Behaviors: 10-Year Health Outcomes From a Health Promotion Program. *American Journal of Preventive Medicine,* 51**,** 1027-1037*,* DOI: 10.1016/j.amepre.2016.09.012.
CABLE, N. 2014. Life Course Approach in Social Epidemiology: An Overview, Application and Future Implications. *Journal of Epidemiology,* 24**,** 347.
CAMPBELL, D. 2016. Obesity is the new smoking. So let's treat it as such. *The Guardian*. http://www.theguardian.com/society/2016/mar/22/obesity-smoking-action-sugar-tax?utm_source=The%20King%27s%20Fund%20newsletters&utm_medium=email&utm_campaign=6884250_HWBB%202016-03-28&dm_i=21A8,43JX6,FLWP3U,EWPAS,1.
CAMPOSTRINI, S. & MCQUEEN, D. 2003. Creating a Synthetic Behavioural Risk Factor Index to Assess Trends in Surveillance Data. *In:* MCQUEEN, D. & PUSKA, P. (eds.) *Global Behavioral Risk Factor Surveillance.* Springer US. DOI: 10.1007/978-1-4615-0071-1_15.
CAPEWELL, S. & LILFORD, R. 2016. Are nanny states healthier states? *BMJ,* 355*,* DOI: 10.1136/bmj.i6341.
CARVALHO, L. 2012. Childhood Circumstances and the Intergenerational Transmission of Socioeconomic Status. *Demography,* 49**,** 913-938*,* DOI: 10.1007/s13524-012-0120-1.
CATFORD, J. 2006. Creating political will: moving from the science to the art of health promotion. *Health Promotion International,* 21**,** 1-4*,* DOI: 10.1093/heapro/dak004.
CHANG, H.-C., YANG, H.-C., CHANG, H.-Y., YEH, C.-J., CHEN, H.-H., HUANG, K.-C. & PAN, W.-H. 2017. Morbid obesity in Taiwan: Prevalence, trends, associated social demographics, and lifestyle factors. *PLOS ONE,* 12**,** e0169577*,* DOI: 10.1371/journal.pone.0169577.
CHASTIN, S. F. M., PALAREA-ALBALADEJO, J., DONTJE, M. L. & SKELTON, D. A. 2015. Combined Effects of Time Spent in Physical Activity, Sedentary Behaviors and Sleep on Obesity and Cardio-Metabolic Health Markers: A Novel Compositional Data Analysis Approach. *PLoS ONE,* 10**,** e0139984*,* DOI: 10.1371/journal.pone.0139984.
CHILDS, K. K. & SULLIVAN, C. J. 2013. Investigating the underlying structure and stability of problem behaviors across adolescence. *Criminal Justice and Behavior,* 40**,** 57-79*,* DOI: 10.1177/0093854812460496.
CHINN, D. J., WHITE, M., HARLAND, J., DRINKWATER, C. & RAYBOULD, S. 1999. Barriers to physical activity and socioeconomic position: implications for health promotion. *Journal of Epidemiology and Community Health,* 53**,** 191-192*,* DOI: 10.1136/jech.53.3.191.
CHIOLERO, A., WIETLISBACH, V., RUFFIEUX, C., PACCAUD, F. & CORNUZ, J. 2006. Clustering of risk behaviors with cigarette consumption: A population-based survey. *Preventive Medicine,* 42**,** 348-353*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2006.01.011.
CHOU, S.-Y., GROSSMAN, M. & SAFFER, H. 2004. An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. *Journal of health economics,* 23**,** 565-587.
CHOW, C. K., JOLLY, S., RAO-MELACINI, P., FOX, K. A., ANAND, S. S. & YUSUF, S. 2010. Association of diet, exercise, and smoking modification with risk of early cardiovascular events after acute coronary syndromes. *Circulation,* 121**,** 750-758.
CHUNG, H., PARK, Y. & LANZA, S. T. 2005. Latent transition analysis with covariates: pubertal timing and substance use behaviours in adolescent females. *Statistics in Medicine,* 24**,** 2895-2910, http://onlinelibrary.wiley.com/store/10.1002/sim.2148/asset/2148_ftp.pdf?v=1&t=i3vem83z&s=e3676c9c0b0d7c40b3ac4c2e68845e7a9a521838.
CLARK, A. J., SALO, P., LANGE, T., JENNUM, P., VIRTANEN, M., PENTTI, J., KIVIMÄKI, M., VAHTERA, J. & ROD, N. H. 2015. Onset of impaired sleep as a predictor of change in health-related behaviours; analysing observational data as a series of non-randomized pseudo-trials. *International Journal of Epidemiology,* 44**,** 1027-1037*,* DOI: 10.1093/ije/dyv063.

CLARK, S. L. & MUTHÉN, B. 2009. Relating latent class analysis results to variables not included in the analysis. *Submitted for publication*.

CLARK, S. L., MUTHÉN, B., KAPRIO, J., D'ONOFRIO, B. M., VIKEN, R. & ROSE, R. J. 2013. Models and Strategies for Factor Mixture Analysis: An Example Concerning the Structure Underlying Psychological Disorders. *Structural equation modeling : a multidisciplinary journal,* 20**,** 10.1080/10705511.2013.824786*,* DOI: 10.1080/10705511.2013.824786.

CLELAND, V., BALL, K. & CRAWFORD, D. 2010. Social and Environmental Determinants of Health Behaviors *In:* STEPTOE, A., FREEDLAND, K. E., JENNINGS, J. R., LLABRE, M. M., MANUCK, S. B., SUSMAN, E. J. & POOLE, L. (eds.) *Handbook of behavioral medicine.* New York: Springer Science & Business Media.

CLOUSTON, S. A. P., RICHARDS, M., CADAR, D. & HOFER, S. M. 2015. Educational Inequalities in Health Behaviors at Midlife: Is There a Role for Early-life Cognition? *Journal of Health and Social Behavior,* 56**,** 323-340*,* DOI: 10.1177/0022146515594188.

CLS 2004. Data Note 2004/5: Investigation into the possible over-reporting of beer consumption at the 1999-2000 NCDS/BCS70 survey. London: Centre for Longitudinal Studies, http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=771&sitesectiontitle=Data+Notes.

CLS 2008a. National Child Development Study: Sweep 5, 1991 *In:* CITY UNIVERSITY. SOCIAL STATISTICS RESEARCH, U. (ed.) 2nd Edition. ed. London: UK Data Service., DOI: http://doi.org/10.5255/UKDA-SN-5567-1.

CLS 2008b. National Child Development Study: Sweep 6, 1999-2000 *In:* JOINT CENTRE FOR LONGITUDINAL, R. (ed.) 2nd Edition. ed. London: UK Data Service. , DOI: http://doi.org/10.5255/UKDA-SN-5578-1.

CLS 2014. National Child Development Study: Childhood Data, Sweeps 0-3, 1958-1974. *In:* NATIONAL BIRTHDAY TRUST FUND, N. C. S. B. (ed.) 3rd Edition. ed. London: UK Data Service, DOI: http://doi.org/10.5255/UKDA-SN-5565-2.

CLS 2016a. 1970 British Cohort Study: Ten-Year Follow-Up, 1980 6th Edition ed. London: UK Data Service, DOI: http://doi.org/10.5255/UKDA-SN-3723-7.

CLS 2016b. 1970 British Cohort Study: Thirty-Four-Year Follow-Up, 2004-2005 4th Edition ed. London: UK Data Service, DOI: http://doi.org/10.5255/UKDA-SN-5585-3.

CLS 2016c. 1970 British Cohort Study: Twenty-Nine-Year Follow-Up, 1999-2000 *In:* JOINT CENTRE FOR LONGITUDINAL, R. (ed.) 4th Edition ed. London: UK Data Service, DOI: http://doi.org/10.5255/UKDA-SN-5558-3.

COBIAC, L. J., TAM, K., VEERMAN, L. & BLAKELY, T. 2017. Taxes and Subsidies for Improving Diet and Population Health in Australia: A Cost-Effectiveness Modelling Study. *PLOS Medicine,* 14**,** e1002232*,* DOI: 10.1371/journal.pmed.1002232.

COCKERHAM, W. C. 2005. Health lifestyle theory and the convergence of agency and structure. *Journal of health and social behavior,* 46**,** 51-67.

COHEN, J. 1992. A power primer. *Psychological bulletin,* 112**,** 155*,* DOI: http://dx.doi.org/10.1037/0033-2909.112.1.155.

COHEN, S., JANICKI-DEVERTS, D., CHEN, E. & MATTHEWS, K. A. 2010. Childhood socioeconomic status and adult health. *Annals of the New York Academy of Sciences,* 1186**,** 37-55.

COLDITZ, G. A., GIOVANNUCCI, E., RIMM, E. B., STAMPFER, M. J., ROSNER, B., SPEIZER, F. E., GORDIS, E. & WILLETT, W. C. 1991. Alcohol intake in relation to diet and obesity in women and men. *The American journal of clinical nutrition,* 54**,** 49-55, http://www.ncbi.nlm.nih.gov/pubmed/2058587.

COLLINS, L. M., GRAHAM, J. W., LONG, J. D. & HANSEN, W. B. 1994. Crossvalidation of latent class models of early substance use onset. *Multivariate Behavioral Research,* 29**,** 165-183.

COLLINS, L. M. & LANZA, S. T. 2010. *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*, John Wiley & Sons.
COLMAN, I., PLOUBIDIS, G. B., WADSWORTH, M. E. J., JONES, P. B. & CROUDACE, T. J. 2007. A Longitudinal Typology of Symptoms of Depression and Anxiety Over the Life Course. *Biological Psychiatry,* 62**,** 1265-1271*,* DOI: http://dx.doi.org/10.1016/j.biopsych.2007.05.012.
CONROY, D. E., RAM, N., PINCUS, A. L., COFFMAN, D. L., LOREK, A. E., REBAR, A. L. & ROCHE, M. J. 2015. Daily Physical Activity and Alcohol Use Across the Adult Lifespan. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association,* 34**,** 653-660*,* DOI: 10.1037/hea0000157.
CONRY, M. C., MORGAN, K., CURRY, P., MCGEE, H., HARRINGTON, J., WARD, M. & SHELLEY, E. 2011. The clustering of health behaviours in Ireland and their relationship with mental health, self-rated health and quality of life. *BMC public health,* 11**,** 692*,* DOI: http://dx.doi.org/10.1186/1471-2458-11-692.
COOPER, R., MISHRA, G. D. & KUH, D. 2011. Physical Activity Across Adulthood and Physical Performance in Midlife Findings from a British Birth Cohort. *American Journal of Preventive Medicine,* 41**,** 376-384*,* DOI: 10.1016/j.amepre.2011.06.035.
COPPO, A., BALDISSERA, S., MIGLIARDI, A., MINARDI, V., QUARCHIONI, E., FERRANTE, G., MOLIN, A., FAGGIANO, F. & GROUP, P. W. 2017. Quit attempts and smoking cessation in Italian adults (25-64 years): factors associated with attempts and successes. *European journal of public health*.
CORBIERE, M., SHEN, J., ROULEAU, M. & DEWA, C. S. 2009. A systematic review of preventive interventions regarding mental health issues in organizations. *Work,* 33**,** 81-116*,* DOI: 10.3233/wor-2009-0846.
CORDER, K., OGILVIE, D. & VAN SLUIJS, E. M. F. 2009. Invited Commentary: Physical Activity Over the Life Course—Whose Behavior Changes, When, and Why? *American Journal of Epidemiology,* 170**,** 1078-1081*,* DOI: 10.1093/aje/kwp273.
CORNA, L. M. 2013. A life course perspective on socioeconomic inequalities in health: A critical review of conceptual frameworks. *Advances in Life Course Research,* 18**,** 150-159*,* DOI: 10.1016/j.alcr.2013.01.002.
CRAWLEY, H. F. & WHILE, D. 1996. Parental smoking and the nutrient intake and food choice of British teenagers aged 16-17 years. *Journal of Epidemiology and Community Health,* 50**,** 306-312*,* DOI: http://dx.doi.org/10.1136/jech.50.3.306.
CROWNE, D. P. & MARLOWE, D. 1960. A new scale of social desirability independent of psychopathology. *Journal of consulting psychology,* 24**,** 349.
CURTICE, J. 2016. Attitudes to obesity Findings from the 2015 British Social Attitudes survey. London: NatCen, http://www.bsa.natcen.ac.uk/media/39132/attitudes-to-obesity.pdf.
CUTLER, D. M., GLAESER, E. L. & SHAPIRO, J. M. 2003. Why have Americans become more obese? *The Journal of Economic Perspectives,* 17**,** 93-118.
CUTLER, D. M. & LLERAS-MUNEY, A. 2010. Understanding differences in health behaviors by education. *Journal of health economics,* 29**,** 1-28, http://ac.els-cdn.com/S0167629609001143/1-s2.0-S0167629609001143-main.pdf?_tid=0cdbfb02-876b-11e4-ade8-00000aab0f02&acdnat=1418985637_d42c58b0410fc793ee0c962842ffdd95.
DAHLGREN, G. & WHITEHEAD, M. 1991. Policies and strategies to promote social equity in health. *Stockholm: Institute for future studies*.

DALE, C. F., FONTANA, V. C. & MARTINEZ, J. A. 2016. What's your 'vice?': a combined approach to drugs and other addictive substances and activities. *Addiction Research & Theory,* 24**,** 366-374*,* DOI: 10.3109/16066359.2015.1133809.

DARE, S., MACKAY, D. F. & PELL, J. P. 2015. Relationship between Smoking and Obesity: A Cross-Sectional Study of 499,504 Middle-Aged Adults in the UK General Population. *PLoS ONE,* 10**,** e0123579*,* DOI: 10.1371/journal.pone.0123579.

DAVIES, J. B. & SHORROCKS, A. F. 2000. The distribution of wealth. *In:* ATKINSON, A. B., F. (ed.) *Handbook of Income Distribution.* Amsterdam, The Netherlands: Elsevier. DOI: http://dx.doi.org/10.1016/S1574-0056(00)80014-7.

DAVIES, S., BURNS, H., JEWELL, T. & MCBRIDE, M. 2011. Start active, stay active: a report on physical activity from the four home countries. *Chief Medical Officers,* 16306**,** 1-62.

DAVIES, S. C. 2016. Annual Report of the Chief Medical Officer 2015. On the State of the Public's Health Baby Boomers: Fit for the Future. London: Department of Health.

DAY, K., GOUGH, B. & MCFADDEN, M. 2004. "Warning! alcohol can seriously damage your feminine health". *Feminist Media Studies,* 4**,** 165-183*,* DOI: 10.1080/1468077042000251238.

DE LEON, J., RENDON, D. M., BACA-GARCIA, E., AIZPURU, F., GONZALEZ-PINTO, A., ANITUA, C. & DIAZ, F. J. 2007. Association between smoking and alcohol use in the general population: stable and unstable odds ratios across two years in two different countries. *Alcohol Alcohol,* 42**,** 252-7*,* DOI: 10.1093/alcalc/agm029.

DE NARDI, M. 2004. Wealth Inequality and Intergenerational Links. *The Review of Economic Studies,* 71**,** 743-768*,* DOI: 10.1111/j.1467-937X.2004.00302.x.

DE VRIES, H., VAN'T RIET, J., SPIGT, M., METSEMAKERS, J., VAN DEN AKKER, M., VERMUNT, J. K. & KREMERS, S. 2008. Clusters of lifestyle behaviors: results from the Dutch SMILE study. *Preventive medicine,* 46**,** 203-208*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2007.08.005.

DE WINTER, A. F., VISSER, L., VERHULST, F. C., VOLLEBERGH, W. A. M. & REIJNEVELD, S. A. 2016. Longitudinal patterns and predictors of multiple health risk behaviors among adolescents: The TRAILS study. *Preventive Medicine,* 84**,** 76-82*,* DOI: 10.1016/j.ypmed.2015.11.028.

DERUITER, W. K., CAIRNEY, J., LEATHERDALE, S. T. & FAULKNER, G. E. J. 2014. A longitudinal examination of the interrelationship of multiple health behaviors. *American Journal of Preventive Medicine,* 47**,** 283-289, http://www.scopus.com/inward/record.url?eid=2-s2.0-84906320981&partnerID=40&md5=f1a6573375ee2889118d63e279a29348.

DEVAUX, M. & SASSI, F. 2015. Social disparities in hazardous alcohol use: self-report bias may lead to incorrect estimates. *The European Journal of Public Health,* DOI: 10.1093/eurpub/ckv190.

DEVINE, C. M. 2005. A life course perspective: Understanding food choices in time, social location, and history. *Journal of Nutrition Education and Behavior,* 37**,** 121-128*,* DOI: 10.1016/s1499-4046(06)60266-2.

DIEPEVEEN, S., LING, T., SUHRCKE, M., ROLAND, M. & MARTEAU, T. M. 2013. Public acceptability of government intervention to change health-related behaviours: a systematic review and narrative synthesis. *BMC Public Health,* 13.

DING, D., DO, A., SCHMIDT, H. M. & BAUMAN, A. E. 2015. A Widening Gap? Changes in Multiple Lifestyle Risk Behaviours by Socioeconomic Status in New South Wales, Australia, 2002-2012. *Plos One,* 10**,** 13*,* DOI: 10.1371/journal.pone.0135338.

DISTEFANO, C., ZHU, M. & MINDRILA, D. 2009. Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation,* 14**,** 1-11.

DIXON, J. & BANWELL, C. 2009. Theory driven research designs for explaining behavioural health risk transitions: The case of smoking. *Social Science & Medicine,* 68**,** 2206-2214*,* DOI: 10.1016/j.socscimed.2009.03.025.

DOH 1995. Sensible drinking: Report of an inter-departmental working group. London: Department of Health, http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/publicationsandstatistics/publications/publicationspolicyandguidance/dh_4084701.

DOH 2011. Changing Behaviour Improving Outcomes: A New Social Marketing Stratgy for Public Health. London: Department of Health, https://www.gov.uk/government/publications/changing-behaviour-improving-outcomes-a-new-social-marketing-strategy-for-public-health.

DOH 2014. Living Well for Longer: National Support for Local Action to Reduce Premature Avoidable Mortality. London: Department of Health, http://webarchive.nationalarchives.gov.uk/20150403155333/http://livinglonger.dh.gov.uk/living-longer/.

DOH 2016. UK Chief Medical Officers' Alcohol Guidelines Review: Summary of the proposed new guidelines. London: Department of Health, https://www.gov.uk/government/consultations/health-risks-from-alcohol-new-guidelines.

DRAGONE, D., MANARESI, F. & SAVORELLI, L. 2015. Obesity and Smoking: can we Kill Two Birds with one Tax? *Health Economics***,** n/a-n/a*,* DOI: 10.1002/hec.3231.

DREWNOWSKI, A., KURTH, C., HOLDEN-WILTSE, J. & SAARI, J. 1992. Food preferences in human obesity: Carbohydrates versus fats. *Appetite,* 18**,** 207-221*,* DOI: http://dx.doi.org/10.1016/0195-6663(92)90198-F.

DRUCKMAN, A. & JACKSON, T. 2008. Household energy consumption in the UK: A highly geographically and socio-economically disaggregated model. *Energy Policy,* 36**,** 3177-3192*,* DOI: http://dx.doi.org/10.1016/j.enpol.2008.03.021.

DUNCAN, G. J. & MAGNUSON, K. 2012. Socioeconomic status and cognitive functioning: moving from correlation to causation. *Wiley Interdisciplinary Reviews: Cognitive Science,* 3**,** 377-386.

DUNCAN, G. J., ZIOL-GUEST, K. M. & KALIL, A. 2010. Early-childhood poverty and adult attainment, behavior, and health. *Child development,* 81**,** 306-325.

DUTRA, L. M., GLANTZ, S. A., LISHA, N. E. & SONG, A. V. 2017. Beyond experimentation: Five trajectories of cigarette smoking in a longitudinal sample of youth. *PLOS ONE,* 12**,** e0171808*,* DOI: 10.1371/journal.pone.0171808.

DZIAK, J. J., COFFMAN, D. L., LANZA, S. T. & LI, R. 2012. Sensitivity and specificity of information criteria. *The Methodology Center and Department of Statistics, Penn State, The Pennsylvania State University*.

EBRAHIM, S., MONTANER, D. & LAWLOR, D. A. 2004. Clustering of risk factors and social class in childhood and adulthood in British women's heart and health study: cross sectional analysis. *Bmj,* 328**,** 861, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC387475/pdf/bmj32800861.pdf.

EDWARDSON, C. L. & GORELY, T. 2010. Parental influences on different types and intensities of physical activity in youth: A systematic review. *Psychology of Sport and Exercise,* 11**,** 522-535.

EKINSMYTH, C., BYNNER, J., MONTGOMERY, S. & SHEPHERD, P. 1992. An Integrated approach to the design and analysis of the 1970 British Cohort Study (BCS70) and the National Child Development Study (NCDS). London: Centre for Longitudinal Studies, http://www.cls.ioe.ac.uk/library-media%5Cdocuments%5Cicwp1.pdf.

ELHAKEEM, A., COOPER, R., BANN, D. & HARDY, R. 2015. Childhood socioeconomic position and adult leisure-time physical activity: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity,* 12**,** 27*,* DOI: 10.1186/s12966-015-0250-0.

ELIAS, P., MCKNIGHT, A. & KINSHOTT, G. 1999. Redefining Skill: Revision of the Standard Occupational Classification (SOC2000). Skills Task Force Research Paper 19.

ELLIOTT, J., DODGEON, B. & ELLIOTT, J. 2007. A Descriptive Analysis of the Drinking Behaviour of the 1958 Cohort at Age 33 and the 1970 Cohort at Age 34. London: Centre for Longitudinal Studies, http://www.cls.ioe.ac.uk/library-media/documents/CLS_WP_2007_3(1).pdf.

ELLIOTT, J. & LAWRENCE, J. 2014. Refining childhood social class measures in the 1958 British cohort study. CLS Cohort Studies Working Paper 2014/1.

ELLIOTT, J. & SHEPHERD, P. 2006. Cohort Profile: 1970 British Birth Cohort (BCS70). *International Journal of Epidemiology,* 35**,** 836-843*,* DOI: http://dx.doi.org/10.1093/ije/dyl174.

ENDERS, C. K. 2010. *Applied missing data analysis*, Guilford Press.

EVANS-POLCE, R., LANZA, S. & MAGGS, J. 2016. Heterogeneity of alcohol, tobacco, and other substance use behaviors in U.S. college students: A latent class analysis. *Addictive Behaviors,* 53**,** 80-85*,* DOI: http://dx.doi.org/10.1016/j.addbeh.2015.10.010.

EZZATI, M., OBERMEYER, Z., TZOULAKI, I., MAYOSI, B. M., ELLIOTT, P. & LEON, D. A. 2015. Contributions of risk factors and medical care to cardiovascular mortality trends. *Nat Rev Cardiol,* 12**,** 508-530*,* DOI: 10.1038/nrcardio.2015.82 http://www.nature.com/nrcardio/journal/v12/n9/abs/nrcardio.2015.82.html#supplementary-information.

FAIRCHILD, A. J. & MCQUILLIN, S. D. 2010. Evaluating mediation and moderation effects in school psychology: A presentation of methods and review of current practice. *Journal of school psychology,* 48**,** 53-84.

FALKSTEDT, D., MOLLER, J., ZEEBARI, Z. & ENGSTROM, K. 2016. Prevalence, co-occurrence, and clustering of health-risk behaviors among people with different socio-economic trajectories: A population-based study. *Preventive Medicine,* 93**,** 64-69*,* DOI: 10.1016/j.ypmed.2016.09.017.

FARRAR, D. E. & GLAUBER, R. R. 1967. Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics***,** 92-107.

FENG, X., CROTEAU, K., KOLT, G. S. & ASTELL-BURT, T. 2016. Does retirement mean more physical activity? A longitudinal study. *BMC Public Health,* 16**,** 1-7*,* DOI: 10.1186/s12889-016-3253-0.

FENNIS, B. M., ANDREASSEN, T. W. & LERVIK-OLSEN, L. 2015. Behavioral Disinhibition Can Foster Intentions to Healthy Lifestyle Change by Overcoming Commitment to Past Behavior. *PLoS ONE,* 10**,** e0142489*,* DOI: 10.1371/journal.pone.0142489.

FERRI, E., BYNNER, J. M., WADSWORTH, M. & EDUCATION, U. O. L. I. O. 2003. *Changing Britain, Changing Lives: Three Generations at the Turn of the Century*, Institute of Education, University of London, http://books.google.co.uk/books?id=PQ7fAAAACAAJ.

FIDLER, J., FERGUSON, S. G., BROWN, J., STAPLETON, J. & WEST, R. 2013. How does rate of smoking cessation vary by age, gender and social grade? Findings from a population survey in England. *Addiction,* 108**,** 1680-5*,* DOI: 10.1111/add.12241.

FILIPPIDIS, F. T., AGAKU, I. T. & VARDAVAS, C. I. 2016. Geographic variation and socio-demographic determinants of the co-occurrence of risky health behaviours in 27 European Union member states. *Journal of Public Health,* 38**,** E13-E20*,* DOI: 10.1093/pubmed/fdv061.

FINCH, H. 2015. A Comparison of Statistics for Assessing Model Invariance in Latent Class Analysis. *Open Journal of Statistics,* 5**,** 191.

FINCH, W. H. & BRONK, K. C. 2011. Conducting Confirmatory Latent Class Analysis Using M plus. *Structural Equation Modeling,* 18**,** 132-151.
FIRESTONE, R., ROWE, C. J., MODI, S. N. & SIEVERS, D. 2016. The effectiveness of social marketing in global health: a systematic review. *Health Policy and Planning,* DOI: 10.1093/heapol/czw088.
FLEIG, L., KÜPER, C., LIPPKE, S., SCHWARZER, R. & WIEDEMANN, A. U. 2015. Cross-behavior associations and multiple health behavior change: A longitudinal study on physical activity and fruit and vegetable intake. *Journal of Health Psychology,* 20**,** 525-534*,* DOI: 10.1177/1359105315574951.
FLEMING, P. J. & AGNEW-BRUNE, C. 2015. Current trends in the study of gender norms and health behaviors. *Current Opinion in Psychology,* 5**,** 72-77*,* DOI: http://dx.doi.org/10.1016/j.copsyc.2015.05.001.
FONE, D. L., FAREWELL, D. M., WHITE, J., LYONS, R. A. & DUNSTAN, F. D. 2013. Socioeconomic patterning of excess alcohol consumption and binge drinking: a cross-sectional study of multilevel associations with neighbourhood deprivation. *BMJ Open,* 3*,* DOI: 10.1136/bmjopen-2012-002337.
FORD, E. S., LI, C., ZHAO, G., PEARSON, W. S., TSAI, J. & GREENLUND, K. J. 2010. Trends in low-risk lifestyle factors among adults in the United States: Findings from the Behavioral Risk Factor Surveillance System 1996–2007. *Preventive Medicine,* 51**,** 403-407*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2010.08.002.
FOTHERGILL, K. E., ENSMINGER, M. E., GREEN, K. M., ROBERTSON, J. A. & JUON, H. S. 2009. Pathways to adult marijuana and cocaine use: a prospective study of African Americans from age 6 to 42. *Journal of health and social behavior,* 50**,** 65-81.
FRANSSON, E. I., HEIKKILÄ, K., NYBERG, S. T., ZINS, M., WESTERLUND, H., WESTERHOLM, P., VÄÄNÄNEN, A., VIRTANEN, M., VAHTERA, J., THEORELL, T., SUOMINEN, S., SINGH-MANOUX, A., SIEGRIST, J., SABIA, S., RUGULIES, R., PENTTI, J., OKSANEN, T., NORDIN, M., NIELSEN, M. L., MARMOT, M. G., MAGNUSSON HANSON, L. L., MADSEN, I. E. H., LUNAU, T., LEINEWEBER, C., KUMARI, M., KOUVONEN, A., KOSKINEN, A., KOSKENVUO, M., KNUTSSON, A., KITTEL, F., JÖCKEL, K.-H., JOENSUU, M., HOUTMAN, I. L., HOOFTMAN, W. E., GOLDBERG, M., GEUSKENS, G. A., FERRIE, J. E., ERBEL, R., DRAGANO, N., DE BACQUER, D., CLAYS, E., CASINI, A., BURR, H., BORRITZ, M., BONENFANT, S., BJORNER, J. B., ALFREDSSON, L., HAMER, M., BATTY, G. D. & KIVIMÄKI, M. 2012. Job Strain as a Risk Factor for Leisure-Time Physical Inactivity: An Individual-Participant Meta-Analysis of Up to 170,000 Men and Women: The IPD-Work Consortium. *American Journal of Epidemiology,* DOI: 10.1093/aje/kws336.
FRENCH, M. T., POPOVICI, I. & MACLEAN, J. C. 2009. Do alcohol consumers exercise more? Findings from a national survey. *American Journal of Health Promotion,* 24**,** 2-10.
FRENCH, S., ROSENBERG, M. & KNUIMAN, M. 2008. The clustering of health risk behaviours in a Western Australian adult population. *Health Promotion Journal of Australia,* 19**,** 203-209.
FRIEL, S., CHOPRA, M. & SATCHER, D. 2007. Unequal weight: equity oriented policy responses to the global obesity epidemic. *BMJ: British Medical Journal,* 335**,** 1241.
FROHLICH, K. L. & POTVIN, L. 2008. Transcending the Known in Public Health Practice. *American Journal of Public Health,* 98**,** 216-221*,* DOI: 10.2105/ajph.2007.114777.
FSA 2008. Consumer Attitudes to Food Standards: Wave 8 UK Report. . London: Food Standards Agency.
GAGNÉ, T., FROHLICH, K. L. & ABEL, T. 2015. Cultural capital and smoking in young adults: applying new indicators to explore social inequalities in health behaviour. *European Journal of Public Health,* 25**,** 818-823*,* DOI: 10.1093/eurpub/ckv069.

GALINDO-RUEDA, F. 2003. The intergenerational effect of parental schooling: Evidence from the British 1947 school leaving age reform. *Centre for Economic Performance, London School of Economics, mimeo*.

GALLIE, D., FELSTEAD, A., GREEN, F. & INANC, H. 2016. The hidden face of job insecurity. *Work, employment and society,* 31**,** 36-53*,* DOI: 10.1177/0950017015624399.

GARRETT, B. E., DUBE, S. R., BABB, S. & MCAFEE, T. 2015. Addressing the Social Determinants of Health to Reduce Tobacco-Related Disparities. *Nicotine & Tobacco Research,* 17**,** 892-897*,* DOI: 10.1093/ntr/ntu266.

GEORG, W. 2016. Transmission of cultural capital and status attainment–an analysis of development between 15 and 45 years of age. *Longitudinal and Life Course Studies,* 7**,** 106-123.

GERONIMUS, A. T., JAMES, S. A., DESTIN, M., GRAHAM, L. F., HATZENBUEHLER, M. L., MURPHY, M. C., PEARSON, J. A., OMARI, A. & THOMPSON, J. P. 2016. Jedi public health: Co-creating an identity-safe culture to promote health equity. *SSM-Population Health,* 2**,** 105-116.

GILMAN, S. E., RENDE, R., BOERGERS, J., ABRAMS, D. B., BUKA, S. L., CLARK, M. A., COLBY, S. M., HITSMAN, B., KAZURA, A. N., LIPSITT, L. P., LLOYD-RICHARDSON, E. E., ROGERS, M. L., STANTON, C. A., STROUD, L. R. & NIAURA, R. S. 2009. Parental Smoking and Adolescent Smoking Initiation: An Intergenerational Perspective on Tobacco Control. *Pediatrics,* 123**,** e274-e281*,* DOI: http://dx.doi.org/10.1542/peds.2008-2251.

GINN, J. & ARBER, S. 1993. Pension Penalties: The Gendered Division of Occupational Welfare. *Work, Employment & Society,* 7**,** 47-70*,* DOI: 10.1177/095001709371003.

GOLDSTEIN, M. G., WHITLOCK, E. P. & DEPUE, J. 2004. Multiple behavioral risk factor interventions in primary care: summary of research evidence. *American journal of preventive medicine,* 27**,** 61-79*,* DOI: 10.1016/j.amepre.2004.04.023.

GOODMAN, A. & GREGG, P. 2010. *Poorer children's educational attainment: how important are attitudes and behaviour?*, Joseph Rowntree Foundation York.

GORARD, S. 2012. Who is eligible for free school meals? Characterising free school meals as a measure of disadvantage in England. *British Educational Research Journal,* 38**,** 1003-1017*,* DOI: 10.1080/01411926.2011.608118.

GRAHAM, H. 2007. *Unequal lives: health and socioeconomic inequalities*, McGraw-Hill Education (UK).

GRAHAM, H. 2012. Smoking, stigma and social class. *Journal of Social Policy,* 41**,** 83-99.

GRAHAM, H., HUTCHINSON, J., LAW, C., PLATT, L. & WARDLE, H. 2016. Multiple health behaviours among mothers and partners in England: Clustering, social patterning and intra-couple concordance. *SSM – Population Health,* DOI: http://dx.doi.org/10.1016/j.ssmph.2016.10.011.

GRAHAM, J. W. 2003. Adding Missing-Data-Relevant Variables to FIML-Based Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal,* 10**,** 80-100*,* DOI: 10.1207/s15328007sem1001_4.

GREEN, M. J. & POPHAM, F. 2016. Life course models: improving interpretation by consideration of total effects. *International Journal of Epidemiology,* DOI: 10.1093/ije/dyw329.

GREENBANK, P. & HEPWORTH, S. 2008. Working Class Students and the Career Decision-making Process: a qualitative study. *Report for the Higher Education Careers Service Unit (HECSU), Manchester, forthcoming*.

GRIFFIN, B., SHERMAN, K., JONES, M. & BAYL-SMITH, P. 2014. The Clustering of Health Behaviours in Older Australians and its Association with Physical and Psychological Status, and Sociodemographic Indicators. *Annals of Behavioral Medicine***,** 1-10*,* DOI: 10.1007/s12160-014-9589-8.

HAGGER-JOHNSON, G., BATTY, G. D., DEARY, I. J. & VON STUMM, S. 2011. Childhood socioeconomic status and adult health: comparing formative and reflective models in the Aberdeen Children of the 1950s Study (prospective cohort study). *Journal of Epidemiology and Community Health,* 65**,** 1024-1029*,* DOI: 10.1136/jech.2010.127696.

HALE, D. R., FITZGERALD-YAU, N. & VINER, R. M. 2014. A Systematic Review of Effective Interventions for Reducing Multiple Health Risk Behaviors in Adolescence. *American Journal of Public Health,* 104**,** e19-e41*,* DOI: 10.2105/ajph.2014.301874.

HALE, D. R. & VINER, R. M. 2016. The correlates and course of multiple health risk behaviour in adolescence. *BMC public health,* 16**,** 1.

HALLERÖD, B. & GUSTAFSSON, J.-E. 2011. A longitudinal analysis of the relationship between changes in socio-economic status and changes in health. *Social Science & Medicine,* 72**,** 116-123*,* DOI: http://dx.doi.org/10.1016/j.socscimed.2010.09.036.

HAMER, M., KIVIMAKI, M. & STEPTOE, A. 2012. Longitudinal patterns in physical activity and sedentary behaviour from mid-life to early old age: a substudy of the Whitehall II cohort. *Journal of Epidemiology and Community Health,* 66**,** 1110-1115*,* DOI: 10.1136/jech-2011-200505.

HARDING, J. F., MORRIS, P. A. & HUGHES, D. 2015. The relationship between maternal education and children's academic outcomes: A theoretical framework. *Journal of Marriage and Family,* 77**,** 60-76.

HARGREAVES, K., AMOS, A., HIGHET, G., MARTIN, C., PLATT, S., RITCHIE, D. & WHITE, M. 2010. The social context of change in tobacco consumption following the introduction of 'smokefree' England legislation: A qualitative, longitudinal study. *Social Science & Medicine,* 71**,** 459-466*,* DOI: 10.1016/j.socscimed.2010.04.025.

HARPER, H. & HALLSWORTH, M. 2016. Counting Calories: How under-reporting can explain the apparent fall in calorie intake. London: Behavioural Insights Team, http://www.behaviouralinsights.co.uk/health/counting-calories-a-new-report-from-bit-on-the-problems-with-official-statistics-on-calorie-intake-and-how-they-can-be-solved/.

HARRINGTON, J., PERRY, I. J., LUTOMSKI, J., FITZGERALD, A. P., SHIELY, F., MCGEE, H., BARRY, M. M., VAN LENTE, E., MORGAN, K. & SHELLEY, E. 2010. Living longer and feeling better: healthy lifestyle, self-rated health, obesity and depression in Ireland. *The European Journal of Public Health,* 20**,** 91-95.

HARRINGTON, J. M., DAHLY, D. L., FITZGERALD, A. P., GILTHORPE, M. S. & PERRY, I. J. 2014. Capturing changes in dietary patterns among older adults: a latent class analysis of an ageing Irish cohort. *Public Health Nutrition,* 17**,** 2674-2686*,* DOI: http://dx.doi.org/10.1017/s1368980014000111.

HARRIS, J., SPRINGETT, J., CROOT, L., BOOTH, A., CAMPBELL, F., THOMPSON, J., GOYDER, E., VAN CLEEMPUT, P., WILKINS, E. & YANG, Y. 2015. Can community-based peer support promote health literacy and reduce inequalities? A realist review.

HAWKES, D. & PLEWIS, I. 2006. Modelling non-response in the national child development study. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* 169**,** 479-491.

HAYES, A. F. 2013. *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*, Guilford Press.

HAYWARD, M. D. & GORMAN, B. K. 2004. The long arm of childhood: The influence of early-life social conditions on men’s mortality. *Demography,* 41**,** 87-107.

HEAD, J., STANSFELD, S. A. & SIEGRIST, J. 2004. The psychosocial work environment and alcohol dependence: a prospective study. *Occupational and Environmental Medicine,* 61**,** 219-224.

HEDEGAARD, L. 2016. Food culture in obesity prevention Europe. *European Journal of Public Health,* DOI: 10.1093/eurpub/ckw048.

HEIKKILÄ, K., NYBERG, S. T., FRANSSON, E. I., ALFREDSSON, L., DE BACQUER, D., BJORNER, J. B., BONENFANT, S., BORRITZ, M., BURR, H., CLAYS, E., CASINI, A., DRAGANO, N., ERBEL, R., GEUSKENS, G. A., GOLDBERG, M., HOOFTMAN, W. E., HOUTMAN, I. L., JOENSUU, M., JÖCKEL, K.-H., KITTEL, F., KNUTSSON, A., KOSKENVUO, M., KOSKINEN, A., KOUVONEN, A., LEINEWEBER, C., LUNAU, T., MADSEN, I. E. H., HANSON, L. L. M., MARMOT, M. G., NIELSEN, M. L., NORDIN, M., PENTTI, J., SALO, P., RUGULIES, R., STEPTOE, A., SIEGRIST, J., SUOMINEN, S., VAHTERA, J., VIRTANEN, M., VÄÄNÄNEN, A., WESTERHOLM, P., WESTERLUND, H., ZINS, M., THEORELL, T., HAMER, M., FERRIE, J. E., SINGH-MANOUX, A., BATTY, G. D., KIVIMÄKI, M. & FOR THE, I. P. D. W. C. 2012. Job Strain and Tobacco Smoking: An Individual-Participant Data Meta-Analysis of 166 130 Adults in 15 European Studies. *PLoS ONE,* 7**,** e35463*,* DOI: 10.1371/journal.pone.0035463.

HEINZ, W. R. & MARSHALL, V. W. 2003. *Social dynamics of the life course: transitions, institutions, and interrelations*, Transaction Publishers.

HERD, P., GOESLING, B. & HOUSE, J. S. 2007. Socioeconomic position and health: the differential effects of education versus income on the onset versus progression of health problems. *Journal of Health and Social Behavior,* 48**,** 223-238.

HERON, J. E., CROUDACE, T. J., BARKER, E. D. & TILLING, K. 2015. A comparison of approaches for assessing covariate effects in latent class analysis. *2015,* 6**,** 15*,* DOI: 10.14301/llcs.v6i4.322.

HEROUX, M., JANSSEN, I., LEE, D. C., SUI, X. M., HEBERT, J. R. & BLAIR, S. N. 2012. Clustering of Unhealthy Behaviors in the Aerobics Center Longitudinal Study. *Prevention Science,* 13**,** 183-195*,* DOI: http://dx.doi.org/10.1007/s11121-011-0255-0.

HESSEL, P. & AVENDANO, M. 2016. Economic downturns during the life-course and late-life health: an analysis of 11 European countries. *The European Journal of Public Health,* DOI: 10.1093/eurpub/ckw063.

HOEK, J. & SMITH, K. 2016. A qualitative analysis of low income smokers' responses to tobacco excise tax increases. *International Journal of Drug Policy,* 37**,** 82-89*,* DOI: http://dx.doi.org/10.1016/j.drugpo.2016.08.010.

HOFFMANN, R., EIKEMO, T. A., KULHÁNOVÁ, I., KULIK, M. C., LOOMAN, C., MENVIELLE, G., DEBOOSERE, P., MARTIKAINEN, P., REGIDOR, E. & MACKENBACH, J. P. 2015. *Obesity and the potential reduction of social inequalities in mortality: evidence from 21 European populations,* DOI: 10.1093/eurpub/ckv090.

HOFSTETTER, H., DUSSELDORP, E., VAN EMPELEN, P. & PAULUSSEN, T. 2014. A primer on the use of cluster analysis or factor analysis to assess co-occurrence of risk behaviors. *Preventive Medicine,* 67**,** 141-146*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2014.07.007.

HOLMES, J., BROWN, J., MEIER, P., BEARD, E., MICHIE, S. & BUYKX, P. 2016. Short-term effects of announcing revised lower risk national drinking guidelines on related awareness and knowledge: a trend analysis of monthly survey data in England. *BMJ Open,* 6*,* DOI: 10.1136/bmjopen-2016-013804.

HORGAN, G. 2007. *The impact of poverty on young children's experience of school*, Citeseer.

HOWE, L. D., LAWLOR, D. A. & PROPPER, C. 2013a. Trajectories of socioeconomic inequalities in health, behaviours and academic achievement across childhood and adolescence. *Journal of epidemiology and community health***,** jech-2012-201892.

HOWE, L. D., SMITH, A. D., MACDONALD-WALLIS, C., ANDERSON, E. L., GALOBARDES, B., LAWLOR, D. A., BEN-SHLOMO, Y., HARDY, R., COOPER, R., TILLING, K. & FRASER, A. 2016. Relationship between mediation analysis and the structured life course approach. *International Journal of Epidemiology,* DOI: 10.1093/ije/dyw254.

HOWE, L. D., TILLING, K., GALOBARDES, B. & LAWLOR, D. A. 2013b. Loss to follow-up in cohort studies: bias in estimates of socioeconomic inequalities. *Epidemiology,* 24**,** 1-9.

HOYLE, R. H. 2012. Path analysis and structural equation modeling with latent variables. *In:* COOPER, H., CAMIC, P. M., LONG, D. L., PANTER, A. T., RINDSKOPF, D. & SHER, K. J. (eds.) *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological.* Washington, DC, US: American Psychological Association.

HOYLE, R. H. & ROBINSON, J. C. 2003. Mediated and moderated effects in social psychological research: Measurement, design, and analysis issues. *In:* SANSONE, C., MORF, C. & PANTER, A. (eds.) *Handbook of methods in social psychology.* Thousand Oaks, CA: Sage.

HSCIC 2014a. Statistics on Obesity, Physical Activity and Diet: England 2014. London: Health and Social Care Information Centre, http://www.hscic.gov.uk/catalogue/PUB13648/Obes-phys-acti-diet-eng-2014-rep.pdf.

HSCIC. 2014b. *Statistics on Smoking, England 2014* [Online]. HSCIC. Available: http://www.hscic.gov.uk/catalogue/PUB14988 [Accessed 20 July 2015].

HU, Y., VAN LENTHE, F. J., PLATT, S., BOSDRIESZ, J. R., LAHELMA, E., MENVIELLE, G., REGIDOR, E., SANTANA, P., DE GELDER, R. & MACKENBACH, J. P. 2016. The impact of tobacco control policies on smoking among socioeconomic groups in nine European countries, 1990-2007. *Nicotine & Tobacco Research,* DOI: 10.1093/ntr/ntw210.

HUNT, K., FORD, G. & MUTRIE, N. 2001. Is sport for all? Exercise and physical activity patterns in early and late middle age in the West of Scotland. *Health Education,* 101**,** 151-158*,* DOI: doi:10.1108/09654280110398725.

HYMAN, D. J., PAVLIK, V. N., TAYLOR, W. C., GOODRICK, G. K. & MOYE, L. 2007. Simultaneous vs sequential counseling for multiple behavior change. *Archives of Internal Medicine,* 167**,** 1152*,* DOI: 10.1001/archinte.167.11.1152.

ILOMÄKI, J., KORHONEN, M. J., LAVIKAINEN, P., LIPTON, R., ENLUND, H. & KAUHANEN, J. 2010. Changes in alcohol consumption and drinking patterns during 11 years of follow-up among ageing men: the FinDrink study. *The European Journal of Public Health,* 20**,** 133-138.

IREDALE, J. M., CLARE, P. J., COURTNEY, R. J., MARTIRE, K. A., BONEVSKI, B., BORLAND, R., SIAHPUSH, M. & MATTICK, R. P. 2016. Associations between behavioural risk factors and smoking, heavy smoking and future smoking among an Australian population-based sample. *Preventive Medicine,* 83**,** 70-76*,* DOI: 10.1016/j.ypmed.2015.11.020.

JACKSON, C. & TINKLER, P. 2007. 'Ladettes' and 'Modern Girls': 'troublesome' young femininities. *The Sociological Review,* 55**,** 251-272*,* DOI: 10.1111/j.1467-954X.2007.00704.x.

JAMES, W. P., NELSON, M., RALPH, A. & LEATHER, S. 1997. Socioeconomic determinants of health. The contribution of nutrition to inequalities in health. *BMJ : British Medical Journal,* 314**,** 1545-1549, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2126753/.

JARVIS, M. J. 1994. A profile of tobacco smoking. *Addiction,* 89**,** 1371-1376*,* DOI: 10.1111/j.1360-0443.1994.tb03732.x.

JEFFERIS, B. J., POWER, C., GRAHAM, H. & MANOR, O. 2004. Changing social gradients in cigarette smoking and cessation over two decades of adult follow-up in a British birth cohort. *Journal of Public Health,* 26**,** 13-18.

JIANG, T., GILTHORPE, M. S., SHIELY, F., HARRINGTON, J. M., PERRY, I. J., KELLEHER, C. C. & TU, Y.-K. 2013. Age-period-cohort analysis for trends in body mass index in Ireland. *BMC Public Health,* 13**,** 1-7*,* DOI: 10.1186/1471-2458-13-889.

JOHNSON, W., LI, L., KUH, D. & HARDY, R. 2015. How Has the Age-Related Process of Overweight or Obesity Development Changed over Time? Co-ordinated Analyses of Individual Participant Data from Five United Kingdom Birth Cohorts. *PLoS Med,* 12**,** e1001828*,* DOI: 10.1371/journal.pmed.1001828.

JOHNSTONE, B. M., LEINO, E. V., AGER, C. R., FERRER, H. & FILLMORE, K. M. 1996. Determinants of life-course variation in the frequency of alcohol consumption: meta-analysis of studies from the collaborative alcohol-related longitudinal project. *J Stud Alcohol,* 57**,** 494-506.

JONES, I. R., PAPACOSTA, O., WHINCUP, P. H., GOYA WANNAMETHEE, S. & MORRIS, R. W. 2011. Class and lifestyle 'lock-in' among middle-aged and older men: a Multiple Correspondence Analysis of the British Regional Heart Study. *Sociology of Health & Illness,* 33**,** 399-419*,* DOI: 10.1111/j.1467-9566.2010.01280.x.

JONES, N. R. V., CONKLIN, A. I., SUHRCKE, M. & MONSIVAIS, P. 2014. The Growing Price Gap between More and Less Healthy Foods: Analysis of a Novel Longitudinal UK Dataset. *PLoS ONE,* 9**,** e109343*,* DOI: 10.1371/journal.pone.0109343.

JUNEAU, C. E., SULLIVAN, A., DODGEON, B., COTE, S., PLOUBIDIS, G. B. & POTVIN, L. 2014. Social class across the life course and physical activity at age 34 years in the 1970 British birth cohort. *Annals of Epidemiology,* 24**,** 641-647*,* DOI: 10.1016/j.annepidem.2014.06.096.

KAMPHUIS, C. B. M., JANSEN, T., MACKENBACH, J. P. & VAN LENTHE, F. J. 2015. Bourdieu's Cultural Capital in Relation to Food Choices: A Systematic Review of Cultural Capital Indicators and an Empirical Proof of Concept. *PLoS ONE,* 10**,** e0130695*,* DOI: 10.1371/journal.pone.0130695.

KAMPHUIS, C. B. M., TURRELL, G., GISKES, K., MACKENBACH, J. P. & VAN LENTHE, F. J. 2013. Life course socioeconomic conditions, adulthood risk factors and cardiovascular mortality among men and women: A 17-year follow up of the GLOBE study. *International Journal of Cardiology,* 168**,** 2207-2213*,* DOI: 10.1016/j.ijcard.2013.01.219.

KAPLAN, M. S., HUGUET, N., FEENY, D., MCFARLAND, B. H., CAETANO, R., BERNIER, J., GIESBRECHT, N., OLIVER, L. & ROSS, N. 2012. Alcohol use patterns and trajectories of health-related quality of life in middle-aged and older adults: a 14-year population-based study. *Journal of studies on alcohol and drugs,* 73**,** 581.

KARVONEN, S., RIMPELA, A. H. & RIMPELA, M. K. 1999. Social mobility and health related behaviours in young people. *Journal of Epidemiology and Community Health,* 53**,** 211-217, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1756866/.

KATIKIREDDI, S. V., HIGGINS, M., SMITH, K. E. & WILLIAMS, G. 2013. Health inequalities: the need to move beyond bad behaviours. *Journal of Epidemiology and Community Health,* DOI: 10.1136/jech-2012-202064.

KELLY, S., MARTIN, S., KUHN, I., COWAN, A., BRAYNE, C. & LAFORTUNE, L. 2016. Barriers and Facilitators to the Uptake and Maintenance of Healthy Behaviours by People at Mid-Life: A Rapid Systematic Review. *PLoS ONE,* 11**,** e0145074*,* DOI: 10.1371/journal.pone.0145074.

KEMM, J. R. 2001. A birth cohort analysis of smoking by adults in Great Britain 1974–1998. *Journal of Public Health,* 23**,** 306-311.

KERR, W. C., FILLMORE, K. M. & BOSTROM, A. 2002. Stability of alcohol consumption over time: evidence from three longitudinal surveys from the United States. *Journal of Studies on Alcohol and Drugs,* 63**,** 325.

KERR, W. C., GREENFIELD, T. K., BOND, J., YE, Y. & REHM, J. 2009. Age-period-cohort modelling of alcohol volume and heavy drinking days in the US National Alcohol Surveys: divergence in younger and older adult trends. *Addiction,* 104**,** 27-37*,* DOI: 10.1111/j.1360-0443.2008.02391.x.

KESTILA, L., MAKI-OPAS, T., KUNST, A. E., BORODULIN, K., RAHKONEN, O. & PRÄTTÄLÄ, R. 2015. Childhood Adversities and Socioeconomic Position as Predictors of Leisure-Time Physical Inactivity in Early Adulthood. *Journal of Physical Activity & Health,* 12**,** 193-199*,* DOI: 10.1123/jpah.2013-0245.

KEYES, K. M., LI, G. & HASIN, D. S. 2011. Birth Cohort Effects and Gender Differences in Alcohol Epidemiology: A Review and Synthesis. *Alcoholism: Clinical and Experimental Research,* 35**,** 2101-2112*,* DOI: 10.1111/j.1530-0277.2011.01562.x.
KHAW, K.-T., WAREHAM, N., BINGHAM, S., WELCH, A., LUBEN, R. & DAY, N. 2008. Combined impact of health behaviours and mortality in men and women: the EPIC-Norfolk prospective population study. *PLoS medicine,* 5**,** e12.
KHOURY, M. J. & EVANS, J. P. 2015. A public health perspective on a national precision medicine cohort: balancing long-term knowledge generation with early health benefit. *Jama,* 313**,** 2117-2118.
KING, A. C., CASTRO, C. M., BUMAN, M. P., HEKLER, E. B., URIZAR, G. G. & AHN, D. K. 2013. Behavioral Impacts of Sequentially versus Simultaneously Delivered Dietary Plus Physical Activity Interventions: the CALM Trial. *Annals of Behavioral Medicine,* 46**,** 157-168*,* DOI: 10.1007/s12160-013-9501-y.
KING, K., MEADER, N., WRIGHT, K., GRAHAM, H., POWER, C., PETTICREW, M., WHITE, M. & SOWDEN, A. J. 2015. Characteristics of Interventions Targeting Multiple Lifestyle Risk Behaviours in Adult Populations: A Systematic Scoping Review. *PLoS ONE,* 10**,** e0117015*,* DOI: 10.1371/journal.pone.0117015.
KIRKWOOD, B. & STERNE, J. 2003. Essential Medical Statistics: Blackwell Science. *West Sussex, UK*.
KIVIMÄKI, M., NYBERG, S. T., FRANSSON, E. I., HEIKKILÄ, K., ALFREDSSON, L., CASINI, A., CLAYS, E., DE BACQUER, D., DRAGANO, N., FERRIE, J. E., GOLDBERG, M., HAMER, M., JOKELA, M., KARASEK, R., KITTEL, F., KNUTSSON, A., KOSKENVUO, M., NORDIN, M., OKSANEN, T., PENTTI, J., RUGULIES, R., SALO, P., SIEGRIST, J., SUOMINEN, S. B., THEORELL, T., VAHTERA, J., VIRTANEN, M., WESTERHOLM, P. J. M., WESTERLUND, H., ZINS, M., STEPTOE, A., SINGH-MANOUX, A. & BATTY, G. D. 2013. Associations of job strain and lifestyle risk factors with risk of coronary artery disease: a meta-analysis of individual participant data. *Canadian Medical Association Journal,* 185**,** 763-769*,* DOI: 10.1503/cmaj.121735.
KLINE, R. B. 2011. *Principles and practice of structural equation modeling*, Guilford press.
KNAI, C., PETTICREW, M., DURAND, M. A., EASTMURE, E., JAMES, L., MEHROTRA, A., SCOTT, C. & MAYS, N. 2015a. Has a public–private partnership resulted in action on healthier diets in England? An analysis of the Public Health Responsibility Deal food pledges. *Food Policy,* 54**,** 1-10*,* DOI: http://dx.doi.org/10.1016/j.foodpol.2015.04.002.
KNAI, C., PETTICREW, M., DURAND, M. A., SCOTT, C., JAMES, L., MEHROTRA, A., EASTMURE, E. & MAYS, N. 2015b. The Public Health Responsibility deal: has a public–private partnership brought about action on alcohol reduction? *Addiction,* 110**,** 1217-1225*,* DOI: 10.1111/add.12892.
KNÄUPER, B., RABIAU, M., COHEN, O. & PATRICIU, N. 2004. Compensatory health beliefs: scale development and psychometric properties. *Psychology & Health,* 19**,** 607-624*,* DOI: 10.1080/0887044042000196737.
KOH, H. K., PIOTROWSKI, J. J., KUMANYIKA, S. & FIELDING, J. E. 2011. Healthy People a 2020 vision for the social determinants approach. *Health Education & Behavior,* 38**,** 551-557.
KOSHY, P., MACKENZIE, M., LESLIE, W., LEAN, M. & HANKEY, C. 2012. Eating the elephant whole or in slices: views of participants in a smoking cessation intervention trial on multiple behaviour changes as sequential or concurrent tasks. *Bmc Public Health,* 12*,* DOI: 10.1186/1471-2458-12-500.
KRAUS, L., TINGHÖG, M. E., LINDELL, A., PABST, A., PIONTEK, D. & ROOM, R. 2015. Age, Period and Cohort Effects on Time Trends in Alcohol Consumption in the Swedish Adult Population 1979–2011. *Alcohol and Alcoholism,* DOI: 10.1093/alcalc/agv013.

KRIEGER, N. & SMITH, G. D. 2016. The tale wagged by the DAG: broadening the scope of causal inference and explanation for epidemiology. *International Journal of Epidemiology*, dyw114.
KRUEGER, P. M. & CHANG, V. W. 2008. Being poor and coping with stress: Health behaviors and the risk of death. *American Journal of Public Health,* 98**,** 889-896*,* DOI: 10.2105/ajph.2007.114454.
KVAAVIK, E., ANDERSEN, L. F. & KLEPP, K. I. 2005. The stability of soft drinks intake from adolescence to adult age and the association between long-term consumption of soft drinks and lifestyle factors and body weight. *Public Health Nutrition,* 8**,** 149-157*,* DOI: 10.1079/phn2004669.
KVAAVIK, E., BATTY, G., URSIN, G., HUXLEY, R. & GALE, C. R. 2010. Influence of individual and combined health behaviors on total and cause-specific mortality in men and women: The united kingdom health and lifestyle survey. *Archives of Internal Medicine,* 170**,** 711-718*,* DOI: http://dx.doi.org/10.1001/archinternmed.2010.76.
KVAAVIK, E., GLYMOUR, M., KLEPP, K. I., TELL, G. S. & BATTY, G. D. 2012. Parental education as a predictor of offspring behavioural and physiological cardiovascular disease risk factors. *European Journal of Public Health,* 22**,** 544-550*,* DOI: 10.1093/eurpub/ckr106.
KVAAVIK, E., MEYER, H. E. & TVERDAL, A. 2004. Food habits, physical activity and body mass index in relation to smoking status in 40–42 year old Norwegian women and men. *Preventive Medicine,* 38**,** 1-5*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2003.09.020.
LAAKSONEN, M., LUOTO, R., HELAKORPI, S. & UUTELA, A. 2002. Associations between health-related behaviors: a 7-year follow-up of adults. *Preventive medicine,* 34**,** 162-170.
LAAKSONEN, M., PRÄTTÄLÄ, R. & KARISTO, A. 2001. Patterns of unhealthy behaviour in Finland. *The European Journal of Public Health,* 11**,** 294-300*,* DOI: http://dx.doi.org/10.1093/eurpub/11.3.294.
LAAKSONEN, M., TALALA, K., MARTELIN, T., RAHKONEN, O., ROOS, E., HELAKORPI, S., LAATIKAINEN, T. & PRÄTTÄLÄ, R. 2008. Health behaviours as explanations for educational level differences in cardiovascular and all-cause mortality: a follow-up of 60 000 men and women over 23 years. *The European Journal of Public Health,* 18**,** 38-43*,* DOI: 10.1093/eurpub/ckm051.
LACEY, R. E., BARTLEY, M., PIKHART, H., STAFFORD, M. & CABLE, N. 2014. Parental separation and adult psychological distress: an investigation of material and relational mechanisms. *BMC Public Health,* 14**,** 1-10*,* DOI: 10.1186/1471-2458-14-272.
LACEY, R. E., CABLE, N., STAFFORD, M., BARTLEY, M. & PIKHART, H. 2010. Childhood socio-economic position and adult smoking: are childhood psychosocial factors important? Evidence from a British birth cohort. *The European Journal of Public Health,* DOI: 10.1093/eurpub/ckq179.
LACHMAN, M. E. & JAMES, J. B. 1997. Charting the course of midlife development: An overview. *In:* LACHMAN, M. E. & JAMES, J. B. (eds.) *Multiple paths of midlife development.* Chicago, US: The University of Chicago.
LACHMAN, M. E., TESHALE, S. & AGRIGOROAEI, S. 2015. Midlife as a Pivotal Period in the Life Course: Balancing Growth and Decline at the Crossroads of Youth and Old Age. *Int J Behav Dev,* 39**,** 20-31*,* DOI: 10.1177/0165025414533223.
LAHTI-KOSKI, M., PIETINEN, P., HELIÖVAARA, M. & VARTIAINEN, E. 2002. Associations of body mass index and obesity with physical activity, food choices, alcohol intake, and smoking in the 1982–1997 FINRISK Studies. *The American Journal of Clinical Nutrition,* 75**,** 809-817, http://ajcn.nutrition.org/content/75/5/809.abstract.

LALLUKKA, T., LAHELMA, E., RAHKONEN, O., ROOS, E., LAAKSONEN, E., MARTIKAINEN, P., HEAD, J., BRUNNER, E., MOSDOL, A., MARMOT, M., SEKINE, M., NASERMOADDELI, A. & KAGAMIMORI, S. 2008. Associations of job strain and working overtime with adverse health behaviors and obesity: Evidence from the Whitehall II Study, Helsinki Health Study, and the Japanese Civil Servants Study. *Social Science & Medicine,* 66**,** 1681-1698*,* DOI: http://dx.doi.org/10.1016/j.socscimed.2007.12.027.

LAMPURE, A., DEGLAIRE, A., SCHLICH, P., CASTETBON, K., PENEAU, S., HERCBERG, S. & MEJEAN, C. 2014. Liking for fat is associated with sociodemographic, psychological, lifestyle and health characteristics. *British Journal of Nutrition,* 112**,** 1353-1363*,* DOI: 10.1017/s0007114514002050.

LAWDER, R., HARDING, O., STOCKTON, D., FISCHBACHER, C., BREWSTER, D. H., CHALMERS, J., FINLAYSON, A. & CONWAY, D. I. 2010. Is the Scottish population living dangerously? Prevalence of multiple risk factors: the Scottish Health Survey 2003. *Bmc Public Health,* 10*,* DOI: 10.1186/1471-2458-10-330.

LAWLOR, D. A., TILLING, K. & SMITH, G. D. 2017. Triangulation in aetiological epidemiology. *International Journal of Epidemiology***,** dyw314.

LEACH, R., PHILLIPSON, C., BIGGS, S. & MONEY, A. 2013. Baby boomers, consumption and social change: the bridging generation? *International Review of Sociology,* 23**,** 104-122.

LEAHY, R. L. 1981. The Development of the Conception of Economic Inequality. I. Descriptions and Comparisons of Rich and Poor People. *Child Development,* 52**,** 523-532*,* DOI: 10.2307/1129170.

LEASURE, J. L., NEIGHBORS, C., HENDERSON, C. & YOUNG, C. 2015. Exercise and alcohol consumption: What we know, what we need to know, and why it's important. *Frontiers in Psychiatry,* 6*,* DOI: 10.3389/fpsyt.2015.00156.

LEDGERWOOD, A. & SHROUT, P. E. 2011. The trade-off between accuracy and precision in latent variable models of mediation processes. *Journal of Personality and Social Psychology,* 101**,** 1174.

LEICESTER, A. W., F. 2004. The 'fat tax': economic incentives to reduce obesity. London: Institute for Fiscal Studies, https://www.ifs.org.uk/publications/1797.

LIM, S. S., VOS, T., FLAXMAN, A. D., DANAEI, G., SHIBUYA, K., ADAIR-ROHANI, H., ALMAZROA, M. A., AMANN, M., ANDERSON, H. R. & ANDREWS, K. G. 2013. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The lancet,* 380**,** 2224-2260.

LINDSTRÖM, M., MODÉN, B. & ROSVALL, M. 2013. A life-course perspective on economic stress and tobacco smoking: a population-based study. *Addiction*.

LISSNER, L., BENGTSSON, C., BJÖRKELUND, C. & WEDEL, H. 1996. Physical Activity Levels and Changes in Relation to Longevity A Prospective Study of Swedish Women. *American Journal of Epidemiology,* 143**,** 54-62.

LOEF, M. & WALACH, H. 2012. The combined effects of healthy lifestyle behaviors on all cause mortality: A systematic review and meta-analysis. *Preventive Medicine,* 55**,** 163-170*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2012.06.017.

LOEWENSTEIN, G., ASCH, D. A., FRIEDMAN, J. Y., MELICHAR, L. A. & VOLPP, K. G. 2012. Can behavioural economics make us healthier? *BMJ,* 344, http://www.cmu.edu/dietrich/sds/docs/loewenstein/CanBEHealthier.pdf.

LOPRINZI, P. D., BRANSCUM, A., HANKS, J. & SMIT, E. 2016. Healthy Lifestyle Characteristics and Their Joint Association With Cardiovascular Disease Biomarkers in US Adults. *Mayo Clinic Proceedings,* 91**,** 432-442*,* DOI: 10.1016/j.mayocp.2016.01.009.

LUPTON, R., TUNSTALL, R., SIGLE-RUSHTON, W., SABATES, R., MESCHI, E., SALTER, E., DODGEON, B., OBOLENSKAYA, P. & KNEALE, D. 2009. Growing up in social housing in Britain: A profile of four generations from 1946 to the present day.

LYNCH, J., SMITH, G. D., HARPER, S. A. M., HILLEMEIER, M., ROSS, N., KAPLAN, G. A. & WOLFSON, M. 2004. Is Income Inequality a Determinant of Population Health? Part 1. A Systematic Review. *Milbank Quarterly,* 82**,** 5-99*,* DOI: 10.1111/j.0887-378X.2004.00302.x.

LYNCH, K. & O'RIORDAN, C. 1998. Inequality in higher education: A study of class barriers. *British Journal of Sociology of education,* 19**,** 445-478.

MACCALLUM, R. C., ZHANG, S., PREACHER, K. J. & RUCKER, D. D. 2002. On the practice of dichotomization of quantitative variables. *Psychological methods,* 7**,** 19.

MACKINNON, D. P., KRULL, J. L. & LOCKWOOD, C. M. 2000. Equivalence of the Mediation, Confounding and Suppression Effect. *Prevention science*, 1(4), 173, DOI: 10.1023/A:1026595011371.

MACLELLAN, J., SUREY, J., ABUBAKAR, I. & STAGG, H. R. 2015. Peer Support Workers in Health: A Qualitative Metasynthesis of Their Experiences. *PLoS ONE,* 10**,** e0141122*,* DOI: 10.1371/journal.pone.0141122.

MAGGS, J. L., PATRICK, M. E. & FEINSTEIN, L. 2008. Childhood and adolescent predictors of alcohol use and problems in adolescence and adulthood in the National Child Development Study. *Addiction,* 103**,** 7-22*,* DOI: 10.1111/j.1360-0443.2008.02173.x.

MAGNUSSON HANSON, L. L., PERISTERA, P., CHUNGKHAM, H. S. & WESTERLUND, H. 2016. Longitudinal Mediation Modeling of Unhealthy Behaviors as Mediators between Workplace Demands/Support and Depressive Symptoms. *PLOS ONE,* 11**,** e0169276*,* DOI: 10.1371/journal.pone.0169276.

MAIBACH, E. W., MAXFIELD, A., LADIN, K. & SLATER, M. 1996. Translating health psychology into effective health communication the american healthstyles audience segmentation project. *Journal of health psychology,* 1**,** 261-277.

MALLER, C. J. 2015. Understanding health through social practices: performance and materiality in everyday life. *Sociology of Health & Illness,* 37**,** 52-66*,* DOI: 10.1111/1467-9566.12178.

MARGOLIS, R. 2013. Educational Differences in Healthy Behavior Changes and Adherence Among Middle-aged Americans. *Journal of Health and Social Behavior,* 54**,** 353-368*,* DOI: 10.1177/0022146513489312.

MARMOT, M. 1997. Inequality, deprivation and alcohol use. *Addiction,* 92**,** 13-20.

MARMOT, M. 2005. Social determinants of health inequalities. *The Lancet,* 365**,** 1099-1104.

MARMOT, M., BELL, R., 2010. Fair Society Health Lives (The Marmot Review). London: UCL Insitute of Health Equity, http://www.instituteofhealthequity.org/.

MARMOT, M. G., STANSFELD, S., PATEL, C., NORTH, F., HEAD, J., WHITE, I., BRUNNER, E., FEENEY, A. & SMITH, G. D. 1991. Health inequalities among British civil servants: the Whitehall II study. *The Lancet,* 337**,** 1387-1393.

MARTIKAINEN, P., HO, J. Y., PRESTON, S. & ELO, I. T. 2013. The changing contribution of smoking to educational differences in life expectancy: indirect estimates for Finnish men and women from 1971 to 2010. *Journal of Epidemiology and Community Health,* 67**,** 219-224*,* DOI: 10.1136/jech-2012-201266.

MARTIN-DIENER, E., MEYER, J., BRAUN, J., TARNUTZER, S., FAEH, D., ROHRMANN, S. & MARTIN, B. W. 2014. The combined effect on survival of four main behavioural risk factors for non-communicable diseases. *Preventive medicine,* 65**,** 148-152.

MARTINEZ, J. A., SHER, K. J., KRULL, J. L. & WOOD, P. K. 2009. Blue-Collar Scholars?: Mediators and Moderators of University Attrition in First-Generation College Students. *Journal of college student development,* 50**,** 87-103*,* DOI: 10.1353/csd.0.0053.

MARTINEZ, M. E., MARSHALL, J. R. & SECHREST, L. 1998. The arbitrary nature of the factor analytical process. *Am J Epidemiol,* 148**,** 17-9.

MARTINO, F. P., MILLER, P. G., COOMBER, K., HANCOCK, L. & KYPRI, K. 2017. Analysis of Alcohol Industry Submissions against Marketing Regulation. *PLOS ONE,* 12**,** e0170366*,* DOI: 10.1371/journal.pone.0170366.

MATHEW, A. R., HECKMAN, B. W., WAHLQUIST, A. E., GARRETT-MAYER, E. & CARPENTER, M. J. 2016. One-Year Smoking Trajectories among Established Adult Smokers with Low Baseline Motivation to Quit. *Nicotine & Tobacco Research,* DOI: 10.1093/ntr/ntw264.

MATJASKO, J. L., CAWLEY, J. H., BAKER-GOERING, M. M. & YOKUM, D. V. Applying Behavioral Economics to Public Health Policy. *American Journal of Preventive Medicine,* 50**,** S13-S19*,* DOI: 10.1016/j.amepre.2016.02.007.

MC SHARRY, J., OLANDER, E. K. & FRENCH, D. P. 2015. Do Single and Multiple Behavior Change Interventions Contain Different Behavior Change Techniques? A Comparison of Interventions Targeting Physical Activity in Obese Populations. *Health Psychology,* 34**,** 960-965*,* DOI: 10.1037/hea0000185.

MCALONEY, K., GRAHAM, H., LAW, C. & PLATT, L. 2013. A scoping review of statistical approaches to the analysis of multiple health-related behaviours. *Preventive Medicine,* 56**,** 365-371*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2013.03.002.

MCAULEY, A., DENNY, C., TAULBUT, M., MITCHELL, R., FISCHBACHER, C., GRAHAM, B., GRANT, I., O'HAGAN, P., MCALLISTER, D. & MCCARTNEY, G. 2016. Informing Investment to Reduce Inequalities: A Modelling Approach. *PLoS ONE,* 11**,** e0159256*,* DOI: 10.1371/journal.pone.0159256.

MCCARTNEY, G., MAHMOOD, L., LEYLAND, A. H., BATTY, G. D. & HUNT, K. 2011. Contribution of smoking-related and alcohol-related deaths to the gender gap in mortality: evidence from 30 European countries. *Tobacco Control,* 20**,** 166-168*,* DOI: 10.1136/tc.2010.037929.

MCDERMOTT, L., DOBSON, A. & OWEN, N. 2009. Determinants of continuity and change over 10 years in young women's smoking. *Addiction,* 104**,** 478-87*,* DOI: 10.1111/j.1360-0443.2008.02452.x.

MCGILL, R., ANWAR, E., ORTON, L., BROMLEY, H., LLOYD-WILLIAMS, F., MARTIN, O., TAYLOR-ROBINSON, D., GUZMAN-CASTILLO, M., GILLESPIE, D. & MOREIRA, P. 2015. Are interventions to promote healthy eating equally effective for all? Systematic review of socioeconomic inequalities in impact. *BMC public health,* 15**,** 457.

MCKENNA, C. S., LAW, C. & PEARCE, A. 2016. Financial strain, parental smoking and the Great Recession – an analysis of the UK Millennium Cohort Study. *Nicotine & Tobacco Research,* DOI: 10.1093/ntr/ntw269.

MEADE, A. W. & LAUTENSCHLAGER, G. J. 2004. A Monte-Carlo Study of Confirmatory Factor Analytic Tests of Measurement Equivalence/Invariance. *Structural Equation Modeling: A Multidisciplinary Journal,* 11**,** 60-72*,* DOI: 10.1207/s15328007sem1101_5.

MEADER, N., KING, K., MOE-BYRNE, T., WRIGHT, K., GRAHAM, H., PETTICREW, M., POWER, C., WHITE, M. & SOWDEN, A. J. 2016. A systematic review on the clustering and co-occurrence of multiple risk behaviours. *BMC Public Health,* 16**,** 1-9*,* DOI: 10.1186/s12889-016-3373-6.

MEIER, P. S., PURSHOUSE, R. & BRENNAN, A. 2010. Policy options for alcohol price regulation: the importance of modelling population heterogeneity. *Addiction,* 105**,** 383-393*,* DOI: 10.1111/j.1360-0443.2009.02721.x.

MÉJEAN, C., MACOUILLARD, P., CASTETBON, K., KESSE-GUYOT, E. & HERCBERG, S. 2011. Socio-economic, demographic, lifestyle and health characteristics associated with consumption of fatty-sweetened and fatty-salted foods in middle-aged French adults. *British journal of nutrition,* 105**,** 776-786.

MENAI, M., FEZEU, L., CHARREIRE, H., KESSE-GUYOT, E., TOUVIER, M., SIMON, C., WEBER, C., ANDREEVA, V. A., HERCBERG, S. & OPPERT, J.-M. 2014. Changes in Sedentary Behaviours and Associations with Physical Activity through Retirement: A 6-Year Longitudinal Study. *PLoS ONE,* 9**,** e106850*,* DOI: 10.1371/journal.pone.0106850.

MENG, Y., HOLMES, J., HILL-MCMANUS, D., BRENNAN, A. & MEIER, P. S. 2014. Trend analysis and modelling of gender-specific age, period and birth cohort effects on alcohol abstention and consumption level for drinkers in Great Britain using the General Lifestyle Survey 1984–2009. *Addiction,* 109**,** 206-215*,* DOI: 10.1111/add.12330.

MENSAH, F. & HOBCRAFT, J. 2008. Childhood deprivation, health and development: associations with adult health in the 1958 and 1970 British prospective birth cohort studies. *Journal of Epidemiology and Community Health,* 62**,** 599-606, http://jech.bmj.com/content/62/7/599.full.pdf.

MERTENS, E., CLARYS, P., MULLIE, P., LEFEVRE, J., CHARLIER, R., KNAEPS, S., HUYBRECHTS, I. & DEFORCHE, B. 2016. Stability of physical activity, fitness components and diet quality indices. *Eur J Clin Nutr,* DOI: 10.1038/ejcn.2016.172.

MEYER, K. A., GUILKEY, D. K., TIEN, H.-C., KIEFE, C. I., POPKIN, B. M. & GORDON-LARSEN, P. 2016. Instrumental-Variables Simultaneous Equations Model of Physical Activity and Body Mass Index: The Coronary Artery Risk Development in Young Adults (CARDIA) Study. *American Journal of Epidemiology,* 184**,** 465-476*,* DOI: 10.1093/aje/kww010.

MISHRA, G. D., BEN-SHLOMO, Y. & KUH, D. 2011. *A Life Course Approach to Health Behaviors: Theory and Methods,* New York, Springer DOI: 10.1007/978-0-387-09488-5_34.

MISHRA, G. D., MCNAUGHTON, S. A., BRAMWELL, G. D. & WADSWORTH, M. E. J. 2006. Longitudinal changes in dietary patterns during adult life. *British Journal of Nutrition,* 96**,** 735-744*,* DOI: 10.1079/bjn20061871.

MIURA, K. & TURRELL, G. 2014. Contribution of Psychosocial Factors to the Association between Socioeconomic Position and Takeaway Food Consumption. *PLoS ONE,* 9**,** e108799*,* DOI: 10.1371/journal.pone.0108799.

MOLANDER, R. C., YONKER, J. A. & KRAHN, D. D. 2010. Age-Related Changes in Drinking Patterns From Mid-to Older Age: Results From the Wisconsin Longitudinal Study. *Alcoholism: Clinical and Experimental Research,* 34**,** 1182-1192.

MORGAN, O. & BAKER, A. 2006. Measuring deprivation in England and Wales using 2001 Carstairs scores. *Health Stat Q,* 31**,** 28-33.

MORRIS, M. C., MIELOCK, A. S. & RAO, U. 2016. Salivary stress biomarkers of recent nicotine use and dependence. *The American Journal of Drug and Alcohol Abuse,* 42**,** 640-648*,* DOI: 10.1080/00952990.2016.1202263.

MOSTAFA, T. & WIGGINS, D. 2014. Handling attrition and non-response in the 1970 British Cohort Study.

MULDER, M., RANCHOR, A. V., SANDERMAN, R., BOUMA, J. & VAN DEN HEUVEL, W. J. 1998. The stability of lifestyle behaviour. *International Journal of Epidemiology,* 27**,** 199-207*,* DOI: http://dx.doi.org/10.1093/ije/27.2.199.

MUNDWILER, J., SCHÜPBACH, U., DIETERLE, T., LEUPPI, J. D., SCHMIDT-TRUCKSÄSS, A., WOLFER, D. P., MIEDINGER, D. & BRIGHENTI-ZOGG, S. 2017. Association of Occupational and Leisure-Time Physical Activity with Aerobic Capacity in a Working Population. *PLOS ONE,* 12**,** e0168683*,* DOI: 10.1371/journal.pone.0168683.

MUTHÉN, B. O. & HSU, J. W. Y. 1993. Selection and predictive validity with latent variable structures†. *British Journal of Mathematical and Statistical Psychology,* 46**,** 255-271.

MUTHÉN, L. K. M., B.O. 2012. Mplus User's Guide. 7th edition ed. Los Angeles, CA: : Muthén & Muthén
MUTHÉN, M. 2014. Mplus: Version 7. Los Angeles, CA: Muthén & Muthén.
NASH, S. G., MCQUEEN, A. & BRAY, J. H. 2005. Pathways to adolescent alcohol use: family environment, peer influence, and parental expectations. *Journal of Adolescent Health,* 37**,** 19-28*,* DOI: http://dx.doi.org/10.1016/j.jadohealth.2004.06.004.
NATCEN 2011. Health Survey for England, 2000. 4th edition ed.: UK Data Service, DOI: http://doi.org/10.5255/UKDA-SN-4487-1.
NEUHOUSER, M. 2010. Dietary Assessment in Behavioral Medicine. *In:* STEPTOE, A., FREEDLAND, K. E., JENNINGS, J. R., LLABRE, M. M., MANUCK, S. B., SUSMAN, E. J. & POOLE, L. (eds.) *Handbook of behavioral medicine.* New York: Springer Science & Business Media.
NI MHURCHU, C., EYLES, H., GENC, M., SCARBOROUGH, P., RAYNER, M., MIZDRAK, A., NNOAHAM, K. & BLAKELY, T. 2015. Effects of Health-Related Food Taxes and Subsidies on Mortality from Diet-Related Disease in New Zealand: An Econometric-Epidemiologic Modelling Study. *PLoS ONE,* 10**,** e0128477*,* DOI: 10.1371/journal.pone.0128477.
NIGG, C. R., ALLEGRANTE, J. P. & ORY, M. 2002. Theory-comparison and multiple-behavior research: common themes advancing health behavior research. *Health Education Research,* 17**,** 670-679*,* DOI: 10.1093/her/17.5.670.
NOBLE, N., PAUL, C., TURON, H. & OLDMEADOW, C. 2015. Which modifiable health risk behaviours are related? A systematic review of the clustering of Smoking, Nutrition, Alcohol and Physical activity ('SNAP') health risk factors. *Preventive Medicine,* 81**,** 16-41*,* DOI: 10.1016/j.ypmed.2015.07.003.
NOEL, J., LAZZARINI, Z., ROBAINA, K. & VENDRAME, A. 2016. Alcohol Industry Self-Regulation: Who is it really protecting? *Addiction*.
NOMAGUCHI, K. M. & BIANCHI, S. M. 2004. Exercise Time: Gender Differences in the Effects of Marriage, Parenthood, and Employment. *Journal of Marriage and Family,* 66**,** 413-430*,* DOI: 10.1111/j.1741-3737.2004.00029.x.
NOONAN, D., LYNA, P., FISH, L. J., BILHEIMER, A. K., GORDON, K. C., ROBERSON, P., GONZALEZ, A. & POLLAK, K. I. 2016. Unintended Effects of a Smoking Cessation Intervention on Latino Fathers' Binge Drinking In Early Postpartum. *Annals of Behavioral Medicine,* 50**,** 622-627*,* DOI: 10.1007/s12160-016-9781-0.
NUGAWELA, M. D., LANGLEY, T., SZATKOWSKI, L. & LEWIS, S. 2016. Measuring Alcohol Consumption in Population Surveys: A Review of International Guidelines and Comparison with Surveys in England. *Alcohol and Alcoholism,* 51**,** 84-92*,* DOI: 10.1093/alcalc/agv073.
NYBERG, G., SUNDBLOM, E., NORMAN, Å., BOHMAN, B., HAGBERG, J. & ELINDER, L. S. 2015. Effectiveness of a Universal Parental Support Programme to Promote Healthy Dietary Habits and Physical Activity and to Prevent Overweight and Obesity in 6-Year-Old Children: The Healthy School Start Study, a Cluster-Randomised Controlled Trial. *PLoS ONE,* 10**,** e0116876*,* DOI: 10.1371/journal.pone.0116876.
NYLUND, K. L., ASPAROUHOV, T. & MUTHÉN, B. O. 2007. Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal,* 14**,** 535-569*,* DOI: 10.1080/10705510701575396.
O'DOHERTY, M. G., SKIDMORE, P. M. L., YOUNG, I. S., MCKINLEY, M. C., CARDWELL, C., YARNELL, J. W. G., GEY, F. K., EVANS, A. & WOODSIDE, J. V. 2011. Dietary Patterns and Smoking in Northern Irish Men: a Population at High Risk of Coronary Heart Disease. *International Journal for Vitamin and Nutrition Research,* 81**,** 21-33*,* DOI: 10.1024/0300-9831/a000047.

O'REILLY, D., ROSATO, M., MAGUIRE, A. & WRIGHT, D. 2015. Caregiving reduces mortality risk for most caregivers: a census-based record linkage study. *International Journal of Epidemiology,* DOI: 10.1093/ije/dyv172.

OKECHUKWU, C., DAVISON, K. & EMMONS, K. 2014. Changing health behaviors in a social context. *Social Epidemiology,* 365.

ONS 2006. General Household Survey, 2000-2001. 3rd Edition ed.: UK Data Service, DOI: http://doi.org/10.5255/UKDA-SN-4518-1.

ONS. 2010a. *SOC2010 volume 3: the National Statistics Socio-economic classification (NS-SEC rebased on SOC2010)* [Online]. London: ONS. Available: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html#7 [Accessed 29/03/2016 2016].

ONS 2010b. Transport. *In:* HUGHES, M. & CHURCH, J. (eds.) *Social trends no. 40.* London: Office for National Statistics. http://webarchive.nationalarchives.gov.uk/.

OUDE GROENIGER, J. & VAN LENTHE, F. J. 2016. Contribution of time-varying measures of health behaviours to socioeconomic inequalities in mortality: how to understand the underlying mechanisms? *Journal of Epidemiology and Community Health,* DOI: 10.1136/jech-2016-207642.

ØVERSVEEN, E., RYDLAND, H. T., BAMBRA, C. & EIKEMO, T. A. 2017. Rethinking the relationship between socio-economic status and health: Making the case for sociological theory in health inequality research. *Scandinavian Journal of Public Health***,** 1403494816686711*,* DOI: 10.1177/1403494816686711.

PAAVOLA, M., VARTIAINEN, E. & HAUKKALA, A. 2004. Smoking, alcohol use, and physical activity: A 13-year longitudinal study ranging from adolescence into adulthood. *Journal of Adolescent Health,* 35**,** 238-244*,* DOI: http://dx.doi.org/10.1016/j.jadohealth.2003.12.004.

PAFFENBARGER JR, R. S., HYDE, R. T., WING, A. L., LEE, I.-M., JUNG, D. L. & KAMPERT, J. B. 1993. The association of changes in physical-activity level and other lifestyle characteristics with mortality among men. *New England Journal of Medicine,* 328**,** 538-545.

PAMPEL, F. C., KRUEGER, P. M. & DENNEY, J. T. 2010. Socioeconomic Disparities in Health Behaviors. *In:* COOK, K. S. & MASSEY, D. S. (eds.) *Annual Review of Sociology, Vol 36.* Palo Alto: Annual Reviews. DOI: 10.1146/annurev.soc.012809.102529.

PARK, M. H., SOVIO, U., VINER, R. M., HARDY, R. J. & KINRA, S. 2013. Overweight in Childhood, Adolescence and Adulthood and Cardiovascular Risk in Later Life: Pooled Analysis of Three British Birth Cohorts. *Plos One,* 8**,** 6*,* DOI: 10.1371/journal.pone.0070684.

PARRY, W. 2013. Experiences of physical activity at age 10 in the 1970 British Cohort Study.

PARSONS, S., SULLIVAN, A. & BROWN, M. 2013. Research on health and health behaviours based on the 1970 British Cohort Study.

PARSONS, T. J., POWER, C. & MANOR, O. 2006. Longitudinal physical activity and diet patterns in the 1958 British Birth Cohort. *Medicine and science in sports and exercise,* 38**,** 547-554*,* DOI: http://dx.doi.org/10.1249/01.mss.0000188446.65651.67.

PATRICK, H. & NICKLAS, T. A. 2005. A review of family and social determinants of children's eating patterns and diet quality. *Journal of the American College of Nutrition,* 24**,** 83-92, <Go to ISI>://WOS:000228529000002.

PATRICK, K., HEKLER, E. & ESTRIN, D. 2016. Rapid rate of technological development and its implications for research on digital behavior change interventions. *Am J Prev Med*.

PATTERSON, C., EMSLIE, C., MASON, O., FERGIE, G. & HILTON, S. 2016. Content analysis of UK newspaper and online news representations of women's and men's 'binge' drinking: a challenge for communicating evidence-based messages about single-episodic drinking? *BMJ Open,* 6*,* DOI: 10.1136/bmjopen-2016-013124.

PATTERSON, E. & ELINDER, L. S. 2015. *Improvements in school meal quality in Sweden after the introduction of new legislation—a 2-year follow-up,* DOI: 10.1093/eurpub/cku184.

PATTERSON, R. E., HAINES, P. S. & POPKIN, B. M. 1994. Health lifestyle patterns of US adults. *Preventive medicine,* 23**,** 453-460.

PAVALKO, E. K. & CAPUTO, J. 2013. Social Inequality and Health Across the Life Course. *American Behavioral Scientist,* DOI: 10.1177/0002764213487344.

PEARSON, N., BIDDLE, S. J. H. & GORELY, T. 2009. Family correlates of breakfast consumption among children and adolescents. A systematic review. *Appetite,* 52**,** 1-7*,* DOI: http://dx.doi.org/10.1016/j.appet.2008.08.006.

PECHEY, R., JEBB, S. A., KELLY, M. P., ALMIRON-ROIG, E., CONDE, S., NAKAMURA, R., SHEMILT, I., SUHRCKE, M. & MARTEAU, T. M. 2013. Socioeconomic differences in purchases of more vs. less healthy foods and beverages: Analysis of over 25,000 British households in 2010. *Social Science & Medicine,* 92**,** 22-26*,* DOI: http://dx.doi.org/10.1016/j.socscimed.2013.05.012.

PEREIRA, S. M. P., LI, L. & POWER, C. 2014. Early-Life Predictors of Leisure-Time Physical Inactivity in Midadulthood: Findings From a Prospective British Birth Cohort. *American Journal of Epidemiology,* 180**,** 1098-1108*,* DOI: 10.1093/aje/kwu254.

PERRY, B., CICIURKAITE, G., BRADY, C. F. & GARCIA, J. 2016. Partner Influence in Diet and Exercise Behaviors: Testing Behavior Modeling, Social Control, and Normative Body Size. *PLOS ONE,* 11**,** e0169193*,* DOI: 10.1371/journal.pone.0169193.

PETTICREW, M., DOUGLAS, N., D'SOUZA, P., SHI, Y., DURAND, M., KNAI, C., EASTMURE, E. & MAYS, N. 2017. Community Alcohol Partnerships with the alcohol industry: what is their purpose and are they effective in reducing alcohol harms? *Journal of Public Health*.

PHILLIMORE, P., BEATTIE, A. & TOWNSEND, P. 1994. Widening inequality of health in northern England, 1981-91. *BMJ,* 308**,** 1125-1128*,* DOI: 10.1136/bmj.308.6937.1125.

PLEWIS, I. 2004. National Child Development Study and 1970 British Cohort Study technical report: Changes in the NCDS and BCS70 populations and samples over time. Centre for Longitudinal Studies, Bedford Group for Lifecourse and Statistical Studies, Institute of Education, University of London.

POORTINGA, W. 2007. The prevalence and clustering of four major lifestyle risk factors in an English adult population. *Preventive Medicine,* 44**,** 124-128*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2006.10.006.

POORTINGA, W., WHITMARSH, L. & SUFFOLK, C. 2013. The introduction of a single-use carrier bag charge in Wales: Attitude change and behavioural spillover effects. *Journal of Environmental Psychology,* 36**,** 240-247*,* DOI: http://dx.doi.org/10.1016/j.jenvp.2013.09.001.

POST 2015. Sugar and Health. *POSTnote.* House of Commons, http://researchbriefings.parliament.uk/ResearchBriefing/Summary/POST-PN-0493.

POWER, C. & ELLIOTT, J. 2006. Cohort profile: 1958 British birth cohort (National Child Development Study). *International Journal of Epidemiology,* 35**,** 34-41*,* DOI: http://dx.doi.org/10.1093/ije/dyi183.

POWER, C., GRAHAM, H., DUE, P., HALLQVIST, J., JOUNG, I., KUH, D. & LYNCH, J. 2005. The contribution of childhood and adult socioeconomic position to adult obesity and smoking behaviour: an international comparison. *International Journal of Epidemiology,* 34**,** 335-344*,* DOI: 10.1093/ije/dyh394.

POWER, C. & HERTZMAN, C. 1997. Social and biological pathways linking early life and adult disease. *British medical bulletin,* 53**,** 210-222.

POWER, C., RODGERS, B. & HOPE, S. 1998. U-shaped relation for alcohol consumption and health in early adulthood and implications for mortality. *The Lancet,* 352**,** 877.

PRANDY, K. 1990. The Revised Cambridge Scale of Occupations. *Sociology,* 24**,** 629-655.

PRANDY, K. 1999. Class, stratification and inequalities in health: a comparison of the Registrar-General's Social Classes and the Cambridge Scale. *Sociology of Health & Illness,* 21**,** 466-484*,* DOI: 10.1111/1467-9566.00167.

PRANDY, K. & LAMBERT, P. 2003. Marriage, social distance and the social space: an alternative derivation and validation of the Cambridge Scale. *Sociology,* 37**,** 397-411.

PRÄTTÄLÄ, R., LAAKSONEN, M. & RAHKONEN, O. 1998. Smoking and unhealthy food habits: How stable is the association? *The European Journal of Public Health,* 8**,** 28-33*,* DOI: 10.1093/eurpub/8.1.28.

PRENTICE, A. M. & JEBB, S. A. 1995. Obesity in Britain: gluttony or sloth? *BMJ,* 311**,** 437-439*,* DOI: 10.1136/bmj.311.7002.437.

PRESTON, S. H. & WANG, H. 2006. Sex mortality differences in the United States: The role of cohort smoking patterns. *Demography,* 43**,** 631-646.

PROCHASKA, J. J., SPRING, B. & NIGG, C. R. 2008a. Multiple health behavior change research: an introduction and overview. *Preventive medicine,* 46**,** 181-188*,* DOI: 10.1016/j.ypmed.2008.02.001.

PROCHASKA, J. J., VELICER, W. F., NIGG, C. R. & PROCHASKA, J. O. 2008b. Methods of quantifying change in multiple risk factor interventions. *Preventive Medicine,* 46**,** 260-265*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2007.07.035.

PUDROVSKA, T. & ANISHKIN, A. 2013. Early-Life Socioeconomic Status and Physical Activity in Later Life Evidence From Structural Equation Models. *Journal of aging and health,* 25**,** 383-404.

PURSHOUSE, R., BRENNAN, A., MOYO, D., NICHOLLS, J. & NORMAN, P. 2017. Typology and Dynamics of Heavier Drinking Styles in Great Britain: 1978-2010. *Alcohol and alcoholism (Oxford, Oxfordshire)*.

PUTTELAAR, J., VERAIN, M. & ONWEZEN, M. 2016. The Potential of Enriching Food Consumption Data by use of Consumer Generated Data: a case from RICHFIELDS. *Proceedings of Measuring Behavior 2016.* Dublin, Ireland, http://www.richfields.eu/wp-content/uploads/2016/05/MB-2016-The_Potential_of_Enriching_Food_Consumption_Data_by_use_of_Consumer_Generated_Data_final2331_FINAL.pdf.

RAQUE-BOGDAN, T. L. & LUCAS, M. S. 2016. Career Aspirations and the First Generation Student: Unraveling the Layers With Social Cognitive Career Theory. *Journal of College Student Development,* 57**,** 248-262.

RCP 2012. Fifty years since smoking and health. London: Royal College of Physicians, https://www.rcplondon.ac.uk/sites/default/files/fifty-years-smoking-health.pdf.

REBOUSSIN, B. A., IP, E. H. & WOLFSON, M. 2008. Locally dependent latent class models with covariates: an application to under-age drinking in the USA. *Journal of the Royal Statistical Society. Series A, (Statistics in Society),* 171**,** 877-897*,* DOI: 10.1111/j.1467-985X.2008.00544.x.

REBOUSSIN, B. A., REBOUSSIN, D. M., LIANG, K.-Y. & ANTHONY, J. C. 1998. Latent transition modeling of progression of health-risk behavior. *Multivariate Behavioral Research,* 33**,** 457-478, http://www.tandfonline.com/doi/pdf/10.1207/s15327906mbr3304_2.

REID, A., CIALDINI, R. & AIKEN, L. 2010. Social Norms and Health Behavior. *In:* STEPTOE, A., FREEDLAND, K. E., JENNINGS, J. R., LLABRE, M. M., MANUCK, S. B., SUSMAN, E. J. & POOLE, L. (eds.) *Handbook of behavioral medicine.* New York: Springer Science & Business Media.

REITHER, E. N., HAUSER, R. M. & YANG, Y. 2009. Do Birth Cohorts Matter? Age-Period-Cohort Analyses of the Obesity Epidemic in the United States. *Social science & medicine (1982),* 69**,** 1439-1448*,* DOI: 10.1016/j.socscimed.2009.08.040.

ROBERTSON-WILSON, J. E., DARGAVEL, M. D., BRYDEN, P. J. & GILES-CORTI, B. 2012. Physical activity policies and legislation in schools: a systematic review. *Am J Prev Med,* 43.

ROLLS, B. J. & MILLER, D. L. 1997. Is the low-fat message giving people a license to eat more? *Journal of the American College of Nutrition,* 16**,** 535-543.

ROOM, R. 2004. Smoking and drinking as complementary behaviours. *Biomedicine & pharmacotherapy,* 58**,** 111-115.

ROSE, D. 2008. *Socio-economic Classifications: Classes and Scales, Measurement and Theories* [Online]. University of Essex: ISER. Available: http://www.iser.essex.ac.uk/files/esec/presentations_and_publications/Measurement_Socstrat.doc [Accessed 14/12/15 2015].

ROSE, G. 1989. High-Risk and Population Strategies of Prevention: Ethical Considerations. *Annals of Medicine,* 21**,** 409-413*,* DOI: 10.3109/07853898909149231.

RSPH 2014. Alcohol calorie labelling. London, UK: RSPH, https://www.rsph.org.uk/our-work/campaigns/alcohol-calorie-labelling-.html.

RUNDLE, A. G., SHEEHAN, D. M., QUINN, J. W., BARTLEY, K., EISENHOWER, D., BADER, M. M. D., LOVASI, G. S. & NECKERMAN, K. M. 2016. Using GPS Data to Study Neighborhood Walkability and Physical Activity. *American Journal of Preventive Medicine,* 50**,** e65-e72*,* DOI: 10.1016/j.amepre.2015.07.033.

RUSSO, F. & WILLIAMSON, J. 2007. Interpreting Causality in the Health Sciences. *International Studies in the Philosophy of Science,* 21**,** 157-170*,* DOI: 10.1080/02698590701498084.

SABIA, S., SINGH-MANOUX, A., HAGGER-JOHNSON, G., CAMBOIS, E., BRUNNER, E. J. & KIVIMAKI, M. 2012. Influence of individual and combined healthy behaviours on successful aging. *Cmaj,* 184**,** 1985-92*,* DOI: 10.1503/cmaj.121080.

SACKER, A., BARTLEY, M., FIRTH, D. & FITZPATRICK, R. 2001. Dimensions of social inequality in the health of women in England: occupational, material and behavioural pathways. *Social science & medicine,* 52**,** 763-781.

SAVAGE, J. S., FISHER, J. O. & BIRCH, L. L. 2007. Parental Influence on Eating Behavior: Conception to Adolescence. *The Journal of Law, Medicine & Ethics,* 35**,** 22-34*,* DOI: 10.1111/j.1748-720X.2007.00111.x.

SAYON-OREA, C., MARTINEZ-GONZALEZ, M. A. & BES-RASTROLLO, M. 2011. Alcohol consumption and body weight: a systematic review. *Nutrition Reviews,* 69**,** 419-431*,* DOI: 10.1111/j.1753-4887.2011.00403.x.

SCHNEIDER, S., HUY, C., SCHUESSLER, M., DIEHL, K. & SCHWARZ, S. 2009. Optimising lifestyle interventions: identification of health behaviour patterns by cluster analysis in a German 50 survey. *European Journal of Public Health,* 19**,** 271-277*,* DOI: http://dx.doi.org/10.1093/eurpub/ckn144.

SCHNITTKER, J. & MCLEOD, J. D. 2005. The Social Psychology of Health Disparities. *Annual Review of Sociology,* 31**,** 75-103*,* DOI: 10.1146/annurev.soc.30.012703.110622.

SCHOOLING, C. M. & KUH, D. 2002. A life course perspective on women's health behaviours. *In:* KUH, D. & HARDY, R. (eds.) *A life course approach to women's health.* Oxford: Oxford University Press.

SCHOON, I. 2006. *Risk and Resilience: Adaptations in Changing Times,* Cambridge, Cambridge University press.
SCHOON, I. & PARSONS, S. 2003. Lifestyle and Health-Related Behaviour. *Changing Britain, Changing Lives.* London: Institute of Education Press.
SCHUIT, A. J., VAN LOON, A. J. M., TIJHUIS, M. & OCKÉ, M. C. 2002. Clustering of Lifestyle Risk Factors in a General Adult Population. *Preventive Medicine,* 35**,** 219-224*,* DOI: http://dx.doi.org/10.1006/pmed.2002.1064.
SCHULZE, A. & MONS, U. 2005. Trends in cigarette smoking initiation and cessation among birth cohorts of 1926–1970 in Germany. *European Journal of Cancer Prevention,* 14**,** 477-483.
SCHUSTER, E., NEGY, C. & TANTLEFF-DUNN, S. 2013. The effects of appearance-related commentary on body dissatisfaction, eating pathology, and body change behaviors in men. *Psychology of Men & Masculinity,* 14**,** 76.
SEILURI, T., LAHTI, J., RAHKONEN, O., LAHELMA, E. & LALLUKKA, T. 2011. Changes in occupational class differences in leisure-time physical activity: a follow-up study. *International Journal of Behavioral Nutrition and Physical Activity,* 8**,** 1.
SHANKAR, A., MCMUNN, A. & STEPTOE, A. 2010. Health-related behaviors in older adults: relationships with socioeconomic status. *American journal of preventive medicine,* 38**,** 39-46.
SHEPHERD, P. 2012a. 1958 National Child Development Study Ethical Review and Consent. London: Centre for Longitudinal Studies, http://www.cls.ioe.ac.uk/Publications.aspx?sitesectionid=93&.
SHEPHERD, P. 2012b. 1970 British Cohort Study Ethical Review and Consent. London: Centre for Longitudinal Studies, http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=1342&sitesectiontitle=Data+Linkage.
SHIROMA, E. J., COOK, N. R., MANSON, J. E., BURING, J. E., RIMM, E. B. & LEE, I. M. 2015. Comparison of Self-Reported and Accelerometer-Assessed Physical Activity in Older Women. *PLoS ONE,* 10**,** e0145950*,* DOI: 10.1371/journal.pone.0145950.
SHORT, S. E. & MOLLBORN, S. 2015. Social determinants and health behaviors: conceptual frames and empirical advances. *Current Opinion in Psychology,* 5**,** 78-84*,* DOI: http://dx.doi.org/10.1016/j.copsyc.2015.05.002.
SHREEVE, V., STEADMAN, K. & BEVAN, S. 2015. Healthy, working economies: Improving the health and wellbeing of the working age population locally. Lancaster, UK: The work foundation, http://www.theworkfoundation.com/Reports/381/Healthy-Working-Economies.
SIJTSMA, F. P., MEYER, K. A., STEFFEN, L. M., SHIKANY, J. M., VAN HORN, L., HARNACK, L., KROMHOUT, D. & JACOBS, D. R. 2012. Longitudinal trends in diet and effects of sex, race, and education on dietary quality score change: the Coronary Artery Risk Development in Young Adults study. *The American Journal of Clinical Nutrition,* 95**,** 580-586*,* DOI: 10.3945/ajcn.111.020719.
SILVA, D. A. S., PERES, K. G., BOING, A. F., GONZÁLEZ-CHICA, D. A. & PERES, M. A. 2013. Clustering of risk behaviors for chronic noncommunicable diseases: A population-based study in southern Brazil. *Preventive Medicine,* 56**,** 20-24*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2012.10.022.
SILVERWOOD, R. J., NITSCH, D., PIERCE, M., KUH, D. & MISHRA, G. D. 2011. Characterizing Longitudinal Patterns of Physical Activity in Mid-Adulthood Using Latent Class Analysis: Results From a Prospective Cohort Study. *American Journal of Epidemiology,* 174**,** 1406-1415*,* DOI: 10.1093/aje/kwr266.

SINGH-MANOUX, A., MARMOT, M. G. & ADLER, N. E. 2005. Does subjective social status predict health and change in health status better than objective status? *Psychosomatic Medicine,* 67**,** 855-861.

SINGH, G. M., MICHA, R., KHATIBZADEH, S., SHI, P., LIM, S., ANDREWS, K. G., ENGELL, R. E., EZZATI, M., MOZAFFARIAN, D., GLOBAL BURDEN OF DISEASES, N. & CHRONIC DISEASES EXPERT, G. 2015. Global, Regional, and National Consumption of Sugar-Sweetened Beverages, Fruit Juices, and Milk: A Systematic Assessment of Beverage Intake in 187 Countries. *PLoS ONE,* 10**,** e0124845*,* DOI: 10.1371/journal.pone.0124845.

SINGHAMMER, J. & MITTELMARK, M. B. 2010. Standard measures of inter-generational social mobility distort actual patterns of mobility and health behaviour: evidence for a better methodology. *Critical Public Health,* 20**,** 223-232*,* DOI: 10.1080/09581590903032611.

SINGLETON, A. D., LONGLEY, P. A., ALLEN, R. & O'BRIEN, O. 2011. Estimating secondary school catchment areas and the spatial equity of access. *Computers, Environment and Urban Systems,* 35**,** 241-249*,* DOI: http://dx.doi.org/10.1016/j.compenvurbsys.2010.09.006.

SLADE, T., CHAPMAN, C., SWIFT, W., KEYES, K., TONKS, Z. & TEESSON, M. 2016. Birth cohort trends in the global epidemiology of alcohol use and alcohol-related harms in men and women: systematic review and metaregression. *BMJ Open,* 6*,* DOI: 10.1136/bmjopen-2016-011827.

SLATER, M. D. & FLORA, J. A. 1991. Health lifestyles: audience segmentation analysis for public health interventions. *Health Education & Behavior,* 18**,** 221-233.

SMED, S., JENSEN, J. D. & DENVER, S. 2007. Socio-economic characteristics and the effect of taxation as a health policy instrument. *Food Policy,* 32**,** 624-639*,* DOI: http://dx.doi.org/10.1016/j.foodpol.2007.03.002.

SMITHERS, A. & ROBINSON, P. 2010. Worlds Apart: social variation among schools. *London: The Sutton Trust*.

SOLARI, C. D. & MARE, R. D. 2012. Housing crowding effects on children's wellbeing. *Social science research,* 41**,** 464-476*,* DOI: 10.1016/j.ssresearch.2011.09.012.

SOMERVILLE, C., MARTEAU, T. M., KINMONTH, A. L. & COHN, S. 2015. *Public attitudes towards pricing policies to change health-related behaviours: a UK focus group study,* DOI: 10.1093/eurpub/ckv077.

SPRING, B., MOLLER, A. C. & COONS, M. J. 2012. Multiple health behaviours: overview and implications. *Journal of public health,* 34**,** i3-i10.

STAIT, E. & CALNAN, M. 2016. Are differential consumption patterns in health-related behaviours an explanation for persistent and widening social inequalities in health in England? *International Journal for Equity in Health,* 15**,** 11*,* DOI: 10.1186/s12939-016-0461-2.

STAMATAKIS, E., EKELUND, U. & WAREHAM, N. J. 2007. Temporal trends in physical activity in England: The Health Survey for England 1991 to 2004. *Preventive Medicine,* 45**,** 416-423*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2006.12.014.

STAPINSKI, L. A., EDWARDS, A. C., HICKMAN, M., ARAYA, R., TEESSON, M., NEWTON, C. N., KENDLER, K. S. & HERON, J. 2016. Drinking to Cope: a Latent Class Analysis of Coping Motives for Alcohol Use in a Large Cohort of Adolescents. *Prevention Science,* 17**,** 584-594*,* DOI: 10.1007/s11121-016-0652-5.

STATACORP 2014. Stata Statistical Software: Release 14 College Station, TX: StataCorp LP.

STEAD, M., MUNAFO, M., FULLER, E., BAULD, L., ANGUS, K., MACDONALD, L., ATTWOOD, A., ATAYA, A. & PICKERING, K. 2013. Scoping and feasibility study to develop and apply a methodology for retrospective adjustment of alcohol consumption data. University of Stirling: Public Health Research Consortium,

http://phrc.lshtm.ac.uk/papers/FINAL_REPORT_PHRC_Scoping_and_feasibility_study_-_alcohol_consumption_data_(October_2013).pdf.
STEIGER, J. H. 1990. Structural model evaluation and modification: An interval estimation approach. *Multivariate behavioral research,* 25**,** 173-180.
STEPTOE, A. 2007. Evaluating strategies for behavior change to reduce cardiovascular risk. *Nat Clin Pract Cardiovasc Med,* 4**,** 598-599, http://dx.doi.org/10.1038/ncpcardio1003.
STOUTENBERG, M., RETHORST, C. D., LAWSON, O. & READ, J. P. 2016. Exercise training – A beneficial intervention in the treatment of alcohol use disorders? *Drug and Alcohol Dependence,* 160**,** 2-11*,* DOI: http://dx.doi.org/10.1016/j.drugalcdep.2015.11.019.
SUDHARSANAN, N., BEHRMAN, J. R. & KOHLER, H.-P. 2016. Limited common origins of multiple adult health-related behaviors: Evidence from U.S. twins. *Social Science & Medicine,* 171**,** 67-83*,* DOI: http://dx.doi.org/10.1016/j.socscimed.2016.11.002.
SULLIVAN, A., BROWN, M., SULLIVAN, A. & BROWN, M. 2013. Overweight and obesity in mid-life: Evidence from the 1970 birth cohort study at age 42. London: Centre for Longitudinal Studies, http://www.cls.ioe.ac.uk/page.aspx?sitesectionid=795sitesectiontitle=Welcome+to+the+1970+British+Cohort+Study+(BCS70).
SUTTON, S., BAUM, A. & JOHNSTON, M. 2004. *The Sage handbook of health psychology*, Sage.
SWINBURN, B. A., SACKS, G., HALL, K. D., MCPHERSON, K., FINEGOOD, D. T., MOODIE, M. L. & GORTMAKER, S. L. 2011. The global obesity pandemic: shaped by global drivers and local environments. *The Lancet,* 378**,** 804-814*,* DOI: http://dx.doi.org/10.1016/S0140-6736(11)60813-1.
TABACHNICK, B. G., FIDELL, L. S. & OSTERLIND, S. J. 2001. Using multivariate statistics.
TAYLOR, A. W., SHI, Z., MONTGOMERIE, A., DAL GRANDE, E. & CAMPOSTRINI, S. 2015. The Use of a Chronic Disease and Risk Factor Surveillance System to Determine the Age, Period and Cohort Effects on the Prevalence of Obesity and Diabetes in South Australian Adults-2003–2013. *PloS one,* 10**,** e0125233.
THOMPSON, B. 2004. *Exploratory and confirmatory factor analysis: Understanding concepts and applications*, American Psychological Association.
THOMPSON, F. E. & SUBAR, A. F. 2008. Dietary assessment methodology. *In:* COULSTON, AM, BOUSHEY & CJ (eds.) *Nutrition in the Prevention and Treatment of Disease.* 2nd ed ed. Burlington, MA: Elsevier Academic Press.
THORPE, M. G., MILTE, C. M., CRAWFORD, D. & MCNAUGHTON, S. A. 2016. A comparison of the dietary patterns derived by principal component analysis and cluster analysis in older Australians. *International Journal of Behavioral Nutrition and Physical Activity,* 13**,** 30*,* DOI: 10.1186/s12966-016-0353-2.
TIAN, J., GALL, S. L., SMITH, K. J., DWYER, T. & VENN, A. J. 2016. Worsening dietary and physical activity behaviours do not readily explain why smokers gain weight after cessation: a cohort study in young adults. *Nicotine & Tobacco Research,* DOI: 10.1093/ntr/ntw196.
TOBIAS, M., JACKSON, G., YEH, L.-C. & HUANG, K. 2007. Do healthy and unhealthy behaviours cluster in New Zealand? *Australian and New Zealand Journal of Public Health,* 31**,** 155-163*,* DOI: http://dx.doi.org/10.1111/j.1753-6405.2007.00034.x.
TOWNSEND, J., RODERICK, P. & COOPER, J. 1994. Cigarette smoking by socioeconomic group, sex, and age: effects of price, income, and health publicity. *BMJ,* 309**,** 923-927*,* DOI: 10.1136/bmj.309.6959.923.
TRAN, N. T., WILLIAMS, G. M., ALATI, R. & NAJMAN, J. M. 2015. Trajectories and predictors of alcohol consumption over 21 years of mothers′ reproductive life course. *SSM – Population Health,* 1**,** 40-47*,* DOI: http://dx.doi.org/10.1016/j.ssmph.2015.11.002.

TSENG, T.-S. & LIN, H.-Y. 2008. Gender and age disparity in health-related behaviors and behavioral patterns based on a National Survey of Taiwan. *International journal of behavioral medicine,* 15**,** 14-20.

TWYMAN, L., BONEVSKI, B., PAUL, C., BRYANT, J., WEST, R., SIAHPUSH, M., D'ESTE, C., OLDMEADOW, C. & PALAZZI, K. 2016. Factors Associated With Concurrent Tobacco Smoking and Heavy Alcohol Consumption Within a Socioeconomically Disadvantaged Australian Sample. *Substance Use & Misuse,* 51**,** 459-470*,* DOI: 10.3109/10826084.2015.1122065.

UNGER, J. B. 1996. Stages of change of smoking cessation: relationships with other health behaviors. *American journal of preventive medicine*.

VALLEJO-TORRES, L., HALE, D., MORRIS, S. & VINER, R. M. 2014. Income-related inequality in health and health-related behaviour: exploring the equalisation hypothesis. *Journal of Epidemiology and Community Health,* DOI: 10.1136/jech-2013-203306.

VAN DE MHEEN, H., STRONKS, K., LOOMAN, C. & MACKENBACH, J. 1998. Does childhood socioeconomic status influence adult health through behavioural factors? *International Journal of Epidemiology,* 27**,** 431-437*,* DOI: 10.1093/ije/27.3.431.

VAN DER VORST, H., ENGELS, R. C. M. E., MEEUS, W., DEKOVIĆ, M. & VAN LEEUWE, J. 2005. The role of alcohol-specific socialization in adolescents' drinking behaviour. *Addiction,* 100**,** 1464-1476*,* DOI: 10.1111/j.1360-0443.2005.01193.x.

VAN NIEUWENHUIJZEN, M., JUNGER, M., VELDERMAN, M. K., WIEFFERINK, K. H., PAULUSSEN, T. W. G. M., HOX, J. & REIJNEVELD, S. A. 2009. Clustering of health-compromising behavior and delinquency in adolescents and adults in the Dutch population. *Preventive Medicine,* 48**,** 572-578*,* DOI: http://dx.doi.org/10.1016/j.ypmed.2009.04.008.

VANDENBERG, B. & SHARMA, A. 2015. Are Alcohol Taxation and Pricing Policies Regressive? Product-Level Effects of a Specific Tax and a Minimum Unit Price for Alcohol. *Alcohol and Alcoholism,* DOI: 10.1093/alcalc/agv133.

VEDØY, T. F. 2014. Tracing the cigarette epidemic: An age-period-cohort study of education, gender and smoking using a pseudo-panel approach. *Social Science Research,* 48**,** 35-47*,* DOI: http://dx.doi.org/10.1016/j.ssresearch.2014.05.005.

VERGER, P., LIONS, C. & VENTELOU, B. 2009. Is depression associated with health risk-related behaviour clusters in adults? *The European Journal of Public Health,* 19**,** 618-624*,* DOI: http://dx.doi.org/10.1093/eurpub/ckp057.

VERMEULEN-SMIT, E., TEN HAVE, M., VAN LAAR, M. & DE GRAAF, R. 2015. Clustering of health risk behaviours and the relationship with mental disorders. *Journal of Affective Disorders,* 171**,** 111-119*,* DOI: http://dx.doi.org/10.1016/j.jad.2014.09.031.

VERMUNT, J. K. 2010. Latent Class Modeling with Covariates: Two Improved Three-Step Approaches. *Political Analysis,* 18**,** 450-469*,* DOI: 10.1093/pan/mpq025.

VERMUNT, J. K. & MAGIDSON, J. 2002. Latent Class Cluster Analysis. *In:* HAGENAARS, J. A. & MCCUTCHEON, A. L. (eds.) *Applied Latent Class Analysis.* Cambridge: Cambridge University Press. DOI: http://dx.doi.org/10.1017/CBO9780511499531.004.

VERPLAETSE, T. L. & MCKEE, S. A. 2017. An overview of alcohol and tobacco/nicotine interactions in the human laboratory. *The American Journal of Drug and Alcohol Abuse,* 43**,** 186-196*,* DOI: 10.1080/00952990.2016.1189927.

VON STUMM, S., DEARY, I. J. & HAGGER-JOHNSON, G. 2013. Life-course pathways to psychological distress: a cohort study. *BMJ open,* 3.

WADSWORTH, M. E. & KUH, D. J. 1997. Childhood influences on adult health: a review of recent work from the British 1946 national birth cohort study, the MRC National Survey of Health and Development. *Paediatric and perinatal epidemiology,* 11**,** 2-20.

WAKEFIELD, M. A., LOKEN, B. & HORNIK, R. C. 2010. Use of mass media campaigns to change health behaviour. *The Lancet,* 376**,** 1261-1271*,* DOI: http://dx.doi.org/10.1016/S0140-6736(10)60809-4.

WALLMANN-SPERLICH, B. & FROBOESE, I. 2014. Physical Activity during Work, Transport and Leisure in Germany – Prevalence and Socio-Demographic Correlates. *PLoS ONE,* 9**,** e112333*,* DOI: 10.1371/journal.pone.0112333.

WANG, J. & WANG, X. 2012. *Structural equation modeling: Applications using Mplus*, John Wiley & Sons.

WANNAMETHEE, S. G., SHAPER, A. G. & WALKER, M. 1998. Changes in physical activity, mortality, and incidence of coronary heart disease in older men. *The Lancet,* 351**,** 1603-1608.

WARDLE, J., HAASE, A., STEPTOE, A., NILLAPUN, M., JONWUTIWES, K. & BELLISIE, F. 2004. Gender differences in food choice: The contribution of health beliefs and dieting. *Annals of Behavioral Medicine,* 27**,** 107-116*,* DOI: 10.1207/s15324796abm2702_5.

WARNER, K. E., CHALOUPKA, F. J., COOK, P. J., MANNING, W. G., NEWHOUSE, J. P., NOVOTNY, T. E., SCHELLING, T. C. & TOWNSEND, J. 1995. Criteria for determining an optimal cigarette tax: the economist's perspective. *Tobacco Control,* 4**,** 380.

WATT, H. C., CARSON, C., LAWLOR, D. A., PATEL, R. & EBRAHIM, S. 2009. Influence of life course socioeconomic position on older women's health behaviors: findings from the British Women's Heart and Health Study. *American journal of public health,* 99**,** 320-327.

WATTS, P., BUCK, D., NETUVELI, G. & RENTON, A. 2015. Clustering of lifestyle risk behaviours among residents of forty deprived neighbourhoods in London: lessons for targeting public health interventions. *Journal of Public Health,* DOI: http://dx.doi.org/10.1093/pubmed/fdv028.

WEST, P. 1997. Health inequalities in the early years: is there equalisation in youth? *Social science & medicine,* 44**,** 833-858.

WEST, P., SWEETING, H. & YOUNG, R. 2010. Transition matters: pupils' experiences of the primary–secondary school transition in the West of Scotland and consequences for well-being and attainment. *Research Papers in Education,* 25**,** 21-50*,* DOI: 10.1080/02671520802308677.

WEYERS, S., DRAGANO, N., RICHTER, M. & BOSMA, H. 2010. How does socio economic position link to health behaviour? Sociological pathways and perspectives for health promotion. *Global Health Promotion,* 17**,** 25-33*,* DOI: 10.1177/1757975910365232.

WHICHELOW, M. J., ERZINCLIOGLU, S. W. & COX, B. 1991. A comparison of the diets of non-smokers and smokers. *British journal of addiction,* 86**,** 71-81.

WHITE, H. R., BRAY, B. C., FLEMING, C. B. & CATALANO, R. F. 2009. Transitions into and out of light and intermittent smoking during emerging adulthood. *Nicotine and Tobacco Research,* 11**,** 211-219*,* DOI: 10.1093/ntr/ntn017.

WHITE, J., REHKOPF, D. & MORTENSEN, L. H. 2016. Trends in Socioeconomic Inequalities in Body Mass Index, Underweight and Obesity among English Children, 2007–2008 to 2011–2012. *PLoS ONE,* 11**,** e0147614*,* DOI: 10.1371/journal.pone.0147614.

WHITLEY, E., BATTY, G. D., HUNT, K., POPHAM, F. & BENZEVAL, M. 2014. The role of health behaviours across the life course in the socioeconomic patterning of all-cause mortality: the west of Scotland twenty-07 prospective cohort study. *Annals of behavioral medicine,* 47**,** 148-157.

WHO 2014. Global Status Report on Non-Communicable Diseases 2014. Geneva: WHO, http://www.who.int/global-coordination-mechanism/publications/global-status-report-ncds-2014-eng.pdf.

WHO 2015. Guideline: Sugars intake for adults and children. Geneva: WHO, http://www.who.int/nutrition/publications/guidelines/sugars_intake/en/.

WICKRAMA, T. & WICKRAMA, K. 2010. Heterogeneity in adolescent depressive symptom trajectories: Implications for young adults' risky lifestyle. *Journal of Adolescent Health,* 47**,** 407-413.

WIIUM, N., BREIVIK, K. & WOLD, B. 2015. Growth Trajectories of Health Behaviors from Adolescence through Young Adulthood. *International Journal of Environmental Research and Public Health,* 12**,** 13711-13729*,* DOI: 10.3390/ijerph121113711.

WILLSON, A. E., SHUEY, K. M. & ELDER JR, G. H. 2007. Cumulative advantage processes as mechanisms of inequality in life course health. *American Journal of Sociology,* 112**,** 1886-1924.

WOOLCOTT, C. G., DISHMAN, R. K., MOTL, R. W., MATTHAI, C. H. & NIGG, C. R. 2013. Physical Activity and Fruit and Vegetable Intake: Correlations Between and Within Adults in a Longitudinal Multiethnic Cohort. *American Journal of Health Promotion,* 28**,** 71-79*,* DOI: 10.4278/ajhp.100917-QUAN-312.

WORSLEY, A., WANG, W. C. & HUNTER, W. 2012. The relationships between eating habits, smoking and alcohol consumption, and body mass index among baby boomers. *Appetite,* 58**,** 74-80*,* DOI: http://dx.doi.org/10.1016/j.appet.2011.09.003.

YANG, S. M., LYNCH, J., SCHULENBERG, J., ROUX, A. V. D. & RAGHUNATHAN, T. 2008. Emergence of socioeconomic inequalities in smoking and overweight and obesity in early adulthood: The National Longitudinal Study of Adolescent Health. *American Journal of Public Health,* 98**,** 468-477*,* DOI: 10.2105/ajph.2007.111609.

YANIV, G., ROSIN, O. & TOBOL, Y. 2009. Junk-food, home cooking, physical activity and obesity: The effect of the fat tax and the thin subsidy. *Journal of Public Economics,* 93**,** 823-830*,* DOI: http://dx.doi.org/10.1016/j.jpubeco.2009.02.004.

YEH, H.-W., ELLERBECK, E. F. & MAHNKEN, J. D. 2012. Simultaneous evaluation of abstinence and relapse using a Markov chain model in smokers enrolled in a two-year randomized trial. *BMC Medical Research Methodology,* 12**,** 95-95*,* DOI: 10.1186/1471-2288-12-95.

YOUNGE, J. O., KOUWENHOVEN-PASMOOIJ, T. A., FREAK-POLI, R., ROOS-HESSELINK, J. W. & MYRIAM HUNINK, M. G. 2015. Randomized study designs for lifestyle interventions: a tutorial. *International Journal of Epidemiology,* DOI: 10.1093/ije/dyv183.

YUSUFOV, M., PROCHASKA, J. O., PAIVA, A. L., ROSSI, J. S., BLISSMER, B., REDDING, C. A. & VELICER, W. F. 2016. Baseline Predictors of Singular Action Among Participants With Multiple Health Behavior Risks. *American Journal of Health Promotion,* 30**,** 365-373*,* DOI: 10.1177/0890117116646341.

ZACNY, J. P. 1989. Behavioral aspects of alcohol-tobacco interactions. *Recent developments in alcoholism: an official publication of the American Medical Society on Alcoholism, the Research Society on Alcoholism, and the National Council on Alcoholism,* 8**,** 205-219.