



A review of ethnicity measures for social research

Vernon Gayle
Roxanne Connelly
Paul Lambert

ABSTRACT

This working paper is a review of issues associated with measuring ethnicity and using ethnicity measures in social science research. The review is orientated towards researchers who undertake secondary analyses of large-scale micro-level social science datasets. The paper begins with an outline of two main approaches used in social surveys to measure ethnicity, the demographic perspective and the ethnic identity perspective. We provide examples of the ethnicity measures used in some of the UK's key social survey data resources. We discuss the importance of intersectionality in the analysis of ethnicity, because ethnicity measures are often associated with other important differences between individuals. We also describe approaches to the use of ethnicity measures in cross-national comparative research. We conclude that researchers should always use existing measures that have agreed upon and well documented standards, and which will facilitate comparability and replication. We also strongly advise researchers not to develop their own ethnicity measures without strong justification, or use existing measures in an un-prescribed or ad hoc manner. We advise that researchers should routinely undertake sensitivity analyses. We also encourage researchers to carefully consider the possible relationships between ethnicity and other important variables in order to avoid spurious interpretations of the effects of ethnicity.

KEYWORDS

Measuring Ethnicity; Social Surveys; Quantitative Data Analysis; UK.

EDITORIAL NOTE

Professor Vernon Gayle is Professor of Sociology and Social Statistics at the University of Edinburgh.

Dr Roxanne Connelly is a Research Fellow at the University of Edinburgh.

Professor Paul Lambert is Professor of Sociology at the University of Stirling.

Corresponding author: Professor Vernon Gayle, Vernon.Gayle@ed.ac.uk.

ACKNOWLEDGEMENTS

We gratefully acknowledge the comments of Professor Geoff Payne, Professor John Field and Dr Alasdair Rutherford.

- © Vernon Gayle, Roxanne Connelly and Paul Lambert all rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

ESRC Centre for Population Change

The ESRC Centre for Population Change (CPC) is a joint initiative between the Universities of Southampton, St Andrews, Edinburgh, Stirling, Strathclyde, in partnership with the Office for National Statistics (ONS) and the National Records of Scotland (NRS). The Centre is funded by the Economic and Social Research Council (ESRC) grant numbers RES-625-28-0001 and ES/K007394/1.

This working paper series publishes independent research, not always funded through the Centre. The views and opinions expressed by authors do not necessarily reflect those of the CPC, ESRC, ONS or NRS.

The Working Paper Series is edited by Teresa McGowan.

[Website](#) | [Email](#) | [Twitter](#) | [Facebook](#) | [Mendeley](#)

A REVIEW OF ETHNICITY MEASURES FOR SOCIAL RESEARCH

TABLE OF CONTENTS

1. INTRODUCTION	1
2. HOW IS ETHNICITY MEASURED?	2
2.1. THE DEMOGRAPHIC APPROACH.....	2
2.2. THE ETHNIC IDENTITY APPROACH.....	6
3. ETHNICITY DATA SOURCES	9
4. INTERSECTIONALITY	10
5. CROSS NATIONAL COMPARISONS	13
6. CONCLUSION	14
REFERENCES.....	16

1. INTRODUCTION

International migration has been a key driver of demographic and social change in the UK and in most other western nations over the last century (Coleman and Salt 1992, Castles and Miller 2009). European countries are predicted to become increasingly ethnically diverse in years to come with continued immigration, fluctuations in the relative size of different minority groups, and increasing numbers of people born from within ‘inter-ethnic’ unions (Coleman 2010, Coleman 2011). Many nationally representative large-scale social surveys collect information on ethnicity. A central aim of this working paper is to provide information for secondary data analysts who are not experts in the field of ethnicity.

Ethnicity is frequently taken to represent a self-claimed or subjective identity linked to a perception of shared ancestry as a result of some combination of nationality, history, cultural origins and possibly religion (Bulmer 1996, Platt 2007, Platt 2011). There is an extensive literature which discusses the meaning and use of the term ethnicity and how this concept differs and overlaps with the neighbouring concepts of race and national identity (for example see Smith 1991, Ratcliffe 1994, Mason 1995, Banton 1998, Cornell and Hartmann 2006, Burton, Nandi et al. 2008, Banton 2014). Ethnicity can sometimes explain substantial patterns of social inequality, and ethnicity is of relevance to a great many sociological enquiries (e.g. Tomlinson 1991, Platt 2005, Heath, Cheung et al. 2007, Platt 2007, Heath, Rethon et al. 2008). In the UK and many other countries, national statistics institutes have developed standardised measures to classify individuals into ethnic groups, often driven by legal requirements to evaluate anti-discrimination policies, the need to monitor change in the social and economic circumstances of ethnic minorities, and the need to have accurate information about the size and nature of the ethnic minority population (Bulmer 1985).

This paper describes the major ways in which ethnicity is measured in social survey research. Experts in the field of ethnicity research are well aware of the opportunities provided by the measures that are available in existing social surveys, as well as their limitations and complexities, but these issues are sometimes not appreciated by those secondary social survey analysts whose interests might lie

outside of ethnicity research. It is obvious that ethnic structures differ radically across nation states, therefore in order to ensure that this paper provides clear prescriptions we confine most of our discussion to the UK (although the issues which we cover are often equally relevant to data from other nations).

2. HOW IS ETHNICITY MEASURED?

Nandi and Platt (2012) describe two main perspectives in the measurement of ethnicity (see also Cornell and Hartmann 2006, Platt 2011). First, social survey data collections often place individuals into ethnic group categories. These are sometimes based upon a taxonomy that can be defined through objective characteristics (e.g. country of birth or parental country of birth). Alternatively, they may involve asking respondents (or even interviewers or proxy respondents) to allocate to an appropriate category on substantially subjective grounds. In either case, to many analysts, the specification of categorical ethnic group measures is a necessity for studying ethnic inequalities using social survey data (Platt 2011). Second, researchers may be interested in understanding the relative strength of an individual's ethnic identity or their feelings of belonging to a particular ethnic group (see Burton, Nandi et al. 2008). As a heuristic device we will refer to the first perspective as the demographic approach and the second perspective as the ethnic identity approach.

2.1. THE DEMOGRAPHIC APPROACH

In the demographic approach the aim is not to understand how people feel about their ethnic identity but rather to fit individuals into the most appropriate category based on relatively fixed characteristics (Burton, Nandi *et al.* 2008). These characteristics may include skin colour or nationality for example (Burton, Nandi *et al.* 2008, Burton, Nandi *et al.* 2010, Platt 2011). In the UK, most demographic measures of ethnicity used in surveys since the 1990's have incorporated subjectivity, and have used standard instruments that allow individuals to self-identify their ethnic group (Martin and Gerber 2006, Office for National Statistics 2013). The UK's tradition of producing social statistics about ethnicity based upon categories of subjective ethnic identity is close to that used in the United States, but is very different to that used in social surveys in many other countries (see Hoffmeyer-Zlotnik and Warner 2010).

In the UK, the Office for National Statistics recommends ethnic group measures for use in social surveys (see table 1). These measures are based on those designed for the UK decennial censuses. The measures are different in different UK territories, reflecting the nature of the minority ethnic population, different legal requirements, and variations in the prominent terminology used (Office for National Statistics 2013). The Office for National Statistics offer clear guidelines for researchers who are analysing ethnicity across the whole of the UK using these territory-specific measures. Their advice is that analysts should combine the different ethnic groups in an organised manner using higher level ethnicity categories to suitably combine groups (see Office for National Statistics 2013).

The ONS standardised measure also allows individuals to classify themselves as belonging to ‘other’ unspecified ethnic groups and to write in their chosen description. The use of this ‘write-in’ section has been shown to improve perceptions of the acceptability of this question to respondents, and to improve response rates (Office for National Statistics 2013). In line with the aims of the demographic approach, data collectors may seek to re-allocate those individuals who have written-in their ethnic identity to one of the existing ethnic group categories based on objective guidelines (see Office for National Statistics 2013).

England	Wales	Scotland	Northern Ireland
White <i>English; Welsh; Scottish; Northern Irish; British; Irish; Gypsy or Irish Traveller; Any other White background*</i>	White <i>Welsh; English; Scottish; North Irish; British; Irish; Gypsy or Irish Traveller; Any other White background*</i>	White <i>Scottish, Other British, Irish, Gypsy/Traveller, Polish, Any other White group*</i>	White <i>White; Irish Traveller</i>
Mixed <i>White and Black Caribbean; White and Black African; White and Asian; Any other Mixed background*</i>	Mixed <i>White and Black Caribbean; White and Black African; White and Asian; Any other Mixed background*</i>	Mixed <i>Any Mixed or Multiple ethnic groups*</i>	Mixed <i>White and Black Caribbean; White and Black African; White and Asian; Any other Mixed/Multiple ethnic group*</i>
Asian or Asian British <i>Indian; Pakistani; Bangladeshi; Chinese; Any other Asian background*</i>	Asian or Asian British <i>Indian; Pakistani; Bangladeshi; Chinese; Any other Asian background*</i>	Asian or Asian British <i>Pakistani; Pakistani Scottish or Pakistani British; Indian; Indian Scottish or Indian British; Bangladeshi; Bangladeshi Scottish or Bangladeshi British; Chinese; Chinese Scottish or Chinese British; Any other Asian*</i>	Asian or Asian British <i>Indian; Pakistani; Bangladeshi; Chinese; Any other Asian background*</i>
Black/African/Caribbean /Black British <i>African; Caribbean; Any other Black/African/Caribbean background*</i>	Black/African/Caribbean /Black British <i>African; Caribbean; Any other Black/African/Caribbean background*</i>	African <i>African; African Scottish or African British; Any other African*</i>	Black/African/Caribbean /Black British <i>African; Caribbean; Any other Black/African/Caribbean background*</i>
		Caribbean or Black <i>Caribbean; Caribbean Scottish or Caribbean British; Black; Black Scottish or Black British; Any other Caribbean or Black*</i>	
Other ethnic group <i>Arab; Any other ethnic group*</i>	Other ethnic group <i>Arab; Any other ethnic group*</i>	Other ethnic group <i>Arab; Arab Scottish or Arab British; Any other ethnic group*</i>	Other ethnic group <i>Arab; Any other ethnic group*</i>

Table 1: The Office for National Statistics harmonised ethnicity measure recommended for use in social surveys (Office for National Statistics 2013).

Note: *If respondents answered 'other' they are given the opportunity to 'write-in' their preferred description of their ethnicity.

These standardised measures have been adopted in a number of large scale social surveys (for a review see Afkhami 2012). For example, in the first wave of the UK Millennium Cohort Study¹ (Institute of Education 2012) the ONS ethnicity measure is provided separately for cohort members living in England, Wales, Scotland and Northern Ireland². Additionally, harmonised UK level ethnicity measures are also provided in 6-category, 8-category and 11-category³ forms. The use of these existing standardised measures will facilitate the comparability of analyses across studies and therefore the use of these standardised measures is encouraged. For many social science investigations using large-scale social survey data the use of these measures within a multivariate analysis, along with other key variables, will fulfil the requirement of providing increased control for the demographic composition of the UK.

A common complication in UK survey research arises when analyses are conducted using non-standard ethnic group categories (i.e. taxonomies that don't correspond with those recommended by the ONS). This sometimes arises when an ethnicity measure is not available in the survey data in the standard format. However, in the case of analysis using large-scale secondary surveys, it is more likely to arise because a decision has been made to re-code the measure of ethnicity into a different taxonomy. Generally speaking we would argue that researchers should resist the impulse to generate derived and variant measures when standardised taxonomies are available (see Lambert, Gayle et al. 2009). The justification for using variant measures may sometimes be ad hoc and insubstantial, and there is an obvious challenge to comparability when using non-standard measurements. Nevertheless, if standardised measures are not available, and/or there are operational reasons that preclude the use of standard categories (e.g. very low numbers of cases in some relevant groups, or an analytical focus on an ethnic division that is not captured by the

¹ For more details see: <http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=851>

² In the first wave of the Millennium Cohort Study the ONS territory specific ethnicity measures for cohort members are: adcea00 (for cohort member 1), adceeb00 (for cohort member 2) and adceec00 (for cohort member 3) for those living in England; adcewa00 (for cohort member 1), adcewb00 (for cohort member 2) and adcewc00 (for cohort member 3) for those living in Wales; adcesa00 (for cohort member 1), adcesb00 (for cohort member 2) and adcesc00 (for cohort member 3) for those living in Scotland; and adcena00 (for cohort member 1), adcenb00 (for cohort member 2) and adrenc00 (for cohort member 3) for those living in Northern Ireland.

³ In the first wave of the Millennium Cohort Study the 6-category UK harmonised ethnicity measures is adc06e00, the 8-category version is adc08e00, and the 11-category version is adc11e00.

standard measure), two further principles should apply. First, researchers should construct variant measures that, as closely as possible, conform to published standards that have been recommended by statistical authorities (e.g. the ONS). Second, researchers should provide clear justification for, and documentation and metadata about, how the measures were derived, in order that other researchers can test the measures and replicate results⁴.

2.2. THE ETHNIC IDENTITY APPROACH

Measuring the extent of an individual's ethnic identity is especially pertinent when investigations require more insight into how individuals view their own ethnic identity, what characteristics make up this identity, or how important this identity is to the individual (Burton, Nandi *et al.* 2008). For example, researchers might be interested in how an ethnic identity is formed and how this changes over the lifecourse (Phinney 1989, Phinney 1990, Phinney and Alipuria 1990, Torres 2003, French, Seidman *et al.* 2006). Similarly investigating how ethnic identity relates to wellbeing (Phinney, Horenczyk *et al.* 2001, Mossakowski 2003), or how self-esteem is influenced by ethnic identity (Bracey, Bamaca *et al.* 2004, Umaña-Taylor 2004, Umaña-Taylor and Shin 2007) will probably require measures which capture more nuanced aspects of ethnic identity.

It has been argued that asking survey respondents to identify themselves on the basis of a single, mutually exclusive, category may overlook some important dimensions of ethnic identity. This is because ethnicity is a multi-dimensional concept which includes a number of elements (e.g. ancestry, national identity, religion and country of birth) (Aspinall 2011). There is some evidence that the importance of different dimensions of ethnic identity may vary across groups. For example, when providing descriptions of their ethnicity in free-text responses, 'Black' groups in the 1991 and 2001 censuses were found to emphasise their national identity (i.e. being British) as a central element of their ethnic identity (Office for National Statistics

⁴ Between 2008 and 2013, the 'GEMDE' service (www.dames.org.uk/gemde) provided an online archive of metadata about ethnicity classifications which was designed specifically to support this purpose – for instance, allowing a researcher to deposit the software code that they used to construct a variant ethnicity measure on a specific survey.

2006). South Asian groups however have been found to emphasise their religion as a central element of their ethnic identity (Modood, Beishon *et al.* 1994).

The use of multiple measures within a social survey, which examine different aspects of the ethnicity concept, offers an effective solution to the measurement of ethnic identity in more detail, since multiple responses across differing characteristics allow respondents to have more control over the expression of various elements of their identity (Burton, Nandi *et al.* 2010). The measures developed for the United Kingdom Household Longitudinal Study (UKHLS, *Understanding Society*⁵) are a prime example of this approach. In addition to measuring ethnicity using standardised demographic measures, described above, a suite of measures were also developed to provide information on various aspects of ethnicity (e.g. country of birth, language, nationality, religion), and also the extent to which the respondents felt these characteristics were important to their identity. Nandi and Platt (2012) developed these questions to offer a comprehensive set of measures suitable for the specialist and detailed study of ethnic identity in the UK. The use of multiple measures has also been advocated as a solution to comparability in cross-national survey research (see Hoffmeyer-Zlotnik and Warner 2010).

In many analyses a standardised categorical ethnicity measure, for example the ONS measure described above, will be sufficient when included in a multivariate analysis with other key variables as this will provide suitable statistical control and represent the demographic structure of the UK. When the focus of the analysis is estimating differences between the major ethnic groups which make up the UK, a standardised categorical ethnicity measure may also be well suited within a multivariate analysis. Indeed, despite their recognition of the relevance of multiple subjective identities, Burton *et al.* (2008) point out that just as in most situations it is possible to study social class inequalities without the need to measure class consciousness, therefore ethnic group inequalities can be studied using social survey data without capturing an individual's conscious ethnic identity.

⁵ For more details see: <https://www.understandingsociety.ac.uk/>.

In other analyses which focus on ethnic experiences, using a standardised categorical ethnicity measure may be less optimal. There are sometimes circumstances where the analytical focus on ethnic inequalities involves important divisions that are not well captured by the standard ethnicity variables. Well documented examples in the UK include the importance of immigrant ‘generation’ to analyses of social inequality (e.g. Heath, Cheung *et al.* 2007), or the distinctiveness of ‘East African Asians’ in the UK in the late twentieth century (e.g. Modood, Berthoud *et al.* 1997). More generally, authors such as Yancy *et al.* (1976) and more recently Aspinall (2012) have suggested that ethnicity is a changeable, complex and multi-dimensional concept which cannot be measured by allocating individuals to a single ethnic group category. Amongst other things, these researchers anticipate the substantial challenges of using ethnicity measures in comparative longitudinal studies, especially because the standard measures designed for the UK Census are revised decennially. In the case of longitudinal comparisons across different time points, equivalent measures might never be feasible. Even if measures are feasible, they may have only ‘nominal equivalence’ and lack ‘functional equivalence’, because the relative meaning of being in a certain category is not the same from one time point to another.

There are some published recommendations of good practice for using standard ethnicity categories for longitudinal comparisons (see Platt, Simpson *et al.* 2005, Simpson and Akinwale 2007, Afkhami 2012). It is often the case that previous comparative studies provide the most appropriate benchmarks for further endeavours. Additionally, in the special case when longitudinal data allows for records to be linked for the same people at different time points, it is possible to derive consistent ethnicity measures by focussing upon the measures used at a particular agreed time point (see Platt 2005). Special cases aside, it is widely accepted that the consistent use of data on ethnicity in survey research is a substantial challenge. In situations when standard measures cannot be used, it is valuable for researchers to clearly describe and document the measurement approach that is adopted.

3. ETHNICITY DATA SOURCES

A notable obstacle for analyses of ethnicity using some social survey datasets is small subsamples. This situation will typically arise in small surveys but can also arise in relatively large surveys because of the fairly small proportion of respondents from minority ethnic groups. Even in large-scale nationally representative surveys the coverage of some smaller ethnic groups can be representative but correspondingly still small. A practical solution which some secondary data analysts adopt when analysing data with a low number of ethnic minority respondents is to aggregate ethnicity groups. For example, it might be beguiling to aggregate data from Indian, Pakistani, Bangladeshi and other Asian respondents if numbers were low in a social survey. In some instances this strategy might provide a practicable solution and do little or no harm to the substantive results. In other analyses there may be important differences between these ethnic groups and such aggregation may result in misleading substantive conclusions. Dale (2008) for example illustrate that Indian women's employment patterns are very different from other women who might reasonably be included in an aggregate 'South Asian' category (Dale 2008). The effects of such aggregations of ethnic categories cannot be easily anticipated or judged *a priori*, therefore we advise that secondary social survey data analysts place suitable thought into the characteristics of the respondents that they are combining within their analysis. To justify aggregations, it is also good practice if analysts undertake detailed exploratory analyses and make these available to the research community.

There are other alternative solutions to dealing with low numbers of ethnic minority survey respondents. These include restricting analytical focus to only the ethnic group or groups which are well represented in the data to hand (see Dale, Lindley *et al.* 2006), or assigning index scores on the basis of external information linked to the ethnic group categories (see Lambert 2005). In certain appropriate circumstances these might be better solutions than completely ignoring ethnicity in the analysis. Care must be taken by data analysts not to introduce confusion or to limit comparability with other studies. We also advocate that any approaches that are used are clearly documented to aid replication and meta-analyses.

In the absence of alternative solutions, researchers interested in ethnicity are usually best advised to utilise datasets which have larger numbers of ethnic minority respondents. UK Census related data resources such as the Sample of Anonymised Records and the ONS Longitudinal Study of England and Wales include data on a large number of individuals from ethnic minority groups. Larger sample surveys such as the Labour Force Surveys also have good coverage. A number of large-scale social surveys have been developed to include ‘boost’ samples of ethnic minorities (e.g. the United Kingdom Household Longitudinal Study and the Millennium Cohort Study). A boost sample comprises an additional set of data collected from specific sub-groups of the survey population.

Survey data has also been collected specifically for the study of the experiences of ethnicity minority populations in the UK (e.g. the Fourth National Survey Ethnic Minorities and the Ethnic Minority British Electoral Survey). These specialised surveys have been designed to collect data from larger samples of ethnic minority individuals. Whilst these surveys have supported detailed and informative research on ethnic minorities in the UK, they differ from the more general (or omnibus) large-scale social surveys that are designed to support multi-purpose secondary data analysis.

4. INTERSECTIONALITY

Alongside definitional issues, an important consideration in most analyses of ethnicity is the way in which ethnic categories are linked to other important differences between people. In multivariate survey analyses, secondary data researchers may be primarily interested in estimating the net effects of ethnicity whilst controlling for the effects of other variables. There are sometimes sizeable correlations between ethnicity and other demographic factors. Platt (2011) provides an approachable review of this issue and, in our view, usefully deploys the term intersectionality to describe the interaction between categories of difference in people’s lives, for example ethnicity, gender and social class (see also Davis 2008). The key point is that a misleading understanding of the influence of ethnicity might emerge if measures are not considered in the context of such correlations (see for example Heath and Martin 2013). In the light of such evidence, secondary data analysts should place a suitable

amount of thought into which key variables and demographic measures they include within models that contain measures of ethnicity, drawing upon extensive theoretically guided exploratory data analyses.

One important element of intersectionality is the strong cohort demographic patterns that are linked to ethnic minority groups due to concentrated waves of immigration (Fryer 1984, Spencer 1997, Panayi 1999, Hansen 2000). Substantial age and regional settlement differences arise between ethnic groups that reflect their immigrant-cohort background. Certain socio-economic measures, such as income, and most outcomes related to health, can vary substantially according to age. Because certain ethnic groups have younger age structures than others there is a pressing need to control for these differences in age when analysing data (for details on the age profiles of ethnic groups in the UK see Haskey 1996, Scott, Pearce *et al.* 2001). In statistical modelling approaches, additional controls for the main effects of age may often be adequate. It is also plausible that an ageing process itself may vary by ethnic group. We advise that survey data analysts should routinely explore interactions between age and ethnicity.

Immigration is another issue that is relevant to studying ethnicity. Individual circumstances and conditions associated with immigration might also be consequential. For example whether or not individuals are born in the country of residence, how long their family has lived in the country, their proficiency in the host country language, and their participation in the education system, may be relevant to some analyses. These patterns may affect important outcomes such as chances and choices in the labour market.

In the social sciences, there is an analytical literature on the differences between the experiences and outcomes of immigrants and those born in the host society (e.g. Raftery, Jones *et al.* 1990, Schnepf 2007, Levels, Dronkers *et al.* 2008, Algan, Dustmann *et al.* 2010). One common approach is to characterise survey respondents into different 'immigrant generations'. Typically the 'first generation' of immigrants are those born abroad, the 'second generation' are those born in the host society whose parents were born abroad, and other categories are occasionally identified such as the 'third' and 'subsequent generations'. In some cases researchers

also define the '1.5 generation', which are those who were born abroad but moved to the host society as young children and had the bulk of their schooling in the host country (e.g. Allensworth 1997, Harklau, Losey *et al.* 1999, Quirke, Potter *et al.* 2010).

In the UK and in many other nations, some ethnic minority groups exhibit sustained differences in patterns of family formation and living arrangements in comparison with other groups. Most well-known in the UK are patterns of larger household size, and younger average ages of marriage and family formation, amongst South Asian ethnic minority groups (e.g. Coleman and Salt 1996). Differences in family patterns are often presumed to reflect the influence of cultural heritage, but there are also complexities to demographic differences between ethnic groups which defy a characterisation based on cultural heritage alone (e.g. Shaw 2014). Family circumstances do have a substantial influence on many other social outcomes within the UK, including employment arrangements, socio-economic circumstances, and patterns of lifestyle and social support. The association between ethnicity and average variations in family circumstances is another aspect of intersectionality. It might often be unsatisfactory therefore to analyse differences by ethnic group without considering other variations in average family circumstances.

There are pronounced ethnic differences in settlement patterns within countries such as the UK (Ratcliffe 1997, Finney and Simpson 2009). Compared with other areas of the UK some urban areas in London, the North East and the North West of England, and the English Midlands have much higher ethnic minority populations. The extent and consequences of regional segregation are sometimes exaggerated (Finney and Simpson 2009). From a survey data analytical perspective there is a persuasive argument for considering geographical issues when studying patterns of ethnic difference. Conventionally, higher level regional measures can be entered into analyses to attempt to separate geographic from ethnicity effects. In some applications, different regional settlement patterns by ethnic groups in the UK are thought to occur because of differential industrial and employment sector opportunities (see Iganski and Payne 1999). In this scenario, an appropriate response to the recognition of regional settlement differences may be to include measures of

industrial sector or employment circumstances in analyses rather than crude measures of geography.

5. CROSS NATIONAL COMPARISONS

The contemporary social sciences are characterised by a great deal of interest in conducting cross-nationally comparative analyses with survey data (see Hoffmeyer-Zlotnik and Wolf 2003, Hoffmeyer-Zlotnik, Harkness *et al.* 2005). In principle the analysis of ethnicity is an intrinsically international theme, given its relationship with international migration. One difficulty arises because different countries have different histories of immigration, meaning that it is unlikely that comparative studies of ethnicity will be dealing with comparable minority groups between countries. A strategy sometimes followed is to identify and compare minorities from the same origin background in different countries (e.g. Crul and Vermeulen 2003, Model 2005). This can still be unsatisfactory because, even in this scenario, it is unlikely that the differences between migrants from the same nation who settled in different countries are independent of other factors. An alternative comparative strategy is simply to study different ethnic minority groups in different countries, and make only carefully qualified comparisons (e.g. Heath, Cheung *et al.* 2007).

Some reviews have suggested ways for specific measurement instruments to be applied consistently to facilitate cross-national comparisons (e.g. Hoffmeyer-Zlotnik 2003, Lambert 2005, Aspinall 2007). This literature confronts a second major challenge for cross-national comparative research on ethnicity, namely that different nations have strong traditions of difference in the measures related to ethnicity that they collect in surveys. In many countries the measurement of ethnicity is highly politicised, and an approach used in one nation might never be considered acceptable in another. To a limited extent, recent social survey instruments are beginning to negotiate this problem, usually by collecting data on multiple measures that could be subsequently adapted to the researcher's needs. Hoffmeyer-Zlotnik and Warner (2010) encourage survey designers to collect measures of characteristics such as citizenship, country of birth, parents' country of birth, and language regardless of national traditions in measuring ethnicity.

6. CONCLUSION

The ethnic composition of the UK like many other western industrialised nations has changed radically as a result of migration, and there is persuasive evidence that most European countries will become increasingly ethnically diverse. Most nationally representative large-scale social surveys collect information on ethnicity due to the central role that it plays in stratifying contemporary societies. We have provided an account of how ethnicity is measured in social surveys, and reflected on the relationship between ethnicity and other social inequalities. Our overall message is that researchers undertaking secondary data analyses should be aware of the scope and challenges that relate to using ethnicity measures as key variables in analyses.

We conclude that researchers should always use existing measures that have agreed upon and well documented standards, and which will facilitate comparability and replication. We also strongly advise researchers not to develop their own ethnicity measures without strong justification, or use existing measures in an un-prescribed or *ad hoc* manner. We have highlighted that general social surveys often contain categorical measures of ethnicity based on the demographic approach. For many sociological research questions these measures will be adequate, especially when all that is required is improved control for underlying patterns in the data that are linked to ethnicity. Our stern advice is that researchers should employ these standard measures.

We have highlighted the potential impact of intersectionality. When estimating statistical models that include ethnicity variables we advise researchers to be cognisant of the potential associations between ethnicity and other key variables. This may help to better elucidate important patterns in social outcomes. If unrecognised such undetected associations can lead to spurious interpretations of ethnicity effects. Secondary data analysts are therefore strongly directed to include ethnicity measures along with other key variables in multivariate analyses. Contemporary data analysis software packages also make it relatively straightforward to estimate and test interaction effects between ethnicity and other key variables.

In another recent working paper we have advocated undertaking sensitivity analyses when analysing occupational measures with social survey data (Gayle,

Connelly *et al.* 2015). In this present working paper we similarly conclude that undertaking sensitivity analyses should be a routine aspect of social survey data analysis. The process of conducting a sensitivity analyses can seem burdensome and even uninspiring, but we are argue that modern software capabilities mean that it is now relatively easy to re-run analyses using different measures. Undertaking sensitivity analyses is of considerable benefit to social science more generally because it increases rigour, and therefore allows more confident interpretation of the results. We hope that this paper provides succinct information on the foundations of existing ethnicity measures in social surveys. In addition we have attempted to provide practical advice that will make a genuine contribution to how existing ethnicity measures in social surveys can be better incorporated into social science analyses.

REFERENCES

- Afkhami, R.** (2012). "Ethnicity: Introductory User Guide." Colchester, Essex, Economic and Social Data Service.
- Algan, Y., Dustmann, C., Glitz, A. and Manning, A.** (2010). The Economic Situation of First and Second-Generation Immigrants in France, Germany and the United Kingdom. *The Economic Journal*, 120 (542), F4-F30.
- Allensworth, E. M.** (1997). "Earnings mobility of first and "1.5" generation Mexican-origin women and men: A comparison with US-born Mexican Americans and non-Hispanic Whites. *International Migration Review*, 386-410.
- Aspinall, P.J.** (2007). Approaches to developing an improved cross-national understanding of concepts and terms relating to ethnicity and race. *International Sociology*, 22 (1), 41-70.
- Aspinall, P.J.** (2011). The utility and validity for public health of ethnicity categorization in the 1991, 2001 and 2011 British Censuses. *Public Health*, 125, 680-687.
- Aspinall, P.J.** (2012). Answer Formats in British Census and Survey Ethnicity Questions: Does Open Response Better Capture 'Superdiversity'? *Sociology*, 46 (2), 354-364.
- Banton, M.** (1998). "Racial theories." Cambridge University Press.
- Banton, M.** (2014). "Ethnic and racial consciousness." Routledge.
- Bracey, J.R., Bamaca, M.Y. and Umana-Taylor, A.J.** (2004). Examining ethnic identity and self-esteem among biracial and monoracial adolescents. *Journal of youth and adolescence*, 33 (2), 123-132.
- Bulmer, M.** (1985). A controversial census topic: race and ethnicity in the British census. *Journal of official statistics*, 2 (4), 471-480.
- Bulmer, M.** (1996). "The ethnic group question in the 1991 census of population." Coleman, J. and Salt, J. *Ethnicity in the 1991 Census, Vol. 1*. London, HMSO, 33-62.
- Burton, J., Nandi, A. and Platt, L.** (2008). Who are the UK's minority ethnic groups? Issues of identification and measurement in a longitudinal study. *ISER Working Paper Series*.
- Burton, J., Nandi, A. and Platt, L.** (2010). Measuring ethnicity: challenges and opportunities for survey research. *Ethnic and Racial Studies*, 33 (8), 1332-1349.
- Castles, S. and Miller, M.J.** (2009). "The Age of Migration: International Population Movements in the Modern World." Basingstoke, Palgrave.
- Coleman, D.** (2010). Projections of the ethnic minority populations of the United Kingdom 2006–2056. *Population and Development Review*, 36 (3), 441-486.
- Coleman, D.** (2011). "Immigration, Population and Ethnicity: The UK in International Perspective." Population.
- Coleman, D. and Salt, J.** (1992). "The British population: patterns trends and processes." Oxford University Press.
- Coleman, D.A. and Salt, J.** (1996). "Ethnicity in the 1991 Census, Volume 1. Demographic characteristics of the ethnic minority population." London, HMSO.
- Cornell, S. and Hartmann, D.** (2006). "Ethnicity and race: Making identities in a changing world." Sage Publications.
- Crul, M. and Vermeulen, H.** (2003). The Second Generation in Europe. *International Migration Review*, 37 (4), 965-986.
- Dale, A.** (2008). Migration, marriage and employment amongst Indian, Pakistani and Bangladeshi residents in the UK. *CCSR Working Paper 2008-02*. University of Manchester, Cathie Marsh Centre for Census and Survey Research.
- Dale, A., Lindley, J. and Dex, S.** (2006). A life-course perspective on ethnic differences in women's economic activity in Britain. *European Sociological Review*, 22 (3), 323-337.
- Davis, K.** (2008). Intersectionality as buzzword: A sociology of science perspective on what makes a feminist theory successful. *Feminist theory*, 9 (1), 67-85.

- Finney, N. and Simpson, L.** (2009). "Sleepwalking to segregation'? Challenging myths about race and migration." Bristol, Policy Press.
- French, S.E., Seidman, E., Allen, L. and Aber, J.L.** (2006). The development of ethnic identity during adolescence. *Developmental psychology*, 42 (1), 1.
- Fryer, P.** (1984). "Staying Power: The history of black people in Britain." London, Pluto Press.
- Gayle, V., Connelly, R. and Lambert, P.** (2015). A review of occupation-based social classifications for social research. *CPC Working Paper 60*. ESRC Centre for Population Change, UK.
- Hansen, R.** (2000). "Citizenship and Immigration in Post-war Britain." Oxford, Oxford University Press.
- Harklau, L., Losey, K.M. and Siegal, M.** (1999). "Generation 1.5 meets college composition: Issues in the teaching of writing to US-educated learners of ESL." Routledge.
- Haskey, J.** (1996). The ethnic minority populations of Great Britain: their estimated sizes and age profiles. *Population Trends*, 84, 33-39.
- Heath, A., Cheung, S.Y. and Smith, S.W.** (2007). "Unequal Chances: Ethnic Minorities in Western Labour Markets." Oxford, Oxford University Press.
- Heath, A. and Martin, J.** (2013). Can religious affiliation explain 'ethnic' inequalities in the labour market? *Ethnic and racial studies*, 36 (6), 1005-1027.
- Heath, A.F., Rethon, C. and Kilpi, E.** (2008). The second generation in Western Europe: Education, unemployment, and occupational attainment. *Annual Review of Sociology*, 34, 211-235.
- Hoffmeyer-Zlotnik, J.H.P.** (2003). "How to Measure Race and Ethnicity." In Hoffmeyer-Zlotnik, J.H.P. and Wolf, C. *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variable*. New York, Kluwer Academic: 267-277.
- Hoffmeyer-Zlotnik, J.H.P., Harkness, J.A. and Zentrum für Umfragen** (2005). "Methodological aspects in cross-national research." Zentrum für Umfragen, Methoden und Analysen (ZUMA), Mannheim.
- Hoffmeyer-Zlotnik, J.H.P. and Warner, U.** (2010). The concept of ethnicity and its operationalisation in cross-national social surveys. *Metodoloski zvezki*, 7(2), 107-132.
- Hoffmeyer-Zlotnik, J.H.P. and Wolf, C.** (2003). "Advances in cross-national comparison: A European working book for demographic and socio-economic variables." Springer Science & Business Media.
- Iganski, P. and G. Payne** (1999). Socio-economic re-structuring and employment: the case of minority ethnic groups. *The British journal of sociology*, 50 (2), 195-215.
- Institute of Education** (2012). "Millennium Cohort Study: First Survey 2001-2003." SN4683. UK Data Archive, Colchester, Essex.
- Lambert, P.S.** (2005). "Ethnicity and the Comparative Analysis of Contemporary Survey Data." In Hoffmeyer-Zlotnik, J.H.P. and Harkness, S. *Methodological Aspects in Cross-National Research*. Mannheim, ZUMA, 259-277.
- Lambert, P.S., Gayle, V., Bowes, A., Blum, J.M., Jones, S.B., Sinnott, R.O., Tan, K.L.L., Turner, K.J. and Warner, G.C.** (2009). "Standards setting when standardizing categorical data." International Conference on e-Social Science, Cologne.
- Levels, M., Dronkers, J. and Kraaykamp, G.** (2008). Immigrant children's educational achievement in western countries: origin, destination, and community effects on mathematical performance. *American Sociological Review*, 73 (5), 835-853.
- Martin, E.A. and Gerber, E.** (2006). "Methodological influences on comparability of race measurements: Several cautionary examples." *Survey Methodology*, 6.
- Mason, D.** (1995). "Race and ethnicity in modern Britain." Oxford University Press.
- Model, S.** (2005). "Non-White Origins, Anglo Destinations: Immigrants in the US and Britain." In Loury, G.C., Modood, T. and Teles, M. *Ethnicity, Social Mobility and Public Policy in*

- the United States and United Kingdom*. Cambridge, Cambridge University Press, 363-392.
- Modood, T., Beishon, S. and Virdee, S.** (1994). "Changing ethnic identities." London, Policy Studies Institute.
- Modood, T., Berthoud, R., Lakey, J., Nazroo, J.Y., Smith, P., Virdee, S. and Beishon, S.** (1997). "Ethnic minorities in Britain: diversity and disadvantage." London, Policy studies Institute.
- Mossakowski, K.N.** (2003). Coping with perceived discrimination: does ethnic identity protect mental health? *Journal of health and social behavior*, 318-331.
- Nandi, A. and Platt, L.** 2012). Developing ethnic identity questions for Understanding Society. *Longitudinal and Life Course Studies*, 3 (1), 80-100.
- Office for National Statistics** (2006). "A guide to comparing 1991 and 2001 Census ethnic group data." London, ONS.
- Office for National Statistics** (2013). "Ethnic Group: Harmonised Concepts and Questions for Social Data Sources." London, Office for National Statistics.
- Panayi, P.** (1999). "The impact of immigration: A documentary history of the effects and experiences of immigrants in Britain since 1945." Manchester, Manchester University Press.
- Phinney, J.S.** (1989). Stages of ethnic identity development in minority group adolescents. *The Journal of Early Adolescence*, 9 (1-2), 34-49.
- Phinney, J.S.** (1990). Ethnic identity in adolescents and adults: review of research. *Psychological bulletin*, 108 (3), 499.
- Phinney, J.S. and Alipuria, L.L.** (1990). Ethnic identity in college students from four ethnic groups. *Journal of adolescence*, 13 (2), 171-183.
- Phinney, J.S., Horenczyk, G., Liebkind, K. and Vedder, P.** (2001). Ethnic identity, immigration, and well-being: An interactional perspective. *Journal of social issues*, 57 (3), 493-510.
- Platt, L.** (2005). "Migration and Social Mobility: The Life Chances of Britain's Minority Ethnic Communities" Bristol, The Policy Press.
- Platt, L.** (2007). "Poverty and Ethnicity in the UK." Policy Press.
- Platt, L.** (2011). "Understanding inequalities: Stratification and difference." Polity.
- Platt, L., Simpson, L. and Akinwale, B.** (2005). Stability and change in ethnic groups in England and Wales. *Population Trends*, 121, 35.
- Quirke, E., Potter, R. and Conway, D.** (2010). Transnationalism and the Caribbean Community in the UK: Theoretical Perspectives. *The Open Geography Journal*, 3, 1-14.
- Raftery, J., Jones, D.R. and Rosato, M.** (1990). The mortality of first and second generation Irish immigrants in the UK. *Social Science & Medicine*, 31 (5), 577-584.
- Ratcliffe, P.** (1994). "'Race', ethnicity and nation: international perspectives on social conflict." UCL Press.
- Ratcliffe, P.** (1997). "Ethnicity in the 1991 Census, Volume 3: Social Geography and Ethnicity in Britain: Geographical spread, spatial concentration and internal migration." London, HMSO.
- Schnepf, S.V.** (2007). Immigrants' educational disadvantage: an examination across ten countries and three surveys. *Journal of Population Economics*, 20 (3), 527-545.
- Scott, A., Pearce, D. and Goldblatt, P.** (2001). The sizes and characteristics of the minority ethnic populations of Great Britain - latest estimates. *Population Trends*, 105, 6-15.
- Shaw, A.** (2014). "Kinship and continuity: Pakistani families in Britain." Routledge.
- Simpson, L. and Akinwale, B.** (2007). Quantifying stability and change in ethnic group. *Journal of Official Statistics*, 23 (2), 185.
- Smith, A.D.** (1991). "National identity." University of Nevada Press.

- Spencer, I.R.G.** (1997). "British Immigration Policy Since 1939: The making of multi-racial Britain." London, Routledge.
- Tomlinson, S.** (1991). Ethnicity and educational attainment in England: An overview. *Anthropology & education quarterly*, 22 (2), 121-139.
- Torres, V.** (2003). Influences on ethnic identity development of Latino college students in the first two years of college. *Journal of College Student Development*, 44 (4), 532-547.
- Umaña-Taylor, A.J.** (2004). Ethnic identity and self-esteem: Examining the role of social context. *Journal of adolescence*, 27 (2), 139-146.
- Umaña-Taylor, A.J. and Shin, N.** (2007). An examination of ethnic identity and self-esteem with diverse populations: Exploring variation by ethnicity and geography. *Cultural Diversity and Ethnic Minority Psychology*, 13 (2), 178.
- Yancey, W.L., Ericksen, E.P. and Juliani, R.N.** (1976). Emergent ethnicity: A review and reformulation. *American Sociological Review*, 391-403.

ESRC Centre for Population Change
 Building 58, Room 2001
 Faculty of Social and Human Sciences
 University of Southampton
 SO17 1BJ

T: +44 (0)2380 592579

E: cpc@soton.ac.uk

www.cpc.ac.uk

To subscribe to the CPC newsletter and keep up-to-date with research activity, news and events, please register online:
www.cpc.ac.uk/newsletter

For our latest research updates you can also follow CPC on Twitter, Facebook and Mendeley:



www.facebook.com/CPCpopulation



www.twitter.com/CPCpopulation



www.mendeley.com/groups/3241781/centre-for-population-change

The ESRC Centre for Population Change (CPC) is a joint initiative between the University of Southampton and a consortium of Scottish universities including St Andrews, Edinburgh, Stirling and Strathclyde, in partnership with the Office for National Statistics and National Records of Scotland.

