Who are the UK's minority ethnic groups? Issues of identification and measurement in a longitudinal study

Jonathan Burton
Alita Nandi
Lucinda Platt

Institute for Social and Economic Research
University of Essex

INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH

ESRC
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

Non-technical summary

Research into ethnicity in the UK is of substantial research and policy interest. A new multi-topic panel survey of around 40,000 households, *Understanding Society: The UK Household Longitudinal Study*, is being developed which has ethnicity as one of its core concerns. Therefore, *Understanding Society* will not only include questions on ethnicity and ethnic identity but will also carry a wide range of questions on issues of relevance to research on the UK's ethnic groups. In addition to the large sample size and its longitudinal design, the survey will also include an over-sample of selected ethnic minority groups to enable research on ethnicity related issues. In order to realise the major potential for ethnicity research that *Understanding Society* offers requires that we resolve conceptual and measurement issues over what we mean by 'ethnicity' and which groups should be included in the over-sample.

In this paper we draw on a wide ranging review of the literature, the results from an intensive consultation exercise, the impact of current policy concerns and priorities, existing practice and development work already carried out for *Understanding Society* and in its Innovation Panel to establish the precise nature of these issues and proposed solutions

The concept of ethnicity can be approached in a number of ways – as commonalities within a group or as differences from 'other' groups. Its multifaceted nature makes its measurement using a single measure almost impossible. Additionally, there is a lack of consensus among researchers about what is to be measured: while social psychologists are interested in understanding individuals' identity, other researchers and policy makers are interested in social stratification where group memberships are seen as shaping group members' outcomes and resulting in different life courses. We conclude that for measuring ethnic identity in *Understanding Society* one way forward is to use a range of measures to capture different dimensions of ethnicity and ethnic group identity.

Developing an ethnic self-identification (categorical) question is a major challenge. In addition to standard survey design issues such as question wording, the design of such a question requires a fine balancing act between trying to get consistent, reliable measures of ethnic identity and capturing people's perception of their own ethnic identity. The response categories need to be meaningful, acceptable and consistent, so that people will be able to respond to them in expected, and predictable, ways. We conclude that in addition to multiple single measures it would be necessary to develop a multiple response question to understand the combinations of identifications that people incorporated in their self-identities and the relative importance attached to these various dimensions.

To capture aspects of ethnic belonging and association it would also be important to ensure a range of general question domains in such areas as social networks, participation, attitudes to own and other groups across the survey.

Different issues arise in relation to the composition of the ethnic minority oversample and the screening question to be used. The screen question is intended to include a number of specific groups that have some internal coherence and are sufficiently numerous for meaningful analysis. Thus its operationalisation is driven by pragmatic rather than analytic concerns.

Overall, *Understanding Society* aims to furnish researchers with a larger repertoire of ethnicity related questions to enable them to carry out analysis according to the framework of their discipline and research questions.

Who are the UK's minority ethnic groups? Issues of identification and measurement in a longitudinal survey.

Jonathan Burton, Alita Nandi, Lucinda Platt
University of Essex

August, 2008

Abstract
In this paper we highlight issues related to measuring ethnicity and ethnic identity. We base our discussion on an extensive review of the literature and an intensive consultation process undertaken as part of the development of the ethnicity focused strand of a major new UK panel study, *Understanding Society*. We conclude that ethnic identity is a multi-dimensional concept and its ideal measure would have to be consistent, reliable as well as capture people's perception of their own ethnic identity. One way forward is to design a multiple response question with different dimensions of ethnicity as response options.

Contact: Alita Nandi, ISER, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK; +44(0) 1206-874699; email anandi@essex.ac.uk

1. Introduction

Research into ethnicity in the UK, both issues of identity, ethnic diversity, and differences between ethnic groups is of substantial research and policy interest. This is attested to by, on the one hand, the extensive analytical literature exploring issues such as the demography of the UK and how it is changing, the differences in life chances across groups, and forms of identification and association, and, on the other hand, the pervasiveness of policy discourses around social cohesion and integration, (in)-equality of employment and other outcomes and points of commonality and tension within the UK population. Nevertheless, there has been no UK survey dedicated to the study of issues concerned with ethnicity and ethnic diversity since the *Fourth National Survey of Ethnic Minorities* was carried out in England and Wales in 1994; and there has never been a panel survey which has explicitly facilitated the measurement of issues concerning ethnicity and differences between ethnic groups. This will change with the development of the new, major panel survey for the UK, *Understanding Society: The UK Household Longitudinal* Study.

This multi-topic survey is ambitious in a number of respects. It offers, through its large sample size (initially 40,000 households), unprecedented opportunities for analysis of different sub-samples/sub-populations including those of the four countries of the UK and different ethnic groups. As a longitudinal survey of households it will enable analysis of change and stability of individuals' lives over time and in relation to other household members. It aims to encourage and enable analysis across disciplines, in particular extending the potential for analysis and users of the data beyond the economists and sociologists who have traditionally made the greatest use of comparable studies, such as the British Household Panel Survey, a UK panel survey that has been running since 1991 and which will be incorporated into *Understanding Society*. This interdisciplinary aspect will be achieved in part through the range of questions asked and also by the collection of different sorts of data than verbal responses to a face-to-face interview, such as linked administrative and geocoded data and, in the future, bio-measures.

What particularly sets this study apart is that it puts ethnicity at the heart of the study. That is, it explicitly attempts to enable research on ethnicity in a longitudinal context.

Researchers will anyway be able to analyse sub-populations over time and across a range of topic areas, given the features of large overall sample sizes and the wide range of directly asked and matched data. But the survey will also explicitly address research and policy interest in ethnicity and ethnic minority groups in three ways:

1. By including questionnaire content about all respondents' ethnicity / ethnic identity

2. By containing questionnaire content or information from other sources about issues of relevance to research on the UK's ethnic groups, including on issues of group formation etc., or on areas of difference between ethnic groups that invite research.

3. By means of an oversample of selected ethnic groups that enables analysis of these groups separately or jointly.

Together we refer to these, which cover both content and design aspects of the UKHLS, as the 'ethnicity strand'.

Delivering across these three areas to enable the full potential of the study for longitudinal analysis does, however, raise a range of conceptual and measurement issues for consideration. What do we mean by ethnicity or ethnic group? What are the most informative questions to ask people about their ethnicity or ethnic identification? What are the most relevant topic areas to the study of ethnicity or ethnic group differences and similarities? What are the topics that are most appropriate to a household panel survey? Which groups should be included in the over-sample and why? How do groups defined for the purpose of the oversample relate to measures of ethnicity asked of the survey sample as a whole?

In this paper we draw on a wide ranging review of the literature, the results from an intensive consultation exercise, the impact of current policy concerns and priorities, existing practice and development work on questions and selection of groups already carried out in the *Understanding Society* Innovation Panel (an experimental sample designed to explore methodological issues in advance of and separately from the main panel) to set out the issues we have faced in relation to these questions relating concepts, measurement and implementation. We discuss how we are aiming to meet the challenges they pose and deliver on the 'ethnicity strand' so that *Understanding*

Society can, indeed, provide the research resource anticipated for the study of ethnicity and ethnic groups within the UK.

The paper is structured as follows. In the next section (Section 2) we reflect on what is meant by ethnicity and ethnic group and also relate that to other related concepts that are used alongside those of ethnicity, as complementary to them or, in other contexts (especially other national contexts), as alternatives or corrections to them, such as race, national identity, country of birth, immigrant status, nationality and religion. We then consider the measurement issues implied by these questions alongside the measurement paradigms already in existence and policy or monitoring imperatives. We explore what single response categorical 'ethnic group' questions can and cannot tell us and reflect on the multiple domains that may constitute an individual's identity, how we can acknowledge these and how we can grasp their relative importance to the individual concerned. For example, a respondent might see themselves simultaneously as Scottish (upbringing, language/accent, politics, local affiliation), Black (politically used to express solidarity with minority, non-white experience), Asian (to suggest antecedents (parents /grandparents or earlier generations from South Asia) and British (to express nationality, citizenship etc.). All these might be important to her, but some aspects might be more important, for example, she may consider the most critical element of her identity to be her 'Scottishness'. Additionally, the relative importance could easily vary with context: whether she is in Glasgow or London, whether she is experiencing harassment, whether she is voting, whether she is visiting her parents or being interviewed by a non-Asian interviewer etc. How we both provide robust measures of relevant characteristics for interrogation and analysis and acknowledge the multiplicity and complexity of owned identity and its development is one of the main measurement challenges for the Ethnicity Strand.

In the light of the discussion of issues of definition and measurement, we then consider (Section 3) the practical issues involved in developing measures of ethnic group for *Understanding Society*, and the solutions we have developed for implementation in the survey and testing. We go on to outline the process and conclusions about the related, but distinct, issue of which 'groups' should be included

in the oversample, and thus be susceptible to more detailed analysis than for minorities sampled across the survey as a whole.

In Section 4, we explore the question of what sort of additional topic areas are most pertinent to the study of ethnicity and ethnic group differences and similarities, and to our understanding of different dimensions of 'ethnicity' itself. These include areas that can allow us to reflect on what being a member of group actually comprises – for example patterns of friendship and association, relationship to 'countries of origin' and attitudes to other 'groups', as well as areas that have shown themselves to be of substantive interest in revealing substantial – and often unexplained, in a statistical sense – differences between ethnic groups, such as educational outcomes, patterns of social mobility, income and poverty. We include consideration in this section of these areas which are also of substantial policy concern and interest. In the final section (Section 5), we draw some brief conclusions.

Section 2: Measuring ethnicity and ethnic groups
2.1 What is ethnicity and what are ethnic groups?
Max Weber, in the early decades of the 20th Century identified defining features of ethnic groups when he wrote:

> We shall call "ethnic groups" those human groups that entertain a subjective belief in their common descent because of similarities of physical type or of customs or both, or because of memories of colonization and migration; this belief must be important for the propagation of group formation; conversely, it does not matter whether or not an objective blood relationship exists. Ethnic membership (*Gemeinsamkeit*) differs from the kinship group precisely by being a presumed identity, not a group with concrete social action, like the latter (Weber 1978:389).

This notion of an identification with common descent, real or imagined, has been utilised in subsequent attempts to define ethnicity; for example, Schermerhorn (1978: 12) defined ethnic group as

> A collective within a larger society having real or putative common ancestry, memories of a shared historical past and a cultural focus on one of more symbolic

elements defined as the epitome of their peoplehood. Examples of such symbolic elements are: kinship patterns, physical contiguity (as in localism or sectionalism), religious affiliation, language or dialect forms, tribal affiliation, nationality, phenotypical features, or any combination of these. A necessary accompaniment is some consciousness of kind among members of the group.[1]

And Cornell and Hartmann (1998), summarising the different possible definitions that have been put forward, conclude by drawing on both Weber and Schermerhorn and stress the importance both of common claims which do not need to be founded in fact and the self-consciousness of the claims. The importance of perception and imagining of antecedents is also core to Anderson's (1991) notion of 'imagined communities'.

As well as defining ethnicity by "a sense of belonging", others have defined groups relatively, *i.e.,* in relation to how they construct difference from others and establish boundaries for the groups (Wimmer 2008). This is a perspective that is associated with Barth's work (Barth 1969); but has also been stressed by the acculturation framework for understanding ethnic identity, which focuses on the juxtaposition or co-existence of two cultures. This juxtaposition makes the concept of ethnic identity meaningful in each. In a society with only one ethnic group, ethnic group identity is meaningless. This has implications for a longitudinal study such as *Understanding Society* because the more integrated a multi-ethnic society gets, the less meaningful the ethnic group identities and group boundaries become. In a more general sense this framework highlights the importance and meaning of group boundaries. Studies looking at ethnic identity in this framework try to understand how persons in a multicultural society, especially with a majority and minority group, choose to identify with an ethnic group (Nekby and Rodin 2007). Some perceive a person's ethnic identification as a linear model i.e., a stronger identification with one group necessitates weaker identification with the other. Others view it as a bimodal model where the strength of a person's identification with the majority and minority ethnic groups are independent of each other.

---

[1] In his subsequent account of 'ethnicity', Schermerhorn did not qualify group memories as real or putative, but it can be assumed that this remains implicit in the reference to (more subjective) 'memories; rather than ancestry' (Schermerhorn 1974).

When we are looking, then, to capture ethnic identity through survey measurement, it swiftly becomes apparent that there are many dimensions bound up in the concept of ethnic groups, including differentiation from others as well as positive association with commonalities, including commonalties across areas that are often considered distinct topics, such as language and religion. Moreover, among those commonalties not all have to be present for an ethnic group – and ethnic identification – to take place.

An issue which vexes the definition and measurement of ethnic groups, and to which we will return, is that there is a lack of consensus among researchers about what is to be measured and why. Broadly speaking there are two interests that get confounded in attempts to articulate ethnicity in a survey, and in resulting analysis. While social psychologists are interested in understanding individuals' identity (and they view ethnicity as part of one's social identity) other researchers and policy makers are interested in social stratification where group memberships is seen as shaping group members' outcomes and resulting in different life courses.

Social psychologists are interested in questions such as how self-esteem of the members of an ethnic group is affected by identification with that group, how identification with a minority ethnic group is affected by the majority ethnic group in that particular society and, on an individual level, they are interested in how a person's own ethnic identity is formed. These studies mainly look at ethnic identity within three main frameworks – social identity theory, acculturation and identity formation (Phinney 1990). In addition, individual ethnic identification is strongly limited by external forces that shape the options, feasibility and attractiveness of various ethnicities (Nagel 1994). These forces themselves are incorporated into models of identity development. Researchers identify three general stages of this identity formation. The first stage is when the person has not yet started thinking of her own ethnic identity or just accepts that of her parents. The second stage is when she starts to investigate it and the final stage is when she arrives at a conclusion about her ethnic identity. Even if the last stage is reached by the time a person reaches adulthood, the different stages may be revisited later.

If these are the sorts of questions and frameworks that drive psychologists' interests in ethnicity, other researchers are far more concerned with understanding the nature of society and population; and also whether there are groups within society that can be shown to be faring poorly or to be discriminated against – and the fate of these groups over time. Such questions are also the primary concern of the policy framework within which patterns of measurement are formed and which in their turn contribute to common-sense understandings of ethnicity and ethnic groups. Governments and the measurement systems they implement in relation to minority groups are driven by concerns other than interest in people's identity and subjective perceptions of their ethnicity and ethnic antecedents.

In the UK, legal attempts to reduce discrimination and disadvantage also require that monitoring takes place to measure the effects of anti-discriminatory policies. The 1976 Race Relations Act and the 2000 Race Relations (Amendment) Act not only attempted to make discrimination illegal but also gave public authorities the duty to monitor policy and service delivery for different ethnic groups. The Acts encouraged authorities to collect statistical data on ethnicity and the likely impact of policies and services on different ethnic groups. For example, schools are required to assess the impact of their policies on ethnic minority pupils, staff and parents. Much of this information is collected through administrative data which also collects ethnic identity as part of the process. To monitor change in the social and economic circumstances of ethnic minorities it is essential for local and national government to have accurate information about the size of the ethnic minority population (Bulmer 1986). The UK census classification, then, aims to provide data to monitor equal access to housing, education, employment for groups who had historically been discriminated against and who are recognised for the purposes of race relations legislation. The situation is similar in the US where the impetus behind the race and ethnicity questions on the census is "for civil rights and other compliance reporting from the public and private sectors and all levels of government" (Anderson and Fienberg, 1999). This compliance reporting emerged in the 1970s from the federal government's responsibilities to enforce civil rights law.

Subjectivity and identity may be questions of great research interest, and a subjective response may be the only way to obtain the information that government seeks.

Nevertheless, official agencies are rarely concerned with how people 'feel' about themselves, but instead are interested in collecting data which helps them understand populations, trends in population characteristics and experiences across subgroups of the population including ethnic (or in US terms, racial) groups. From the perspective of social psychologists and those concerned with identity and identity development, these sociological, policy-oriented and demographic investigations employ ethnicity, ethnic origin or ethnic identity as interchangeable terms to measure a multi-dimensional and fluid concept as if it were a uni-dimensional, fixed and stable concept; which is perhaps the key problem of such categorical ethnic identity measures (Aspinall 2001, Yancey, Ericksen and Juliani 1976).

These differences in approach and research purpose make the task of establishing 'good measures' of ethnic group a complex one: variability in response to measures by the same individual can be seen as either of intrinsic interest or a problem of the instrument depending on one's perspective.

One way forward is to use a range of measures to capture different definitions of ethnicity and ethnic group, though, as the ensuing discussion shows, the attempt to separate subjective from objective measures is never fully possible – the two must always be in a symbiotic relationship with one another, given that both identity and constructions of difference are formed in a social context. Nevertheless, there remains substantial scope for disentangling the different meanings or research purposes attributed to ethnic identity, and identifying them separately for measurement.

Phinney (1990) attempts to unravel some of the contestation over the different understandings and 'uses' of ethnicity by unpacking the composite concept and its meaning. She identifies five components, one or more of which different studies implicitly refer to when citing ethnicity/ethnic group / ethnic identity. The first of these components she calls *ethnicity* by which she means a person's heritage, parents' ethnicity, country of origin. Secondly, people *self-identify* as a member of an ethnic group which may be different from their ethnicity. Third is *ethnic belonging* which refers to a sense of belonging to the self-identified ethnic group/s. *Ethnic involvement* is all about the actual participation and practice of being a member of an ethnic group, such as the language spoken, the ethnicity of friends and social networks,

participation in ethnic social groups, cultural practices, etc. Finally, there is *ethnic attitude* which refers to one's feeling (positive or negative) towards one's self-identified ethnic group. These are useful distinctions to keep in mind when discussing methodological issues of measuring ethnicity/ethnic group identity. The analytical perspective can help us understand which one (or more) of these component/s is intended for particular research purposes and consequently how it would best be captured. Whether it is actually possible to 'unpack' ethnicity at all is debated (McKenzie 1998). Nevertheless, McKenzie himself then goes on to say that "unpacking could be of great importance if it is used to produce multi-dimensional instead of categorical representations of culture and ethnicity."

Phinney's first and second components are in general what ethnic group questions set out to or are assumed to measure, though there is some overlap between the second and the third components. We focus on the components of antecedents, self-identity and, to a lesser extent, sense of belonging in this section, discussing the development of an ethnic group question or questions. The component of *involvement* might be susceptible to direct questioning but is arguably best measured by exploring people's practices and social networks more generally. We return to this issue in Section 4. In addition, the question of *ethnic attitude* can probably best be linked to general perceptions on groups in society and on self-esteem measures, questions which have an interest beyond the ethnicity strand and which speak to the cross-disciplinary emphasis of *Understanding Society*.

Phinney's two components of ethnic identity that we treat in this section (*ethnicity* and *self-identity*) could be seen to correspond broadly to the interests in demographic and social facts on the one hand and in identification and identity formation on the other, that we have already outlined. Nevertheless, though the subdivision of what Phinney calls 'ethnicity' might give some indications of how to measure it in a stable framework, in practice a clear separation between the objectiveness of *ethnicit*y and the subjectivity of *self-identity* is not readily achieved, as past practice shows us.

For example, an emphasis on the centrality of subjectivity and identification has been incorporated into the standard accounts of how ethnicity should be measured in censuses and surveys (Martin and Gerber 2006; Office for National Statistics 2003),

and is consistently reasserted. Yet despite the emphasis placed on this aspect of the measurement process (along with the corresponding acceptance that there will be change in self-identification and change in meaning of categories over time), and despite the fact that it would be impossible to consider any other way of ascertaining people's ethnic group other than by asking them, subjective questions in censuses are primarily concerned with attempting to achieve stable, if not entirely objective, measurement. Or, in Manuel's (2006) terms, such questions incorporate high levels of 'confusion' about both the purpose of obtaining ethnic group information and what answers are intended to elicit.

The stress on subjectivity is in tension with the concern for reliability, stability and standardisation evident in official approaches, and the aims of collecting demographic and monitoring information. The subjective approach in census questions also risks disguising the fact that the primary aim of the data collection is rather blunt and is less concerned with 'ethnicity' itself than with other criteria that may or may not be a component of ethnic identity – specifically 'colour'. For example, in the UK's first attempt to ask ethnic group in a census, in 1991, the acknowledgement that what was being collected was subjective and the claim that it measured ethnic group, while simultaneously emphasising that what was of interest was the non-white population (Coleman and Salt 1996; Salt 1996), led to extensive criticism (Ballard 1996; Ballard 1997; Ratcliffe 1996).

There are of course a whole set of assumptions about who is non-white and how they can be allocated into 'groups', which draws more on concepts and process of ascription than of identity. And the very establishment of a discrete set of categories within which people are expected to fit themselves is itself an on-going process of ascription and management of boundaries (Barth 1969). The relevance of a priori assumptions in relation to what constitutes an 'ethnic minority group' for the purposes of census classification can further be seen in the way in which 'write-in' answers are reallocated to the main categories on offer (Office for National Statistics 2003). What is at stake then, in these cases is not understanding people's identity or the extent to which they have a particular 'race consciousness' (themselves conceptually very different as Manuel (2006) points out); but finding a set of categories that meet official conceptions of ethnic group and that are sufficiently meaningful or

comprehensible that people will be able to respond to them in expected – and predictable – ways.

This does not mean that questions of identification, consciousness and 'imagined community' cease to be of interest, but makes it unlikely that the categories put to people to ascertain their ethnic group in a standard way will be informative, in and of themselves, about other components. Just as class stratification can be demonstrated without class consciousness being required, so ethnic group stratification can be revealed without the underlying assumption that these groupings are meaningful in and of themselves to those who select them. Conversely, it is possible to identify strongly with a particular group or ethnicity without participating in the average or general experience of members of that group as a whole. However, while it may be possible to move towards better ways of capturing subjective identities, the very creation of official categories may have implications for responses to questions which attempt to capture different aspects of ethnicity. We go on to discuss such measurement implications below.

First, however, we briefly outline the relationship of ethnicity and ethnic group to other concepts that are used sometimes as complementary, sometimes as equivalent and sometimes as alternative framings of 'difference', depending not only on the elements of 'ethnicity' that are being considered, but also according to national context and research orientation. If ethnicity has two main components, one concerned with stable characteristics associated with heritage, ancestry or country of origin, and the other concerned with identification and the development of identity, it also overlaps with a number of other concepts that are often implicit in discussions of ethnicity or which are offered as alternative ways of understanding social reality and social difference. It is worth outlining these briefly, both for clarity and to enable consideration of the extent to which research interests do truly correspond even when using different terminology. Additionally, to the extent that some measures such as religion or national origin can be considered potential components of ethnicity, they allow us to consider how competing definitions can be resourced from a 'portfolio' of measures that can be used separately or in combination to represent or explore different understandings of 'ethnicity'. These related concepts are: *race'; national identity; parentage; nationality; religion and language*.

2.2 Related concepts

2.2.1 Race

The concept of race has been shown to have no biological basis and it is widely accepted that 'races' do not exist. As Cornell and Hartmann (1998:23) put it, 'most contemporary scholars dismiss the entire idea of race as a meaningful biological category that can be applied to separate groups of human beings' (See also (1996:5)).Yet the language of race persists widely and many of the ideas associated with racial ideology continue to find forms of expression and are subject to serious discussion (Herrnstein and Murray 1994). Banton (1998) has discussed the ways that the language of race has been used historically and the development of ideas of racial types, often associated with the work of Gobineau and the rise of the eugenics movement. He cogently and forcefully argues that since such ideas have been discredited and have no scientific basis, sociologists should eschew the language of race. 'Folk concepts' he argues have no place in sociological discourse or theorising and to attempt to become too tied up with everyday language is to inhibit the development of more appropriate ways of understanding and theorising forms of social relations, in some of which physical differences play a prominent role in defining group boundaries and inclusiveness and exclusiveness. Despite this, however, practical problems remain in achieving a stringent use of theoretical and conceptual framework when not only everyday language – and the political structures and prescriptions that go with that (such as 'Race Relations' acts) – but also academic discourse itself, particularly in the US, uses that language of 'race' as unambiguously meaningful and self-evident. For example, the US Office of Management and Budget defines the census categories as "a socio-political construct designed for collecting data on the race and ethnicity of broad population groups in this country, and are not anthropologically or scientifically based" (quoted in Anderson and Finberg 1999). Thus, despite an emphatic conceptual rejection of 'race', it retains common currency in both policy and academic literatures, which is heightened when researchers attempt to look comparatively across countries using a common terminology.

While some persist in rejecting the language of race (Gilroy 2000; Miles 1993), not only does its pervasiveness make this difficult, but some commentators would see the attempted rejection of the language of race as misguided and also as missing the

primary importance of 'race' as an organizing principle within societies. As Loury (2004) writes "to establish the scientific invalidity of racial taxonomy demonstrates neither the irrationality nor the *immorality* of adhering to a social convention of racial classification…. [T]he social convention of thinking about other people and about ourselves as belonging to different 'races' is such a longstanding and deeply ingrained one in our political culture that it has taken on a life of its own" (p.76) (see also Mason 1995). In the UK, 'race' continues to be used relatively freely among the policy community and the general public, while within the academic community there has emerged a tendency to hedge the question by using the formulation 'race and ethnicity', in which the combination of the two terms implies not so much complementary concepts as a conflation of alternative terminologies, which are assumed to approximate to similar social facts.

To the extent that there is distinctiveness, race is more generally associated with inherited characteristics and specifically skin colour, and ethnicity with 'cultural' characteristics. This itself, however, is not a clear distinction: it can be impossible to distinguish physical difference at the margins and distinctiveness is also context specific, depending itself on what is defined as a 'group' and the perceived salience of such a group. For example, in the UK the focus on 'ethnic minorities' for official purposes is on those from (former) commonwealth countries. It is these which are effectively racialised through forms of recognition in classification systems (such as the census) that focus on 'non-white minorities', through particular language to describe them and through expectations of their existence. Other groups, such as those of Turkish origin, which might be considered 'non-white' or racialised in other contexts or countries, are not racialised in that way in the UK – presenting ambiguities about whether indeed they constitute a non-white ethnic minority for purposes of classificatory systems. (In the UK Census of 2001 about half of those who were born in Turkey ticked the 'white other' box and most of the rest ticked the 'any other ethnic group' box.)

Moreover, distinctiveness can be associated with forms of dress or religious practice – and indeed with language or accent – that are not necessarily considered any more separable from 'identity' than skin colour to the individuals concerned. In Canada, racialised groups are referred to as 'visible minorities'. Visibility is perhaps helpful

here in that it implies more than black–white differences as visibility can reside in a range of attributes. It also potentially enables a stress on the extent to which 'visibility' comes into being through the eye of the beholder, the ways in which groups are given meaning and significance and are considered to be bounded, and the ways in which they may then claim that group status for political or mobilising purposes.

Interests in 'race' thus incorporate interests in who is non-white – which by definition implies particular constructions of 'whiteness' and of dominant 'English' ethnicity (Bonnet 1998, Young 1990, 2008), in being Black as both a political and classificatory category, in the racialisation of particular groups which will be context (and generation) specific and in ancestry or perceived geographical antecedents.

2.2.2 National identity

For some, national and ethnic identity are closely connected concepts with many of the same defining features as ethnicity (Smith 1991). For others, national identity is regarded as a substitute for more 'primordial' ties or ethnic associations. National identity can be expressed either as a perceived or actual link to a particular geographic region or to a 'nation' defined as a collectivity. It thus has many overlaps with concepts of subjectively identified ethnicity, with perception of historical antecedents playing an important role. Language (and/or accent or dialect) may also be an important component of national identity in some contexts or formulations. In the UK context, national identity can also refer to particular association with one of the four countries of the UK: England, Wales, Scotland or Northern Ireland. Thus national identity can offer one dimension of belonging that does not necessarily presuppose other aspects of ethnic identity, but implies the possibility of multiple identities – or multiple dimensions – for example to see oneself as Scottish (upbringing, language/accent, politics, local affiliation) does not preclude also seeing oneself as Black and/or Indian and/or British.

2.2.3 Parentage, ancestry and heritage

As noted, ethnicity and ethnic group is linked to parentage or ancestry in a number of ways. Definitions of ethnic group ancestry – whether real or putative – can be an important element of understanding and identification. Thus the physical location of

origins is often a crucial way of understanding oneself and commonalities and differences. Moreover, as with analysis of immigrants and immigration, it is the direct fact of someone – or their parents – coming into one country from another one that is often regarded as of critical importance in recognising and understanding difference, and in accounting for differences in life chances and outcomes (Massey 1993). In addition, given the concern with visible difference in many categorical systems and in accounts of discrimination, it is often the heritability of characteristics that is deemed critical to ethnicity and to minority status. Thus there are clearly overlaps between notions of parentage, ancestry or heritage and those of ethnicity. There are also links to common conceptualisations of 'race' and to national identity, as discussed above. However, there are also criticisms of using such concepts or measures to define groups. It has been argued that they place too much weight on 'blood' and biology rather than recognising the social context in which ethnic identity is formed and reformed. Such criticisms treat questions on such issues as if they are not subject to the same level of subjective (re)-interpretation as any ethnic group question. But if self-identification is a subjective process that can potentially vary, then arguably the same could be true of attribution to parents or antecedents. Nevertheless, it is indisputable that measures are likely to be more 'stable' than self-identity measures.

2.2.4 Immigration status, citizenship, nationality

There is a very large research agenda around immigration and the 'immigrant' that is often confounded with considerations of ethnicity. This is due to the fact that many minority groups are defined through their 'origins' in a different country (though clearly this is not the case for indigenous minorities). It also presupposes that the fact of immigration tends towards some commonalities across immigrant groups regardless of their country of origin. Much of the discussion of immigrants and immigration has been informed by the assimilation paradigm (Duncan and Lieberson 1959) that has been a dominant analytical framework in much U.S. research and has been applied to other contexts in research as well. While some commentators have regarded assimilation models as no longer pertinent to current patterns of immigration and settlement, there is still a substantial body of research supporting the retention of an assimilation framework in some form for analysing the outcomes of immigrants and their descendants (see for example, Alba and Nee 2003; Chiswick 2006).

Assimilation frameworks are concerned with the adaptation of the immigrant to the patterns of experience and norms of existence in the receiving country. Thus, they emphasise the potential for the loss of distinctive ethnicity or ethnic attributes, and hence stress a fluid and contextual understanding of ethnicity. On the other hand, as much of the concern with assimilation is to do with change over time there is an emphasis on the oxymoronic construct of the 'second generation immigrant', which conversely implies the importance of descent – even if to reveal that the children of immigrants are different from their parents.

In some European countries (certain) immigrants have been defined as 'foreigners', relating insider or outsider status to notions of 'ethnic belonging', which almost paradoxically can transcend national boundaries. Thus, in Germany 'ethnic Germans' from other countries had a different status to immigrants from, for example, Turkey. Legal status and citizenship can then reinforce such distinctions, and such 'foreignness' can be passed down through generations, in situations where citizenship criteria do not follow from birth. Where 'foreigners' are also regarded as 'visible' by the dominant society, the transmission of foreignness in legal terms is conflated with notions of transmission of visible characteristics, with such visibility being viewed by construction as the primary definition of a 'foreigner'. The importance of context, national history and patterns of immigration to particular countries also informs notions of which immigrants get to be seen as 'foreign' or translated into 'ethnic minorities', though in research into immigration this is often not made explicit. Research into 'immigration' is thus concerned with identifying not only own but also parents' (and possibly grandparents') country of birth, alongside some measure of racialisation or of ethnicity.

2.2.5 Religious affiliation

Increasing attention is now paid to religious affiliation and religious differences in populations (Brierley 2000; Purdam et al. 2007). While religion has been a long-standing source of discrimination, differential rights and life chances and of cleavage and conflict, recently much greater attention has been paid to the religion of ethnic minorities or new immigrants, and particularly to Muslims rather than to long-standing differences (such as those between Catholics and Protestants). This has been facilitated in the UK by the incorporation of a religion question into the Census

(Dobbs et al. 2006), itself in part a response to the need to monitor and to identify discrimination on religious grounds as required by legislation.

In the British case law that has accompanied the 30 years of the 1976 Race Relations Act (and its Amendment in 2000), many groups have been defined as, or as not, coming under the Act. While religion itself was not originally covered by anti-discrimination legislation, some groups defined according to religion (notably Jews and Sikhs) were considered as coming under its remit through their constitution as cultural/ethnic groups. Here we see the ways in which conventional accounts of religion as a matter of individual conscience do not necessarily match on to the reality of ethno-religious groups bound by common expressions of religious observance, group belonging or cultural commonalities that are closely linked to the religious groups' histories, and responses to them over time, or to the 'visibility' of expressions of religious belonging or of religion as ancestry. It is clear that for those who actively practise, the ethnic and the religious are often not clearly separable, with places of worship often being distinguished along ethnic / country of origin / language lines as well as via the actual religion or denomination (Salway et al. 2007).

Religion and religious activity can then be of interest in its own right, but can also be seen as an important dimension of ethnicity and the cultural realisation of ethnic identity.

2.2.6 Language

Similarly, (first) language or mother tongue is of independent interest to a large number of researchers in relation to its influence on employment prospects and experience. Fluency is one issue which is seen as relevant to labour market success (Chiswick 2008; Dustmann and Fabbri 2003), but accent, regardless of fluency has also been seen to be a potential source of discrimination (and this could apply to regional accents as well). In addition, studies of interviews have shown how speakers of other languages may bring different conventions about 'decorum' or how to express themselves appropriate to particular contexts to such situations, and thus may not be able to 'deliver' in the ways that the conventions of interviews expect (Roberts and Campbell 2006).

However, beyond such substantive issues concerned with linguistic practice itself, language can also form an important component of ethnic or national identity. Language and language use can be crucial to how people seem themselves, how they identify (Leslie and Lindley 2001; Shields and Wheatley Price 2001;) and how they form and sustain social networks – and thus give 'groupness' its meaning. It is through speech that people often are recognised as 'foreign' regardless of other 'markers' such as name or appearance. First language can thus be an important determinant of 'visibility' even if it is not picked up by classification systems.

2.3 Measurement issues

As the previous section made clear, at the heart of the measurement issue is a distinction that is often not expressed between interest in understanding people's sense of self and the desire to accurately capture an assumed social and demographic reality. In the first case, variation and instability are not only anticipated in the definition of ethnicity but become of interest in their own right, and are in some cases associated with particular life stages or generalised identity development processes. The fluidity of the very concept of ethnic identity results in these so-called 'inconsistencies'. In the latter case, the emphasis is on reliability of measures and being able to relate them adequately to existing policy and research questions and constructs. Instability in measures in terms of high rates of test-retest inconsistency are problematic and are only of interest to the extent of highlighting the need for alternatives.

Within-person inconsistency may well be the result of development of ethnic identity (along with other aspects of identity) over a person's life (Phinney and Alipuria 1990). A related issue is that ethnic group labels of the same ethnic (minority) group may convey different meanings. Consequently a person's choice of ethnic group label may be correlated with how she views her ethnic group vis-à-vis the mainstream dominant group (Hecht and Ribeu 1991). Studies have also found differences in responses depending on the interview location such as at home vs. at school (Harris

and Sim 2002, Harris 2002). Responses also differ depending on who reports, e.g., interviewer vs. respondent (Telles and Lim 1997, Harris and Sim 2000)[2].

The concept of ethnicity may change both at the individual and group level (Yancey, Ericksen and Juliani 1976). Xie and Goyette (1997) find that for persons with one Asian and one non-Asian parent between the 1st and 2nd generation there is a process of assimilation (decrease in percentage identifying as Asian), between the 2nd and 3rd generation there is an increase in percentage identifying as Asian and this increase is even greater for higher educated parents. Political considerations (such as the need of ethnic minority groups to be heard in a democracy) may lead to formation of new ethnic identities or redefining existing ones (Urla 1993).

Another source of inconsistency in ethnic identity measures are the socio-demographic characteristics of persons. Hecht and Ribeu (1991) find differences in the choice of ethnic label by gender, age and family income. Travassos and Williams (2004) also find differences in ethnic group identification by socio-demographic characteristics. In Brazil wealthier people classify themselves - and are classified by others - in lighter-skinned categories. Telles and Lim (1997) report similar findings for interviewer reported race/ethnicity with respect to education and income.

The problem of finding a consistent measure is magnified for children of inter-ethnic marriage or cohabitation and persons who have recently immigrated. Persons of mixed parentage may choose different ethnic group identities depending on which parent is present at the time of the interview or the proportion of persons of the majority group present (Martin and Gerber 2006, Phinney and Alipuria 1992, Xie and Goyete 1997). Others may wish to claim more than one of the offered categories and may resist being forced to choose (Aspinall 2000) or allocate themselves arbitrarily between the options proffered. Those who have recently migrated retain the concept of ethnicity of their country of origin. If that is very different from the concept of ethnicity of the host country then there are reportage problems. For example Martin and Gerber (2006) find that persons from Central and South America and the

---

[2] It should be noted that observer reporting is no longer accepted except when researchers want to find out 'wider society's perception of ethnicity' and in case of "very young and permanently confused psychiatric patients" (Aspinall 2001).

Caribbean do not think race and ethnicity are separate concepts and would like to choose the country they have come from as their race/ethnicity (see also Massey 1993).

Historically the two most common measures of ethnicity have been country of origin, own/parents/grandparents' country of birth and self-identified ethnic group. Aspinall (2000) places these measures on a 'continuum of stability' with country of birth or origin being the most stable and self-identified ethnic group using open response being the least. While even country of birth can be susceptible to recall error or reconstruction, it is uncommon. Therefore an interest purely in stability might place substantial emphasis on ascertaining geographical antecedents of individuals to capture ethnic group distributions and variations. On the other hand, country of birth information may well not capture what is intended for monitoring purposes. For example, knowing someone was born in Zimbabwe may not be informative if the interest is in knowing who is 'non-white' and potentially subject to discrimination on that basis.

3 Designing ethnic group question(s)

3.1 Operationalising ethnic group questions

The design of questions to measure ethnic identity of a person becomes, then, an attempt to bridge the need for consistent, reliable measures of ethnic identity and for capturing people's perception of their own ethnic identity. While an open text question serves the latter it fails miserably in the former, quite apart from the coding issues it introduces. Explaining the context of the question in an open text question, through the use of examples may be helpful (Martin 2002); but she points out that they need to be used with caution since in some contexts examples can encourage responses specifically for the example categories.

Pre-designated categories will reduce inconsistency in reporting but not eliminate it especially when the underlying concept itself changes. Pre-designated categories come with their own baggage of worries. Respondents, especially those of mixed parentage, may choose different identities depending on the context of the survey – personal vs. official. It is claimed by some that 'official' surveys have created 'fictive unities', which are categories that respondents have learnt to use when identifying

themselves in official situations (Werbner 1990, cited in Aspinall 2007). When asking an open-text unprompted ethnicity question and the existing Census ethnicity question, Pringle and Rothera (1996) found that there was an exact or near exact match in only 28% of the valid responses. Martin and Gerber (2006) raise a similar issue. They point out that pre-designated categories imply that these are the 'correct' categories and unless there is a clear understanding in the society what these 'correct' categories comprise they will constrain people to report within categories that they do not identify with. Bates et al. (2006) also reveal the substantial frustration that can arise when the categories are felt to deny preferred identifications.

For some purposes as long as people can match themselves to the correct categories, the extent to which they identify with that category may not be considered to be an issue. For example, in the US, the desire to measure changing ethnic identity appears to come second to the political need to collect data on certain groups and so the focus is more on reliability (has the person marked the 'correct' box) rather than validity (is the answer given one that truly reflects the identity?). As Anderson and Feinberg (1999) note, there is "tension between the choices presented by the boxes in a census questionnaire, which define the possible answers according to administrative requirements, and the complex lived identities of individuals". In the US, growing criticism of the ability of the Census to measure new immigrants or the children of interracial marriages prompted the Office for Management of the Budget (OMB) to review the ethnic categories. On the one side there were well-established civil rights organisations who lobbied for the categories to stay as they were, whilst on the other side newly-formed groups of 'multiracial' Americans campaigned for the addition of a 'multiracial' category. In addition, a number of other groups lobbied for the break-up of the "White" category to reflect geographical heritage, including the reclassification of people of Middle Eastern origin from "white" to a new category. In the end, the OMB balanced these political interests and came out in favour of allowing people to code more than one option (Nobles, 2002).

In the UK, the census questions for the relevant census decade get asked across a range of surveys and administrative contexts and are recommended categories for the monitoring of employees and service users/clients that is mandatory for public bodies. Thus people are likely to 'learn' a response for these categories (or alternatively learn

to resist it), potentially increasing the stability but not necessarily the meaningfulness of the response. If *Understanding Society* uses the relevant census categories, then it may also impact on responses other questions that follow, not only alternative questions about ethnicity, but also about other 'dimensions' of identity, such as religion. For example, the surprisingly high rates of Christians reporting in the 2001 Census was attributed in part to the placing of the religion question after the ethnic group question: responses as 'Christian', particularly given the wording of the question which did not ask about practice or belonging and which did not specify separate Christian denominations (in England and Wales), were taken to imply a cultural interpretation of Christian, equated with white English (or Welsh) ethnicity.

Positioning of questions is explicitly used to increase the numbers completing the 'right' box in both UK and U.S. census. Thus, Hispanic 'ethnicity' is now offered prior to race in the U.S. census to enable the initial assertion of this identity, to improve rates of response to the question and to encourage (with partial success) fewer responses of 'Hispanic' or 'Latino' to the subsequent 'race' question (Bates et al. 2006; Martin et al. 1990). Similarly, the proposed placing of a national identity question prior to the ethnic group question in the 2011 Census is designed to allow all those who wish to assert their Britishness to have an initial opportunity to do so, with the intended result that fewer minorities will 'incorrectly' subsequently select the British sub-option of 'white' or use the write-in sections to define themselves as British. Nevertheless, while such design improvements may improve the utility and accuracy of a classification system for monitoring it may only weakly capture subjective identity or the different weights attached to the different elements of 'self'.

In general, when persons of mixed parentage are asked to choose one category they find it very difficult and there is high inconsistency of reporting (Harris and Sim 2000). Perhaps using open texts and multiple responses are better solutions (Wallman *et. al.* 2000, Lee 2001). On the other hand the addition of specific 'mixed' categories in the 2001 UK Census (a change from the 1991 Census) did reduce the number of people allocating themselves to the residual 'other' categories, which are the least stable and which do not constitute meaningful groups (Simpson and Akinwale 2007). Asking country of origin (or ethnic origin) of parents as well as of self will help identify persons of mixed parentage (Berthoud 1998). However, it should be noted

that their own ethnic identity may be different from those of either of their ancestors (Aspinall 2001, Martin and Gerber 2006). To adequately capture mixed ethnicities in *Understanding Society* or to capture them in a way that meets the research needs of different users, may therefore imply multiple approaches.

Similarly, new immigrants with different concepts of ethnicity may find choosing the categories prevalent in the host country meaningless. For example, Martin and Gerber (2006) finds that persons from Central and South America are expected to choose Hispanic as the 'correct official ethnicity' and to choose a separate race category, such as white, black, Asian etc. However, such persons find this classification scheme very problematic. They are not familiar with the term Hispanic, they do not think race and ethnicity are separate concepts and would like to choose the country they have come from as their race/ethnicity (Massey 1993). They find it difficult to choose between white and black and would prefer to choose some colour in between. Travassos and Williams (2004) also point out that difference in ethnicity concepts between US/Brazil/UK makes cross-country comparisons meaningless. For example, until recently Asian Indians (UK classification) in the US were classified as "White"; those from the Middle East still are.

As mentioned earlier the strong need for a political voice by ethnic minority groups in a democracy sometimes makes the choice of ethnic categories heavily dependent on the numbers game. The introduction of the multiple response race question in the U.S. census was met with opposition from black interest groups who feared that this would lead to a reduction in the proportion of blacks and consequently a reduction in their political power (Farley 2002, Skerry 2002). Goldstein and Morning (2000) estimated that upon the introduction of the multiple response option in the U.S. Census, 8 to 18 million people who earlier had chosen single race categories would choose more than one race category. In reality 6.8 million people chose two or more races (Lee 2001). The perception by ethnic minority groups of the purpose of collecting this information has serious implications for item nonresponse. An example of an extreme case is the dropping of this question in the 1981 British Census because of very high item non-response in the 1979 Test Census resulting from fear (among ethnic minority groups) that these data might be used to deport persons 'not of British descent' to their 'home countries' (Bulmer 1986).

In addition, the form of a categorical question can constrain the opportunity for assertion of a political minority identity. In the UK 'black' is primarily used to refer to those of African and Caribbean origins, but has also been employed as a unifying term for the positive assertion of minority status across all 'non-white' minorities (Hall 1992) and is, for example, used in this way by the Trades Union Congress and its campaigns for equality for 'Black workers'. This dual usage has led to it being resisted as a means for describing 'Asian' groups (which in the UK primarily refers to those of South Asian origin, itself another source of potential constraint or confusion) (Modood 1992), with the argument that the tracing of patterns of ethnic minority disadvantage (or advantage) with the assistance of, for example, ethnic monitoring, needed to be kept separate from people's political consciousness and over-arching identification. Such arguments appear to have been successful in the way that census questions are constructed to link Black with Caribbean and African. However, if self-identity as 'black' in this politicised sense is central to people's self-perception they may put themselves in the 'wrong' box or experience frustration at the inability to express this sense of self, by contrast with a set of categories that deals primarily with geographical areas.

Consistency over time requires carrying the same ethnic group categories over the years but these 'old' categories may not be meaningful over time. While these 'old' categories reflect the historical processes that have resulted in particular ethnic relations and categorisations, the 'new' categories reflect self-identification of the ethnic groups as of today (Aspinall 2001). Also, in the UK the 'older' categories although useful according to some in research on health and illness, have strong racial overtones. In a longitudinal survey such as *Understanding Society*, it is possible to investigate the extent to which people redefine themselves over time – that is, the extent to which changes in distributions of categories reflect demographic changes or changes in identification.

As we can see a crucial criterion in deciding which categories to include in closed form ethnic self-identification questions is the acceptability of the ethnic identity question (Sillitoe and White 1992). In general, if the categories used do not make sense to the respondents then there will be a lot of item non-response or 'other'

reporting. Also when the interviewers find some of these questions ambiguous they try to resolve these issues differently, resulting in interviewer effects on the responses (Martin and Gerber 2006).

As well as the determination of categories themselves, acceptability issues and whether to use single or multiple response questions, there are a number of other practical issues associated with designing an effective ethnic group question. Pre-designated categories have list effects (recency and primacy effects). Long lists also exacerbate mode effects, especially between telephone and face-to-face. Minor changes in the transition from pen and paper to CAPI may also affect responses (Smith 2008). Finally, mode effects may be confounded by self-selection (Martin and Gerber 2006) if, for example, those who do not understand English are both less likely to complete a mail questionnaire (and are thus interviewed face to face) and tend to report ethnicity differently from those who do have a strong grasp of English.

Question wording can also be important. Martin and Gerber's (2006) study is based on U.S. race and ethnicity questions but the issues they raise may also be pertinent to the UK. The word 'describe' brings up images and so respondents may think that they are expected to report on how others 'see' them. "best indicate …'s race" is problematic because it may be interpreted as the race that is considered to be superior to others. Using the word 'better' would also create similar problems. They suggest using terms such as 'identify' and 'consider him/herself to be'. However, if combined with multiple options or disparate categories, creating a wording that is both clear and meaningful to respondents and avoids these sorts of leading biases is potentially demanding, requiring careful development work. When the question contains options for choosing more than one category for the first time, respondents may not realise this change as they are used to choosing one category. Martin and Gerber emphasise the efforts that need to be made to highlight that addition, by putting it in more than once in the question.

3.2 Implementation of ethnic group in the UKHLS

To capture some of the different elements of ethnicity or ethnic group identity and the different interests of research users in what measures of these provide, it was considered that it would be necessary to incorporate multiple, even if overlapping,

ways of capturing people's ethnic group or origins in the survey context. It is also clear from the discussion above that it would be important to consider sequencing of questions, (non-)availability of politically salient categories, multiple response options as opposed to single categories, and the wording and practical design of questions (for example, the length of list options and the ordering of options within these).

We recognised that inclusion of a categorical question in the form of the Census question[3] would be important not only for range of research purposes concerned with analysis of patterns of association between ethnic groups and other characteristics, as well as detailed understanding of population characteristics and their relationship over time, but that it would also be important for purposes of comparison and calibration with other sources, both census and other surveys. In a complex longitudinal survey such as this, multiple weights will play an important role in constructing different samples for analysis and core variables such as ethnic group categories will be necessary to construct these weights.

However, alongside this it was equally clear to us that such a categorical question would not represent ethnic identification or belonging and would not meet the research needs of social psychologists or of those interested in understanding in more detail the meaning and coherence of groups. For such purposes it would be important to develop a question which tackled who people think they are and how important they consider that to be. Thus it would need to incorporate flexibility, through multiple response options, allow respondents to give some weight to the different elements specified. Such a question requires development and would not be asked at the first wave of the survey. Moreover, as we discuss below, there could be additional advantages in having a multiple response question at wave 2.

At the same time, it was clear that it would be important to address existing and alternative research agendas through a range of questions (some of them novel in their own right) that in combination or separately would meet expressed research needs and cover both the possible components or dimensions of 'ethnic group' and the

---

[3] We talk of the Census question, but in fact the ethnic group question differed in the 1991 and 2001 censuses and will differ again in the 2011 census. We are adopting a version that reflects the main changes to be carried in 2011.

complementary and related concepts covered in Section 2. Thus the proposal was to field questions which would, both between them and in combination with the other data collected in the survey, meet a wide range or research interests in ethnicity, nationality, immigration, religion and so on. Moreover many of these questions would only need to be asked once or rarely and thus their inclusion in wave 1, with its emphasis on collecting histories and 'one-off' information is particularly suitable.

These dimensions asked in this first wave could then feed into the multiple response options for an ethnic identity question to be asked in Wave 2, and avoid the need for long (and necessarily partial) lists of categories Thus, individuals could be asked about whether they see 'your religion' 'your mother's/father's country of birth', 'your nationality' etc as part of their ethnic identity, and how important each identified component is. Given the conditioning effects of ethnic group questions and the importance of sequencing, it was also clear that such a question would need to be placed at a temporal distance from a (census type) classificatory question, and given the potential sensitivity of responses to wording attention would need to be paid to both how people were asked to respond, to the order the different elements were placed, and to the ways in which they were invited and /or encouraged to provide more than one response. Such a question is currently undergoing development for inclusion in Wave 2 of the UKHLS. This will also allow it to be clearly separated from the categorical question (to be asked in Wave 1), and ensure that it can use the information supplied separately on country of birth, religion etc. to feed into its construction. Cognitive testing has already established the acceptability and comprehensibility of a multiple response question and attention is now being given to the precise wording, ordering and son on. Given the proposals for a mixed mode design for Wave 2, development work will also need to be attentive to minimising mode effects and the development of a question suitable for asking in multiple modes. Such a multiple response ethnic identity question (or questions) would provide a new departure in UK survey research on ethnicity and, combined with the longitudinal nature of the survey, and the potential for repeating the question in subsequent years (at reasonably long intervals to minimise conditioning effects) is likely to open up a whole new domain of research possibilities.

Meanwhile all the ethnicity related questions to be asked in Wave 1 have been included in the innovation panel and have been shown to be able to produce meaningful and generally consistent results (where inconsistencies themselves are a potential source of interest in the survey as a whole). A number of them have been also been cognitively tested with minority as well as majority group respondents and their comprehensibility and acceptability have thus been broadly confirmed. They are outlined in more detail below and the ways they can be utilised by researchers to investigate particular constructs is also illustrated.

Proposed ethnicity-related questions

1. Census classification question with single response categories

As noted, this question will enable comparison of distribution and characteristics of ethnic groups in the UKHLS with other surveys. It will also be important for matching to demographic information on the UK (and its changing population) and for the calculation of suitable weights. In relation to what it will tell us about ethnicity and ethnic groups, it sits in an ambiguous position between a classification based on fixed or stable attributes (such as country of birth) and a subjective measure of ethnic identity. Thus in Phinney's terms it will be partially informative about both 'ethnicity' and 'self-identity' though will not inform us comprehensively about either. What it will provide is reasonably 'stable' ethnic group information, in the sense that people are generally fairly consistent in their response to the categories and additionally learn responses. It will allow a classification of the population into mutually exclusive categories. When combined with country of birth information (see below), it will also allow a more detailed understanding of the residual 'other categories', at least in terms of national origins. It will provide information (alongside its long-standing objective, on 'non-white' or 'visible' minorities) through the separate classification of white groups including a 'white other' group. However, as noted, it will not resolve the issue for some groups which cannot satisfactory allocate themselves within the existing classification and may regard themselves (and be regarded) either as 'white other' or as 'other ethnic group'. The classification should, broadly speaking, match onto those groups at risk of discrimination and for whom anti-discrimination legislation is designed, especially when combined with information on religion (see below). When combined with country of birth information it will provide the intersection of ethnicity with immigration that is central to a large body of research.

Overall this question is likely to be used by a large number of researchers interested in diversity of outcomes across ethnic groups; the potential role of discrimination in understanding differential outcomes, and allowing for ethnic difference in analysis dynamics and durations of experiences as well as understanding possible diversity in the consequences of particular events (such as divorce), particular behaviours (such as geographical mobility) or particular attitudes (such as job or educational aspirations).

2. National identity

As highlighted in the discussion above, it will probably be important that the ethnic group classification question should follow one on 'national identity' covering identification with any of the four countries of the UK and/or with British. This is intended to ensure more stable responses both from minorities who wish to assert their Britishness and, for example, from Scots who wish to assert their Scottishness (and will thus have been already provided an opportunity to do so) to the subsequent ethnic group categories. In addition, the extent to which people identify with different countries of the UK and express such national identities, whether or not living in the countries concerned is of interest in its own right to researchers both concerned with nationality and with the complementary and complex nature of identification.

3. Own, parents' and grandparents' country of birth

Many surveys ask about country of birth (and date of arrival for those born abroad) and thus enable the distinction between immigrants and UK born, a critical distinction for a large stream of immigration-focused analysis. However, while ethnicity and country of birth information can be informative about changes between first and subsequent generations it cannot distinguish between second and subsequent generations. Asking additionally about parents' and grandparents' country of birth can provide new understandings of generational difference as well as shedding light on the complexity of people's antecedents and transitions across generations. Such information will provide a measure of Phinney's 'ethnicity' dimension and be informative for those interested in ancestry, heritage and also provide some information for those interested in national identity.

4. Religious affiliation/ upbringing and importance of religion to self.

These questions between them shed light both on religion as a cultural attribute or an element of ethnic identity and religiosity as an important element of identity more broadly. They will be of interest to those interested in religion as a primary source of identity and to those who see it as an important dimension of ethnicity – or who are interested in the intersection between the two and relative experience of those with different ethno-religious backgrounds. In a longitudinal context it will allow disentangling of age and cohort effects in religious affiliation (Voas and Crockett 2005) comparatively for different groups, as well as the transmission of religiosity across generations. It may well be relevant to understanding coherence of ethnic groups and networks and attitudes to 'in' and 'out' groups. Religious practice/religiosity is likely to explain more variation in different socio-economic behaviours such as education, marital and fertility choice than simply religious affiliation.

5. Mother's and father's ethnic group and identification with mother's and father's ethnic group.

This question is also a novel departure, and is intended to facilitate a much greater understanding of 'mixed' or multiple identities (Aspinall 2000). It will both complement and shed light on responses to the 'mixed' (and non-mixed) categories in the census classification and additionally, allows people to express strength of identification. It thus will convey some idea not only of 'self-identity' (in Phinney's terms) but also 'belonging'. It will also be informative about origins or heritage /ancestry, and allow new research developments in terms of how people's identity differs from that of their parents or the relative significance of father's and mother's (perceived identity). Given the household context of the survey, for some respondents it will also be possible to compare children's perceptions of their parents' ethnic group with the parent's own ethnic classification, and it will also be able to explore variations across circumstances and ages, and how this relates to the development of identity along other dimensions. It represents a 'middle ground' between stable categorical measures of ancestry and those of identity and identification.

6. Language and fluency in English

For some ethnic groups language is a critical or primary unifying dimension. In addition, for those interested in immigration and the penalties associated with being

non-UK born (for example, in employment) fluency in English is a significant consideration. Moreover language and English fluency can be an important consideration to understanding social networks and patterns of association. The longitudinal nature of the survey will enable a greater understanding of the causal relationship between fluency and employment outcomes or social networks, that is does people's fluency increase because of their contacts or work or do their networks and employment opportunities expand following increased fluency. What the current question does not cover is all the potential languages that respondents might speak. Though, for those with only limited spoken English fluency we will have some information on the language the interview was conducted in.

7. Citizenship

Measuring citizenship and resident status gives insights into formal aspects of belonging or residence rights (to the UK or another country or both). It can be informative about national identity (according to some formulations) and also about relative vulnerability or marginality in terms of security or status. Over time it may be relevant to understanding patterns of onward or return migration or the circumstances under which people change their status.

This rich selection of measures will enable, separately or in combination a wide range of research questions into ethnicity to be addressed as well as comparison between respondent categories on other questions. Figure 1, gives an illustration of how different constructs can be derived from cross-classifications of two measures. The row and column headings contain a selection of the ethnicity related measures to be carried in *Understanding Society*. The body of the table illustrates possible constructs that could be derived from the cross-classification. It is straightforward to conceive of a multi-dimensional table which takes advantage of three or more dimensions and which cross-classifies in alternative ways. The Figure, while only suggestive, is indicative of the strength and value of allowing multiple points of reference for considerations of 'ethnicity' rather than attempting to develop and rationalise one measure to categorise the population in a single, mutually exclusive fashion.

Figure 1: Illustration of possible combinations of ethnicity related measures to build up analytical constructs

	Religion	Country of birth	Identification with parents' ethnicity
Ethnic group (cat)	*Ethno-religious groups*	*Immigrant analysis*	*Identity; understanding of 'mixedness'*
Parents' country of birth	*Religious / identity development*	*Antecedents, heritage and generation*	*Identity development*
'Britishness'	*Competing / complementary identities?*	*'Assimilation'*	*Orientation; distance*

3.3 Who should be included in the ethnic minority oversample?

As noted, *Understanding Society* is to contain an oversample of 'key' ethnic minority groups. Much consideration has been given to which groups should be included in this oversample. Proposals for a Longitudinal Study of Ethnic Minorities (LSEM) (Nazroo 2005) had progressed a considerable distance before it was decided to incorporate that study into plans for the large, new, population-coverage household panel study. Many of the original principles and proposals were retained in some of the core design elements of *Understanding Society*. These were included in some of the specifications of the oversample: it had been decided that the LSEM should be a general purpose longitudinal survey of *ethnicity* rather than of culture or immigration, which had implications for which and what sort of groups were prioritised. It led to a focus on those groups (or categories) which could be considered reasonably meaningful, distinct and homogenous as groups. The LSEM also acknowledged the practical constraints that it only made sense to oversample those groups where sufficient numbers could be effectively sampled (at not unreasonable cost) to permit effective analysis. It would be 'wasteful' to sample small numbers of a range of groups which could not be analysed separately. This tended towards the inclusion of groups which

were both relatively numerous and showed some degree of geographical concentration, and who could be expected to respond in a consistent fashion to a simple screening question. The designated groups from the LSEM proposals were Indian, Pakistani, Bangladeshi, Caribbean and Black African. Clearly the Black African group did not meet the criterion of homogeneity including people from different countries, ethnicities and with widely varying migration histories, languages and historical relationships with the UK. Nevertheless it was considered, essential to include those of African origin within the sample and the blanket coverage was deemed preferable to a focus on just one or two African nationalities, which would then be highly contested. The largest African groups would still have the potential for discrete analysis. And some researchers would still choose to aggregate all African respondents, particularly if they were concerned about consistency with other sources. These five groups were carried over into *Understanding Society* specification by the proposal that the oversample should be designed in such a way as to ensure 1000 adults would be sampled from each of these groups. Moreover the advantage of the new study would be that other groups would still be sampled proportionate to their populations in the main sample of *Understanding Society*.

However, these five groups did not define the whole coverage of the oversample and in developing the design and throughout the consultation process for *Understanding Society*, consideration was given to the ways in which the oversample *in combination with the main sample* could best meet research interests and priorities. A crucial insight was that the criteria used to construct the oversample were largely independent of any ethnicity-related questions that were then asked. For example, research interests in religion and comparisons of religiosity would require an oversample with a sufficient numbers from a range of religious affiliations but to achieve that, giving existing knowledge about the broad distribution of religion across ethnic and national groups, it would not be necessary to select people by asking them about their religion. Nevertheless, it was clear that the screen question would be a categorical, classificatory question and that many of the factors relevant to categorisations of ethnicity as discussed above remained relevant. For example, issues of stability and of acceptability, of willingness to respond were highly pertinent. Moreover, given the particular nature of a screening question for an oversample – asked on the doorstep of one household member only – ease of response in such circumstances and the ability

of a single household member to speak for all were significant considerations. For the UK at least, colour (or other visible physical characteristics) are argued to be a defining factor of ethnic relations within the UK (Nazroo and Williams, 2005), as we have discussed. And initially, for the LSEM, it had been decided not to sample white minorities. However, as we have also discussed, there is in some cases substantial ambiguities about who is and isn't white: and it was clear that researchers as well as potential respondents meant different things by 'white minorities': some referring exclusively to European (largely relatively recent migrants) and some referring to Turks and North Africans. Moreover many were keen to include such groups in the oversample. To resolve some of these issues of inclusion in the boost sample, we returned to the issue of the extent to which meaningful groups for analysis would be captured by the oversample. This implied that the oversample should

- o select groups which were positively identified as having some internal coherence; and conversely exclude categories which were aggregates of widely different groups;
- o use categories that would be meaningful to respondents including household respondents defining other members of the household;
- o select groups which were sufficiently numerous and / or showed sufficient levels of geographical segregation that a screening process based on a selection of geographical areas would supply sufficient numbers for effective analysis; and conversely exclude categories which would produce numbers of respondents insufficient for analysis;
- o select groups where a screening process which was restricted in terms of geographical coverage would include those who were, broadly speaking, representative of the overall population of that group (or who were, at least, not atypical);
- o include groups of substantial research and or policy interest *in a longitudinal context*; and conversely exclude categories which would have limited long-term meaning or research and policy salience.

The screening question developed on these criteria (see Figure 2) and which is currently being piloted had then some commonalities with a standard census-style categorical question – and with the purpose of such a question. It also shared with most categorisations and with the original LSEM proposals the aggregation of those

of African origins into one group, somewhat at odds with our aims for meaningful and non-aggregate categories. However, by maximising the numbers from this group we aimed to provide analysts with sufficient numbers from some of the larger national origin groups to enable discrete analysis of those groups. But the screening question also had some important differences from the census categorical question. In particular, it did not use the words black or white (as a means for inclusion or exclusion) except when trying to define African respondents more specifically; it positively identified those groups and only those groups which it intended to include: these included the five groups originally identified for the LSEM and some additional categories that identified as substantial long-term interest, reasonably numerous and not well-captured by existing categorisations (Turkish, Middle Eastern, North African and Sri Lankan); it asked about ethnic group in a way that emphasised ancestry (parents or grandparents) rather than asserting subjectivity; it had no 'mixed' categories, but set out to maximise the number of those with one Indian or Caribbean parent and one parent from another ethnic group (given the extensive research interest in 'mixedness' and its interest in longitudinal perspective) by specifically asking about this combination of parents' ethnicities for those with at least one parent/grandparent from one of these groups. Moreover, the screening question is designed for screening and not for analysis. There is no reason why *as a question* it should meet criteria for a particular construct or dimension of ethnicity. It may be of interest, particularly from a validation and methodological point of view to see how screen categories map on to ethnicity-related questions; and this has already been done with a version of the screen question that was asked of all household respondents in the innovation panel. However, its intention is not to be of analytical value.

Figure 2: The form of the screen question for the ethnic minority oversample

Does anyone living at this address come from, or have parents or grandparents from, any of the following ethnic groups or origins?

A) Indian
B) Pakistani
C) Bangladeshi
D) Sri Lankan
E) Chinese
F) Far Eastern
G) Turkish
H) Middle Eastern and Iranian
I) Caribbean
J) African (including North African)
K) Other minority group
L) No – none of these
(Note: K and L included for completeness, but none selected for inclusion in the oversample)

If A (and not C-H) then ask: and does that include anyone from mixed Indian and other ethnic group origins?

If I (and not C-H) then ask: and does that include anyone from mixed Caribbean and other ethnic group origins?

If J (and not C-H) then ask: and which of the following most closely describes those African origins: i) North African; ii) African Asian;
iii) Black African; iv) White African.

4. What are relevant topic areas for the study of ethnicity and ethnic group differences?

4.1 Key domains

If the measures discussed above only capture certain dimensions of ethnicity, then additional measures may capture other areas that are relevant to ethnicity and ethnic group, but which it would be impossible and inappropriate to include in a set of categorical questions. Such additional areas may also be ones that are more susceptible to variation over time and therefore warrant being asked more frequently, or where developments in relation to household formation and dissolution may be particularly relevant, for example how attitudes to suitable marriage partners may

relate to actual marriage patterns. Experiences, attributes and attitudes that are pertinent to ethnicity and ethnic group membership include

- o discriminatory and prejudicial attitudes and behaviours, that may vary over time both in intensity and in terms of the groups they are directed against, and may correspondingly shape the experience and behaviour of other groups
- o patterns of friendship, association and social participation, and the extent to which these are mono-ethnic (or mono-religious)
- o romantic relationships and attitudes to romantic relationships of self or others, across ethnic or religious boundaries
- o connections among minorities and/or immigrants with country of origin or family connection via, for example, visits, the sending of remittances, marriage patterns and so on.
- o political affiliation, participation and activity and use of alternatives to formal political processes, e.g. through activism.
- o expressions of religious, ethnic or cultural identity through, for example, dress, celebration or fasting, charitable giving, religious observance, additional educational activities such as religious instruction or language learning and so on.
- o expressions of self-worth and self-esteem

These are all susceptible in some form to measurement in a survey and some are of particular interest in the longitudinal context, such as changes in patterns of friendship or their relationship to other outcomes (Kalter 2006). They can also be seen as shaping as well as being shaped by ethnic identity. These areas also relate to considerations of ethnicity and ethnic group and ethnic identity formation that are not effectively explored in the range of questions outlined above as potential measures for dimensions of ethnicity. We discuss three of them in more detail below.

4.1.1 Discrimination and Prejudice

One aspect of the British Race Relations legislation discussed above resulted in the need to monitor the situation of ethnic minorities. However, the main aims were to tackle discrimination which emerged with the migration of people from the Indian sub-continent and the Caribbean. There is plenty of cross-sectional evidence in the

UK of an 'ethnic penalty' (Heath and McMahon 1997) across minority groups in employment and earnings (see e.g. Heath and Cheung 2006, and the discussion in Platt 2007). And part (or sometimes all) of that penalty has been assumed to be discrimination on the part of employers. Discrimination has been directly investigated in 'tests' involving ethnically differentiated but otherwise identical responses to job applications but such studies are now not current and so do not reflect the current political and employment context. Attitudes to other groups and prejudiced or discriminatory views are collected in some attitudinal surveys, but we have only a limited understanding of how that relates to individuals' own group belonging or identity.

Feelings of being discriminated against, and having to deal with prejudice, are likely to be one way in which an individual can identify themselves as part of a group, an "us" being discriminated by "them". The attitudes and behaviour of an external group (not necessarily the majority group) may help in promoting solidarity among a group of people who come to see themselves as having something in common. It is important that these issues of prejudice and discrimination are measured in the UKHLS because they may affect how an individual sees themselves (self-esteem), others who share their ethnicity and their attitude to other groups. These perceptions may change over time depending on the changing circumstances and experiences of the individual. Moreover experiences of discrimination may affect patterns of behaviour over time, something we currently have little understanding of.

Attitudes to own group, sense of groupness and issues of religion and value will also be likely to affect attitudes to engagement in romantic relationships, including marriage and leaving home as well as attitudes to those of others.

4.1.2 Patterns of friendship, association and participation

Nazroo (2005) notes that "issues of exclusion, integration and social participation (how different ethnic groups orientate themselves to each other, local communities, and civic life) are of great academic and policy interest". There is substantial academic and policy interest in understanding people's networks and how they do or do not contribute to mobility and life choices. These include friendship patterns, employment contacts and leisure activities and patterns of volunteering and civic

participation. Currently information on ethnic similarity and difference across these areas and their consequences for individuals and communities is partial and is derived only from a limited number of questions in cross-sectional surveys.

Interest in 'community relations' has been variable over time, but recently the issue has moved up the agenda, often referred to as 'social cohesion'. According to monthly Ipsos-MORI opinion polling the issue of "immigration, race relations" is, at the time of writing (April 2008), now the most important issue facing Britain today (42% of respondents gave it as a response). And community cohesion has been subject to sustained policy attention. Key to robust analysis of these issues is the need to have data which is collected with information on ethnicity. The Commission on Integration and Cohesion state that at the moment much information already collected is not coded by ethnicity and "if data is not coded by ethnicity it means that there can be no diagnosis of problems, no targeting for this group and no monitoring of the impact on a particular community".

It is clear that, in addition, to the ability to address current policy concerns a more developed understanding of people's networks and how they develop over time across ethnic groups is informative about the formation and reformation of 'groups', their meaning as groups and stability or change in that meaning. Such information can address important sociological questions as to the extent to which those who ostensibly belong to the same 'group' also act as a group. Measurement of transnational networks and inter-household patterns of exchange and sharing also allows consideration of networks beyond everyday, face-to-face contacts and potentially complements a narrower perspective on people's participation and sources of support and friendship, as well as challenging purely localised definitions of 'community'.

4.1.3 Poverty, Disadvantage and Exclusion
Poverty and exclusion are major research interests and policy concerns. The interest in the extent to which disadvantage and poverty are differentially experienced across ethnic groups is not, however, matched by equivalent data availability. The National Strategy for Neighbourhood Renewal (2000) noted that "some of the groups in society that are the most vulnerable to becoming victims of social exclusion are forgotten

simply because not enough is known about their particular circumstances" and advised that the Office for National Statistics should ensure that administrative data is coded by ethnicity. While poverty, in particular child poverty, has risen up the political agenda in recent years, there is only limited scope for detailed analysis by ethnic group, particularly longitudinal analysis (Platt 2007); and ethnicity has only recently been explicitly recognised as a dimension of the child poverty agenda.

Poverty and exclusion, to the extent that they are differentially experienced, can justify the classification and distinction of groups even in the absence of strong identification with a particular classification. If belonging to a particular ethnic group is associated with poorer life chances, greater chances of exclusion and higher risks of poverty then it invites exploration of that 'group' in and of itself, and as a category having meaning in these terms. The argument is analogous to, though not identical with, Lambert's (2002, Lambert and Penn 2001) proposals for measuring ethnicity as a function of social stratification and thereby illuminating the meanings we give to 'ethnicity'. To return to the earlier analogy with class, it is relevant that those from lower social classes have poorer health outcomes than those from higher social classes, whether or not members of either category identify themselves as belonging to a class. Conversely, where life chances are indistinguishable from the general the arguments for separate classification of that group are weaker or may take a different form, in terms of interest in identity and identification per se. Thus, we can see groups that share differential life chances as analytically meaningfully groups, whether or not they are sources of individual identification or socially constructed as groups, and such distinctions give weight and rationale to classification systems which reveal such systematic differences. Understanding how these patterns of disadvantage play out in a longitudinal context will help shed light both on how they come to concentrate in particular groups and also whether that analytical meaning is sustained when we consider dynamics and durations of disadvantage and poverty.

4.2 How do we measure them?
Such topic areas, then that are informative about ethnic diversity in key areas of interest and about the nature of groups, group boundaries and 'groupness' itself (Wimmer 2008). In some areas good models exist for questions to capture these different domains (for example, questions which allow the measurement of poverty;

questions on political affiliation and participation). In other areas, there are more limited models or it is not clear that existing questions have been cross-culturally validated – that is, they might be differently understood across groups or elicit different types of answers. In these cases, for example in relation to social networks, participation, bases of interaction and exchange, it is important that we pay attention to the development of appropriate questions, considering as we did for ethnic group questions issues of wording, mode and presentation as well as the precise constructs we are aiming to measure. We are therefore planning a development and testing process for these areas of question content. This development process will need to balance the general interests and ambitions for such questions among those not specifically interested in questions of ethnicity with the ambition to understand their contribution to constructs relating to ethnic 'groupness'.

5. Conclusions and next steps.

Understanding Society has the potential to be an unprecedented resource for understanding of ethnic identity and ethnic groups, and their formation and development. It will enable researchers to pursue new research questions (in the UK context), both descriptive and analytical. These will include both questions that explore the meaning and composition of groups and identities themselves. It will also include questions that relate to comparisons in other outcomes across different ethnic groups (or identities or ethnicity-related characteristics such as religion, language, generation). For example, researchers will be able to analyse ethnic differences in poverty dynamics or the consequences of partnership breakdown, or the intergenerational transmission of beliefs or resources across ethnic groups.

In order to achieve this capacity and provide such a resource, it will be essential to have good measures of the constructs researchers wish to employ and that there is clarity about what they are and are not measuring. As we have demonstrated, there is no single question which would provide an optimum measure of 'ethnic group' or which would meet the diverse needs and demands of researchers and policy makers. Instead we have mapped out how a range of questions can respond either individually or in combination to many of the underlying interests in an ethnic group question. These will enable individuals to identify with multiple dimensions, for example, a person can identify herself as Scottish (upbringing, language/accent, politics, local

affiliation), Black (politically used to express solidarity with minority, non-white experience), Asian (to suggest antecedents (parents /grandparents or earlier generations from South Asia) and British (to express nationality, citizenship etc.). We have also demonstrated that such a range of questions can provide flexibility in terms of different research perspectives and preoccupations, enabling more precise comparisons with particular bodies of literature, and relating to specific research questions.

It is only when analysts come to use the data that we will be able to demonstrate with what success *Understanding Society* has enhanced the study of ethnicity in the UK and expanded both the range and the nature of questions open to investigation, as well as the ways in which ethnicity as a variable is conceived. But due to the longitudinal design of the study, we have the chance to build up over successive ways an increasingly sophisticated battery of measures of the complex ways in which people view themselves and their ethnicity. Meanwhile the work of question development and testing and sample design to attempt to ensure the quality of the survey is ongoing.

References

Abdelal, R., Herrera, Y. M., Johnston, A. I. and Martin, T. (2001) 'Treating identity as a variable: measuring the content, intensity, and contestation of identity', *American Political Science Association,* San Francisco.

Alba, R. and Nee, V. (2003) *Remaking the American Mainstream: Assimilation and contemporary immigration.* Boston, MA: Harvard University Press.

Allen, J. P. and Turner, E. (2001) 'Bridging 1990 and 2000 Census race data: fractional assignment of multiracial populations', *Population Research and Policy Review,* 20(6): 513-533.

Allen, R. L., Dawson, M. C. and Brown, R. E. (1991) 'A schema based approach to modeling an African American racial belief system', *American Political Science Review,* 83: 421-441.

Altonji, J. G. and Blank, R. M. (1999) 'Race and gender in the labor market' in O. C. Ashenfelter and D. Card (eds) *Handbook of Labor Economics.* Vol. 3C. New York: Elsevier Science Press, pp. 3143-3259

Anderson, B. (1991) *Imagined Communities: Reflections on the Origins and Spread of Nationalism.* London: Verso.

Anderson, M. and Feinberg, S. (1999) *Who Counts: The Politics of Census-Taking.* New York: Russell Sage.

Aspinall, P., Song, M. and Hashem, F. (2006) *Mixed Race in Britain: a Survey of the Preferences of Mixed Race People for Terminology and Classifications.* The Inheritance Project, Canterbury: University of Kent.

Aspinall, P. (2000) 'The challenges of measuring the ethno-cultural diversity of Britain in the new millennium', *Policy & Politics,* 28(1): 109-118.

— (2001) 'Operationalising the collection of ethnicity data in studies of the sociology of health and illness', *Sociology of Health and Illness,* 23(6): 829-862.

— (2007) 'Approaches to developing an improved cross-national understanding of concepts and terms relating to ethnicity and race', *International Sociology,* 22(1): 41-70.

Avery, D. R., Tonidandel, S., Thomas, K. M., Johnson, C. D. and Mack, D. A. (2007) 'Assessing the Multigroup Ethnic Identity Measure for measurement equivalence across racial and ethnic groups', *Educational and Psychological Measurement,* 67: 877-888.

Ballard, R. (1996) 'Negotiating race and ethnicity: exploring the implications of the 1991 Census', *Patterns of Prejudice,* 30(3): 3-34.

— (1997) 'The construction of a conceptual vision: "ethnic groups" and the 1991 UK Census', *Ethnic and Racial Studies,* 20: 182-194.

Banton, M. (1998) *Racial Theories.* Cambridge: Cambridge University Press.

Barth, F. (1969) *Ethnic Groups and Boundaries: The Social Organisation of Culture Difference.* London: Allen & Unwin.

Bates, N., Martin, E. A., DeMaio, T. J. and de la Puente, M. (2006) *Questionnaire Effects on Measurement of Race and Spanish Origin*: U.S. Census Bureau Research Report Series: Survey Methodology No.2006-12.

Berinsky, A. (2004) *Silent Voices: Public Opinion and Political Participation in America.* Princeton: Princeton University Press.

Berthoud, R. (1998) 'Defining ethnic groups: origin or identity', *Patterns of Prejudice,* 32(2): 53-63.

— (2000) 'Ethnic employment penalties in Britain', *Journal of Ethnic and Migration Studies,* 26(3): 389-416.

Biemer, P. P. and Woltman, H. (2001) 'Estimating reliability and bias from reinterviews with application to the census race question', *Federal Committee on Statistics Methodolgy Research Conference.*

Bonnett, A. (1998) 'Who was white? The disappearance of non-European white identities and the formation of European racial whiteness', *Ethnic and Racial Studies,* 21(6): 1029-1055.

Booth, H. (1985) 'Which "ethnic question"? The development of questions identifying ethnic origin in official statistics', *The Sociological Review,* 33(2): 254-274.

Brierley, P. (2000) 'Religion' in A. H. W. Halsey, J. (ed) *Twentieth-Century British Social Trends.* Houndmills: Macmillan, pp. pp. 650-674

Brown, M. S. (2000) 'Religion and economic activity in the South Asian population', *Ethnic and Racial Studies,* 23(6): 1035–1061.

Bulmer, M. (1986) 'A controversial census topic: race and ethnicity in the British Census', *Journal of Official Statistics,* 2(4): 471-480.

— (1996) 'The ethnic group question in the 1991 Census of population' in D. Coleman and J. Salt (eds) *Ethnicity in the 1991 Census: Volume One: Demographic Characteristics of the Ethnic Minority Populations.* London: HMSO, pp. 33-62.

Chiswick, C. U. (2006) *The Economic Determinants of Ethnic Assimilation*. IZA Discussion Paper 212. Bonn: IZA.

Chiswick, B. R. (2008) *The Economics of Language: An introduction and overview.* IZA Discussion Paper No. 3568. Bonn: IZA.

Coleman, D. and Salt, J. (1996) 'The Ethnic Group Question in the 1991 Census: A New Landmark in British Social Statistics' in D. Coleman and J. Salt (eds) *Ethnicity in the 1991 Census: Volume One: Demographic Characteristics of the Ethnic Minority Populations.* London: HMSO, pp. 1-32.

Connolly, H. and White, A. (2006) 'The different experiences of the UK's ethnic and religious populations', *Social Trends* 36.

Cornell, S. and Hartmann, D. (1998) *Ethnicity and Race: Making Identities in a Changing World.* Thousand Oaks: Pine Forge Press.

Cross, W. E. (1991) *Shades of Black: Diversity in African American Identity.* Philadelphia, PA: Temple University Press.

Dobbs, J., Green, H. and Zealey, L. (eds) (2006) *Focus on Ethnicity and Religion: 2006 Edition.* Basingstoke: Palgrave Macmillan.

Duncan, O.D. and Lieberson, S. (1959) 'Ethnic segregation and assimilation', *The American Journal of Sociology,* 64(4): 364-374.

Dustmann, C. and Fabbri, F. (2003) 'Language Proficiency and Labour Market Performance of Immigrants in the UK', *Economic Journal,* 113(489): 695-717.

Ellison, C. G. (1992) 'Military background, racial orientations and political participation among black adult males', *Social Science Quarterly,* 73: 360-378.

Gardener, D. and Connolly, H. (2005) *Who are the 'Other' Ethnic Groups?* London: Office for National Statistics.

Gilroy, P. (2000) *Between Camps: Race, Identity and Nationalism at the End of the Colour Line.* Harmondsworth: Allen Lane The Penguin Press.

Glazer, N. (2002) 'Do we need the census race question?' *The Public Interest,* Fall: 21-31.

Hahn, R., Mulinare, J. and Teutsch, S. M. (1992) 'Inconsistencies in coding of race and ethnicity between birth and death in U.S. infants: a new look at infant mortality, 1983 through 1985', *Journal of the American Medical Association,* 267: 259-263.

Hall, S. (1992) 'New ethnicities' in J. Donald and A. Rattansi (eds) *'Race', culture and difference.* London: Sage.

Harris, D. and Sims, J. J. (2002) 'Who is multi-racial? Assessing the complexity of lived race', *American Sociological Review,* 67(4): 614-627.

Heath, A. and Cheung, S. Y. (2006) *Ethnic Penalties in the Labour Market: Employers and discrimination.* Department for Work and Pensions Research Report 341. Leeds: Corporate Document Services.

Heath, A. and McMahon, D. (1997) 'Education and occupational attainments: the impact of ethnic origins' in V. Karn (ed) *Ethnicity in the 1991 Census: Volume Four: Employment, education and housing among the ethnic minority populations of Britain.* London: HMSO, pp. 91-113.

Heath, A. and McMahon, D. (2005) 'Social mobility of ethnic minorities' in G. C.

Loury, T. Modood and S. M. Teles (eds) *Ethnicity, Social Mobility and Public Policy: Comparing the US and UK.* Cambridge: Cambridge University Press.

Hecht, M. and Ribeau, S. (1991) 'Sociological roots of ethnic identity: A look at black America', *Journal of Black Studies,* 21: 501-513.

Heckman, J. J., Lyond, T. M. and Todd, P. E. (2000) 'Understanding black-white wage differentials: 1960-1990', *American Economic Review,* 90(2): 344-349.

Herrnstein, R. J. and Murray, C. (1994) *The Bell Curve: Intelligence and class structure in American life.* New York: The Free Press.

Hymans, J. (2002) 'Applying social identity theory to the study of international politics: a caution and an agenda', *International Studies Association convention,* New Orleans, LA.

Kalter, F. (2006) ‚Auf der Suche nach einer Erklärung für die spezifischen Arbeitsmarktnachteile von Jugendlichen türkischer Herkunft', *Zeitschrift für Soziologie,* 35(2): 144-160.

Kertzer, D. and Dominique, A. (2002) *Census and Identity: The Politics of Race, Ethnicity, and Language in National Censuses.* New York: Cambridge University Press.

Lambert, P. (2002) 'Quantitative representations of ethnic difference'. Paper prepared for presentation to ISA RC28 on Social Stratification and Mobility, Oxford, 10-13th April 2002

Lambert, P. and Penn, R. P. (2001) *SOR Models and Ethnicity Data in LIS and LES: Country by Country Report.* LIS Working Paper No 21.

Law, I. (1996) *Racism, Ethnicity and Social Policy.* Prentice Hall.

Lee, J. and Bean, F. D. (2004) 'America's changing color lines: immigration, race/ethnicity, and multi-racial identification', *Annual Review of Sociology,* 30: 221-242.

Leslie, D. and Lindley, J. (2001) 'The impact of language ability on employment and earnings of Britain's ethnic communities', *Economica,* 68(272)

Leslie, T., Raglin, D. and Schwede, L. (2002) 'Understanding the effects of interviewer behavior on the collection of race data', *Proceedings of the American Association for Public Opinion Research (Survey Research Methods Section of American Statistical Association)*, 2063-2068.

Light, A. and Nandi, A. (2007) 'A. Identifying Race and Ethnicity in the 1979 National Longitudinal Survey of Youth', *Population Research and Policy Review,* 26(2): 125-144.

Lindley, J. (2002) 'Race or religion? The impact of religion on the employment and earnings of Britain's ethnic communities', *Journal of Ethnic and Migration Studies,* 28(3): 427–442.

Loury, G. C. (2004) '*The Anatomy of Racial Inequality*: the author's account', *The Review of Black Political Economy,* Fall 2004: 75-88.

Manuel, R. C. (1982) 'The dimensions of ethnic minority identification: An exploratory analysis among elderly black Americans' in R. C. Manuel (ed) *Minority Aging: Sociological and Social Psychological Issues.*, Westport, CT: Greenwood Press, pp.

Manuel, R. C. (2006) 'The conceptualization and measurement of race: confusion and beyond', *Perspectives,* 6(2): 7-17.

Martin, E. (2002) 'The effects of questionnaire design on report of detailed Hispanic origin in Census 2000 mail questionnaires', *Public Opinion Quarterly,* 66: 582-593.

Martin, E., Bates, N., DeMaio, T. and de la Puente, M. (1995) 'Questionnaire effects on measurement of race and Hispanic origin', *Journal of Official Statistics,* 11: 435-462.

Martin, E., DeMaio, T. J. and Campanelli, P. C. (1990) 'Context effects for census measures of race and Hispanic origin', *Public Opinion Quarterly,* 54(4): 551-566.

Martin, E. A. and Gerber, E. (2006) *Methodological Influences on Comparability of Race Measurements: Several Cautionary Examples.* Washington DC: Statistical Research Division, U.S. Census Bureau.

Mason, D. (1995) *Race and Ethnicity in Modern Britain.* Oxford: Oxford University Press.

Massey, D. S. (1993) 'Latinos, poverty and the underclass: a new agenda for research', *Hispanic Journal of Behavioral Sciences,* 15(4): 449-475.

Meyers, B. (1984) 'Minority group: An ideological formation', *Social Problems,* 32: 1-15.

Miles, R. (1982) *Racism and Migrant Labour.* London: Routledge and Kegan Paul.

Miles, R. (1993) *Racism after 'race relations.* New York: Routledge.

Modood, T. (1992) *Not Easy being British: Colour, Culture and Citizenship.* Stoke-on-Trent: Runnymede Trust and Trentham Books.

Modood, T., Berthoud, R., Lakey, J., Nazroo, J., Smith, P., Virdee, S., and Beishon, S. (1997) *Ethnic Minorities in Britain: Diversity and Disadvantage.* London: PSI.

Nagel, J. (1994) 'Constructing ethnicity: creating and recreating ethnic identity and culture', *Social Problems,* 41(1): 152-176.

Nazroo, J. (2005) *A Longitudinal Survey of Ethnic Minorities: Focus and Design.* Final Report to the ESRC and ONS, June 2005.

Nazroo, J. Y. and Williams, D. R. (2005) 'The social determination of ethnic/racial inequalities in health'' in M. Marmot and R. G. Wilkinson (eds) *Social Determinants of Health.* 2nd edition. Oxford: Oxford University Press, pp. 238-266

Office for National Statistics (2003a) *Ethnic Group Statistics: A guide for the collection and classification of ethnicity data.* London: ONS.

— (2003b) *Ethnic group statistics. A guide for the collection and classification of ethnicity data.* London: HMSO.

Peach, C. (2006) 'Islam, ethnicity and South Asian religions in the London 2001 census', *Transactions of the Institute of British Geographers,* 31(3): 353-370.

Perlmann, J. and Waters, M. (2002) *The New Race Question: How the Census Counts Multi-racial Individuals.* New York: Russell Sage.

Phinney, J. S. (1990) 'Ethnic identity in adolescents and adults: review of research', *Psychological Bulletin,* 108(3): 499-514.

— (1992) 'The Multigroup Ethnic Measure: a new Scale for use with diverse groups', *Journal of Adolescent Research,* 7: 156-176.

Phinney, J. S. and Alipuria, L. L. (1990) 'Ethnic identity in college students from four ethnic groups', *Journal of Adolescence,* 13: 171-183.

— (1996) 'At the interface of cultures: multiethnic/multiracial high school and college students', *Journal of Social Psychology,* 36: 139-158.

Phoenix and Tizard (1995) 'The identity of mixed parentage adolescents', *Journal of Child Psychology and Psychiatry,* 36(8): 1399-1410.

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.* Bristol: The Policy Press.

Ponterotto, J. G., Gretchen, D., Utsey, S. O., Starcuzzi, T. and Saya, R. (2003) 'The Multigroup Ethnic Identity Measure (MEIM): psychometric review and further validity testing', *Educational and Psychological Measurement,* 63: 502-515.

Purdam, K., Afkhami, R., Crockett, A. and Olsen, W. (2007) 'Religion in the UK: An overview of equality statistics and evidence gaps', *Journal of Contemporary Religion,* 22(2): 147-168.

Ratcliffe, P. (1996) 'Social geography and ethnicity: a theoretical, conceptual and substantive overview' in P. Ratcliffe (ed) *Ethnicity in the 1991 Census: Volume Three: Social geography and ethnicity in Britain: geographical spread, spatial concentration and internal migration.* London: HMSO, pp. 1-22.

Roberts, C. and Campbell, S. (2006) *Talk on Trial: Job interviews, language and ethnicity.* DWP Research Report 344. Leeds: Corporate Document Services.

Salt, J. (1996) 'Immigration and ethnic group' in D. Coleman and J. Salt (eds) *Ethnicity in the 1991 Census: Volume One: Demographic characteristics of the ethnic minority populations.* London: HMSO, pp. 124-150.

Salway, S., Platt, L., Chowbey, P. and Harriss, K. (2007) *Long-term Illness, Poverty and Ethnicity: Understanding the experiences of sufferers and their families.* Bristol: The Policy Press.

Schuman, H. and Presser, S. (1996) *Questions and Answers in Attitude Surveys.* Thousand Oaks, CA: Sage Publications.

Sellers, R. S., Cooke, D.Y., Chavous, T.M., Rowley, S.A.J. and Smith, M.A. (1998) 'A multidimensional model of racial identity: Assumptions, findings and future directions' in R. L. Jones (ed) *African American Identity Development: Theory, Development and Intervention.* Hampton, VA: Cobb & Henry.

Shields, M. A. and Wheatley Price, S. (2001) 'Language fluency and immigrant employment prospects: evidence from Britain's ethnic minorities', *Applied Economics Letters,* 8: 741-745.

Sillitoe, K. (1987) 'Developing Questions on Ethnicity and related topics for the Census', *OPCS Occasional Paper 36.* London: Office of Population Censuses and Surveys.

Sillitoe, K. and White, P. H. (1992) 'Ethnic group and the British Census: the search for a question', *Journal of the Royal Statistical Society, series A,* 155(1): 141-163.

Simpson, L. and Akinwale, B. (2006) *Quantifying Stability and Change in Ethnic Group*, CCSR Working Paper 2006-05. Manchester: University of Manchester:

Skerry, P. (2000) *Counting on the Census? Race, Group Identity, and the Evasion of Politics.* Washington, DC: Brookings Institution.

Smith, A. (1981) *The Ethnic Revival.* Cambridge: Cambridge University Press.

Smith, A. D. (1991) *National Identity.* Harmondsworth: Penguin.

Smith, J. P. and Welch, F. (1989) 'Black economic progress after Myrdal', *Journal of Economic Literature,* 27(2): 519-564.

Smith, R. (2004) 'Identity, interests, and the future of political science', *Perspectives on Politics,* 2(2): 301-312.

Smith, T.W. (2008) 'An experimental comparison of methods of measuring ethnicity', *Field Methods*, 20: 171-178

Thompson, V. L. S. (1991) 'Perspectives of race and race relations which affect

African American identification', *Journal of Applied Social Psychology,* 21: 1502-1516.

Tilley, V. (1997) 'The terms of the debate: untangling language about ethnicity and ethnic movements', *Ethnic and Racial Studies,* 20: 497-522.

Tourangeau, R., Rips, L. J. and Rasinski, K. (2000) *The Psychology of Survey Response.* New York: Cambridge University Press.

Travassos, C. and Williams, D. R. (2004) 'The concept and measurement of race and their relationship to public health: a review focused on Brazil and the United States', *Cad. Saúde Pública,* 20(3): 660-678.

Trejo, S. J. (1997) 'Why do Mexican Americans earn low wages?' *Journal of Political Economy,* 105(6): 1235-1268.

Verkuyten, M. (1991) 'Self-definition and ingroup formation among ethnic minorities in the Netherlands', *Social Psychology Quarterly,* 54: 280-286.

Voas, D. and Crockett, A. (2005) 'Religion in Britain: neither believing nor belonging ', *Sociology,* 39(1): 11-28.

Weber, M. (1978) *Economy and Society.* Edited by G. Roth and C. Wittich. Berkeley: University of California Press.

White, C. L. and Burke, P. J. (1987) 'Ethnic role identity among black and white college students: An interactionist approach', *Sociological Perspectives,* 30: 310-331.

Wimmer, A. (2008) 'The making and unmaking of ethnic boundaries: a multilevel process theory', *American Journal of Sociology,* 113(4): 970-1022.

Winters, L. and DeBose, H. (2003) *New Faces in a Changing America: Multi-racial Identity in the 21st Century.* Thousand Oaks, CA: Sage Publications.

Xie, Y. and Goyete, K. (1997) 'The racial identification of biracial children with one Asian parent: evidence from the 1990 Census', *Social Forces,* 76: 547-570.

Yancey, W., Ericksen, E. and Juliani, R. (1976) 'Emergent ethnicity: a review and reformulation', *American Sociological Review,* 41(3): 391-403.

Young, R. (1990) *White Mythologies: Writing History and the West.* London: Routledge.

Young, R. (2008) *The Idea of English Ethnicity*. Oxford: Blackwell.

Zuberi, T. (2001) *Thicker than Blood: How Racial Statistics Lie.* Minneapolis: University of Minnesota Press.