

**OPERATIONALISING THE COLLECTION OF ETHNICITY DATA IN STUDIES OF THE
SOCIOLOGY OF HEALTH AND ILLNESS**

Peter J Aspinall

Guy's, King's, and St. Thomas's School of Medicine,

Oak Lodge,

David Salomons Estate,

Broomhill Road,

TUNBRIDGE WELLS,

Kent TN3 0TG

Abstract:

In response to the burgeoning interest in ethnic health issues and related published research, a number of recent contributors have attempted to clarify or systematize the usage of overarching terminology like 'ethnicity', 'race', 'culture', and 'racism', including the development of guidelines. However, the operational problems of how to collect ethnicity data in studies of the sociology of health and illness have not been satisfactorily addressed. This paper explores conceptual issues, notably, the meanings of ethnic identity and ethnic origin/ancestry; methodological approaches, including which dimensions to collect, multidimensional versus global measures, and exclusive groups versus optional ethnicity; and also practical issues such as method of assignment. The approach calls for a stronger development of the theoretical understandings of ethnicity and work on how best ethnicity should be conceptualised and measured in the different approaches to explaining ethnic inequalities in health.

Keywords: ethnicity, race, terminology, classifications, inequalities, censuses, surveys

Introduction

Much concern has been expressed in the literature about the terminology used to describe ethnic groups (McKenzie and Crowcroft 1996, Bhopal, Kohli, *et al.*1997) and the usage of analytical concepts like 'ethnicity', 'race', and 'racism' (Bradby 1995), including 'race' and 'ethnicity' as variables in epidemiological research (Senior and Bhopal 1994) and in the varying approaches to ethnic inequalities in health (Smaje 1996, Nazroo 1998a). This comes at a time of burgeoning interest in ethnic health issues, stimulated by the unprecedented access enjoyed by the research community to statistics on the socioeconomic position of ethnic groups in Great Britain, including small-area Census data, and the high priority accorded to the reduction of inequalities in health in Government policy.

Given the interest generated by the 1991 Census question (Coleman and Salt 1996, Ratcliffe 1996) and a growing research literature (Rudat 1994, Modood, Berthoud, *et al.*1997), it is timely that a number of recent contributions have attempted to clarify or systematise the meaning and usage of terms and concepts. Bradby (1995) has focused on definitions of selected overarching terminology like 'ethnicity', 'racism', and 'racialisation' in studies of the sociology of health and illness, her call for the rejection of the term 'race' but continued use of 'ethnicity' (including that of the majority) according with a growing consensus of opinion, reflected in the fact that 'race' has been largely superseded by 'ethnic group' in the health literature. However, the subsumption of race by ethnicity and the utility of the concept of ethnicity continues to be contested in broader sociological discussions of race and ethnicity (Bourne 1985, Gilroy 1987, 1993, Jenkins 1994, and Smith 1996).

A further literature has focussed on the way some of these broader concepts are interpreted in research settings. For example, the *BMJ* - with the help of Drs Kwame McKenzie and Natasha Crowcroft - have responded to growing concern about the *imprecision* of definitions of ethnic group in published research by producing guidelines on the use of ethnic, racial, and cultural descriptions which recommend that a range

of information is best collected (Anonymous 1996). The pragmatic and parsimonious approach they adopt to the use of ethnic group in scientific writing accords with approaches developed in health-related fields (Singh 1997, Okazaki and Sue 1995, Jenkins and Parron 1995, Entwistle and Astone 1994). As they make clear, much writing on this topic has been rendered difficult to interpret or useless by the lack of clarity with which ethnic group is determined. The guidelines are necessarily general and it seems correct to link the concepts and classification to the researcher's hypothesis, rather than specify a list of terms or labels that would inevitably be contested.

Most recently and of particular value has been Nazroo (1998a)'s critical consideration of the varying approaches to ethnic inequalities in health, including genetic and cultural difference that frequently characterises epidemiological approaches, class (or material) disadvantage that is generally the focus of structural approaches, and approaches that focus on ethnic identity. Both Nazroo's and McKenzie and Crowcroft's contributions help to clarify the context in which investigators must make the necessary choices about how to collect ethnic group data. The operational problems of doing so in studies of the sociology of health and illness are the focus of this paper. The issues addressed are conceptual (meanings of ethnic group and ethnic origin/ancestry), methodological (which dimensions to collect; multidimensional versus global measures; and exclusive groups versus optional ethnicity), and practical (method of assignment).

Conceptual problems confronting medical sociologists in collecting ethnicity data: Ethnic group versus ethnic origin

Two methodologically distinct, albeit broad, but mutually implicated concepts of ethnicity are generally used in a health and illness context. Firstly, there is self-perceived *ethnic group* which is based on a conception of social group membership. This concept is widely used in survey research, including the 1991 Census, the harmonised set of Government social surveys, and the Fourth National Survey of Ethnic

Minorities (Modood, Berthoud, *et al.* 1997). This conceptualisation is different from what Bradby (1995) terms an 'externally allocated' definition relating to 'a common ancestry or place of origin', regardless of the individual's own identification, although for some respondents a common ancestry may be the factor that most strongly contributes to their ethnic group identification. Indeed, in a sample of 489 patients, amongst factors deciding a person's ethnic group 62% thought that 'it is passed on from your parents' (Pringle and Rothera 1995). Clearly, *ethnic origin* focuses the question back in time and conveys an historical and frequently geographic context. The strength of focus on origins or descent depends largely on the wording of the question. For some, ethnic origin and ethnic group are interchangeable terms while, in other contexts, the former may access family origins or actual ancestry. The first (Rudat 1994) and second (Johnson, Owen *et al.* 2000) Black and Minority Ethnic Group (BMEG) surveys asked respondents to describe their 'race or ethnic origin'. 'Ethnic origin' was also used in a survey of the health and lifestyles of the Chinese population (Sproston, Pitson *et al.* 1999). Family origins, a more explicit focus on origin, was adopted by the Fourth National Survey, responses being highly related to perceived ethnicity (ethnic group) for most categories. An even more discriminating question has been proposed but not yet tested by Berthoud (1998), which asks for the ethnic origins of the respondent's mother's family and father's family.

Other studies and surveys have used the 'ethnic or cultural group(s)' to which the 'person's ancestors belong' (1991 and 1996 Censuses of Canada), an ancestor being defined as 'someone from whom a person is descended and is usually more distant than a grandparent' (Statistics Canada 1996); and a person's 'ancestry or ethnic origin' (1990 and 2000 US Censuses), where (in 1990) ancestry was 'the person's ethnic origin or descent, "roots", or heritage' but also might refer to 'the country of birth of the person or the person's parents or ancestors before their arrival in the United States'. 'Ancestry', with its connotation of roots and descent or lineage, is a concept not generally used in a Great Britain context (compared with ethnic origin), although used in the instruction to the 1991 Census question and tested in the Office for National Statistics' development programme for the 2001 Census.

Surveys and censuses have also asked for country of birth of respondent (Dunn and Fahy 1990) and his/her

parents (McKenzie, van Os, *et al.* 1995, 1971 Great Britain Census, 1971 Canadian Census, 1998 Canadian National Census Test, and 1996 Canadian General Social Survey) or grandparents (Rudat 1994, 1998 Canadian National Census Test). However, these questions provide only a supplement or proxy to ethnic origin or ancestry questions, but can effectively identify first-, second-, and third-generations of a community when country of birth is collected for all three (respondent, parents, and grandparents). The two types of question are quite different conceptually. Country of birth questions collect data for geographically defined countries of birth, while ethnic origin/ancestry questions collect data for ethnic groups which may be located within one or many countries, with some countries being home to more than one distinct ethnic or cultural group. The two variables represent different, but in some cases linked, sub-sections of the population. For relatively new ethnic communities, birthplace of parents can be a useful marker of ethnicity. However, its utility diminishes with the length of time that groups have been established, third, fourth, and other generations not being identified by this method. Also, the birthplace of parents cannot provide information for distinct ethnic groups which are not associated with a specific country of birth, such as Jews or Kurds, or where changes to the state conceal ethnic origins, for example, Serbs and Croats where country of birth of parents is Yugoslavia; or persons of East European ethnic extraction born in Germany to displaced Second World War refugee parents.

These various conceptualisations of ethnicity can be placed along a hypothetical continuum of stability, with ethnic self-identification using open response likely to be the least stable, producing high gross error rates on retest, and operational definitions (those embodying criteria that are used to identify and classify individual members of a set) of family origins, such as Berthoud's question, the most stable, with low gross error rates. For example, questions such as 'What ethnic group do you identify with?' which have sought respondents' free text and unprompted response, have yielded non-standard data. In a survey of 109 general practice patients in New Zealand (Kljakovic 1993), the information was difficult to obtain in 22 per cent of patients, either because they did not understand the question or thought it was odd to be asked, and responses produced 35 separate self-identified categories. In Pringle and Rothera (1995)'s survey of general

practice patients in Lincolnshire and Leicestershire, 14.2 per cent of respondents did not know what an ethnic group was and a further 9.9 per cent were unsure which group to belong to. The question produced 56 different self-identified responses. Even where classifications are used to capture self-perceived ethnic group, the responses obtained are inconsistent on retest for some groups. For example, the 1991 Census Validation Survey (ONS [SSD] 1996) (the follow-up survey carried out by OPCS to check the accuracy of the data collected by the Census) revealed a *gross error rate* - that is, the proportion of times the response on the Census form was not the same as that given in the CVS interview - for minority ethnic groups (excluding those who answered 'White' in both the Census and the CVS, which were the vast majority of the 13,080 persons participating in the CVS) of 13.2 per cent. Inconsistent reporting was not uniformly manifested across the different ethnic groups. While 99.6 per cent of those who identified as White in the 1991 Census so identified in the CVS, the proportion consistently identifying in the Black groups was 88.0 per cent, in the Indian subcontinent groups 98.7 per cent, and in other groups, 78.1 per cent. Although the pattern of switches between the options on the form was not reported, most involved the final 'other' group. Comments of these respondents (especially mixed parentage persons) showed that they found the categories hard to understand or were unsure where they fitted in.

Ethnic *origin* or ancestry is commonly perceived as being inherently more stable and less labile than ethnic *group* and the experience of census agencies suggests that, even when open response is used, albeit with the provision of examples, data quality is generally high. For example, the non-response rate in the 1998 Canadian National Census Test was 4.5 per cent (similar to about 5 per cent in the 1996 Canadian Census), the rate of invalid responses being minimal (0.2 per cent). However, questions that combine the concepts of 'identification' and 'ancestry', for example, the USA 1980 and 1990 census ancestry questions (done in the instruction guides), are problematic (Liebersohn 1993). Higher levels of inconsistent reporting of ancestry in these censuses have been reported, with non-response rates of 10.2 per cent and 9.6 per cent, respectively (Farley 1991). This combination is maintained in the US 2000 Census, where guidance informs respondents that 'This *self-identification* approach recognises that strong *ethnic identity* is not

limited to just first- and second-generation immigrants' (emphasis added). A question assessed in the 1993 Australian Census Test, asking if the person *identifies* with an *ancestry* different from their country of birth, produced a non-response rate of 12.2 per cent and follow-up found the combination particularly confusing for respondents (Australian Bureau of Statistics [ABS] 1994). The ABS concluded that census data on ethnic origin or ancestry would be subject to serious problems of interpretation, validity, and stability, based on accumulated evidence that the concept is not clearly understood and consistently reported by the community (ABS 1994). Open response ancestry questions are also subject to strong 'example effects', the inclusion of 'Canadian' as an example in the 1993 Canadian Census increasing the count of Canadians from 763,000 (1991) to 5,193,000 (Renaud and Costa 1994). The use of classifications with predesignated categories increases the stability of conceptualisations of identity and origins/ancestry. However, 'established best practice' may propel investigators to ask respondents to indicate their own personal definition of their ethnic origin without prompting and this approach can be successful when additional survey evidence is used to map such responses to a set of groupings (Johnson, Owen, *et al.* 2000).

Clearly, the extent to which these findings of censuses and surveys in North America and elsewhere are generalizable to the British context may be limited by the differing cultural contexts, especially the specificity of each country's historical processes of ethnogenesis. Nevertheless, cognitive research by the ONS (Rainford 1997) has shown that ethnic origin is seen as different from ethnic group by some respondents, yet is still characterised by a range of interpretations. Moreover, respondents have different understandings of 'ancestry', including their parents, their grandparents, or earlier generations. Not knowing how far back in time to go meant that some Black British and Black Caribbeans interpreted it as ancestry in Africa rather than origins in the Caribbean. Without the context of the question being made clear by guidance and pre-designated options (and possibly examples), concepts of ethnic origin are likely to be as difficult to operationalise as ethnic group. Perhaps their attraction and utility lies in their ability to allow researchers to capture white subgroups in Britain with origins in Europe, such as the Portuguese, Poles,

Italians, Irish, Greeks, Turks, and other minority groups and to better define the Black and mixed heritage groups.

Clearly, the conceptualisation of ethnicity to be used will depend on the context of the research. Frequently, especially in epidemiology, measures of ethnic origin are needed that will produce well-defined populations that are representations of the 'full size' of a community or population group rather than a potentially biased set caused by selective attribution of ethnicity based on individuals' perceptions of *identity*. For example, although limited as a method of allocating ethnicity, registered country of birth was used to investigate reported health, lifestyles, and use of health care amongst immigrants in the Netherlands (Reijneveld 1998), thus avoiding such bias. Parents' country of birth, as recorded in the Longitudinal Study, is a better strategy for investigating morbidity and mortality amongst second generation Irish living in Britain (Harding and Balarajan 1996) than a measure based on self-perceived ethnicity. Similarly, the use of ethnicity as a primary screening tool for sickle trait would require the use of a stable concept like ancestry (Zeuner *et al.* 1999), whilst the ethnic monitoring of the uptake of haemoglobinopathy screening services needs to consider culture and religion (Chapple and Anionwu 1998). Where ethnic group is the concept of choice, a more considered approach to data collection is needed. As Nazroo (1998a) has cautioned, the frequently crude approach to the allocation of individuals into such groups that characterises much epidemiological investigation risks construing ethnicity as a natural and fixed division between social groups that ignores its contextual and labile characteristics and invoking explanations based on assumed and unmeasured cultural or genetic factors. A similar point is made by Kaufman, Cooper, *et al.* (1997) who argue that inadequate statistical attempts to control for socio-economic status and their inappropriate interpretation reflect a social ideology that accepts genetic and biological differences as the 'right answer'. In spite of race's untenable status as a biological entity, such deconstruction may not have fully wiped out this legacy in terms of modes of explanation. There is, however, an equal danger of dismissing any role (as opposed to a residual role) for genetic explanations for ethnic patterns of health. In certain cases there may be a partial and geographically mediated correlation between human genetic

diversity and socially defined racial or ethnic categorisations that should be accorded importance.

The limitations of the ethnic data available in existing surveys and research

(i) *Insensitive categories:* Many current surveys, including the harmonised question set in Government social surveys, utilise the 'ethnic group' question that was new to the 1991 Great Britain Census. The question was developed from empirical criteria and justified on pragmatic grounds including its acceptability by the public (Sillitoe and White, 1996). The categorisation was based substantially on colour (Black, White), national origins (Indian, Pakistani), and geographical origins (Caribbean, African), the conceptual approach encompassing both self-identification and descent/ancestry.

Where *ethnic identity* is the concept of choice and an unprompted open response question is not used, the accurate measurement of this concept depends heavily on the categories used. 'Questionnaire ethnicity' is vulnerable to the 'list effect'. People frequently have several ways of describing their identity, Kumari (1998) arguing from focus group research with adolescent Asian pupils that it is possible for individuals to subscribe to dual ethnicities through developing the ability to shift between cultural frames of reference. The choice of identities may depend on whether the context is personal (where a term may be used that is recognised by members of the group) or official (a description constituted in its recognition by observers), as Ifekwunigwe (1997) has shown for mixed parentage individuals. These two kinds of abstraction are brought into focus by the role of the state in the creation of what some commentators have called 'fictive unities', categories at the level of, say, 'Indian' which, although only partially and situationally meaningful to people thus labelled, are nonetheless employed by them in their dealings with officialdom. Pringle and Rothera (1996) show that respondents' personally constructed identities expressed as unprompted self-descriptions are different from the choices they make amongst 'unities' in a Census classification. When general practice staff asked patients to give their ethnic group without any prompting 851 (96.7 per cent) offered a response. Significantly, when prompted with OPCS's classification 855 (97.2 per cent) of patients

selected a category, but in only 236 (27.7 per cent) of valid cases was this an exact or very close match to the self-reported description. Rankin and Bhopal (1999) have also questioned the utility of 1991 Census labels in questions about ethnic self-identity. In a South Tyneside study which asked respondents for their ethnicity first using the census question and then as an open choice, around only two-thirds of respondents choosing 'Indian' (81/130) and Pakistani (29/43) from the 1991 Census options described themselves unprompted in these ways.

Clearly, the concept of ethnicity which the Census seeks to capture is much broader and more summary than just that of identity. Attempts to assess subjective descriptions of identity show that responses in open text are too complex and heterogeneous for any uniform ethnic terminology to be applicable and the kind of solution adopted by Johnson, Owen, *et al.* (2000) is best followed. Further, question development programmes for Great Britain censuses have been primarily observer-led, where respondents assume an essentially 'reactive' role. This approach and the necessary limitations upon choice, however, sometimes result in a discordance between the category options and respondents' preferences, the 1991 Great Britain Census revealing that one in four members in ethnic groups other than 'White' eschewed the seven pre-designated options and wrote in a self-description in one of the two free-text fields (Aspinall 1995). Research by the Bureau of Labor Statistics (Tucker, McKay, *et al.* 1996) in the USA on large population samples (60,000 households) has demonstrated that the different 'cultural and national origin' groups have specific preferences for racial terms that are not always the same as those used by officialdom. The need in censuses for some unitary measure of ethnic relations in the country at large may result in the use of ideal-type constructs like 'Hispanic', derived from research on clusters of objective characteristics.

Additionally, importance needs to be accorded to the fact that ethnic identity changes through the life-course, making its capture in cross-sectional surveys a measure of 'state' rather than 'stage'. Empirical findings for Britain are limited, although the linking of the ONS Longitudinal Study cohort members with their 2001 Census schedules will enable the stability of self-ascribed ethnic group to be assessed over a

decade. As Frable (1997) has noted, such longitudinal research is rare but crucial in order to follow identity paths. The main body of empirical work is based on identity development in adolescents and adults using Cross (1978)'s black identity model and Phinney (1989)'s model of ethnic identity development, delineating a progression from an unexamined ethnic identity, through a period of exploration, to an achieved or committed ethnic identity. The genre of autobiographical writing (Thompson and Tyagi 1996) has also contributed a growing body of in-depth findings. Identity changes over the life-course may happen for many reasons, including the ebb and flow in popularity of some identities, marriage/union formation and dissolution, exposure to social and political movements, and the experience of discrimination. Longitudinal research designs are essential to measure this fluidity, the form it takes, and how it may be interpreted.

(ii) Problematic concepts: For purposes such as access to health and social care services, it may be argued that 'ethnicity' as a term describing broad historical processes of colonialism, migration, and discrimination - as reflected in the ethnic categories of the Census and large social surveys - is useful. Disadvantage mediated by discrimination may be impervious to fine distinctions in personal ethnic identification and better reflected in broad, *socially constructed* 'race' labels like 'White', 'Black-Caribbean', 'Black-African', and South Asian categories like 'Indian' that substantially reflect societal perceptions of ethnic differences. These summary Census categories provide a point of access to the longer-term historical processes that have influenced and shaped the nature of ethnic relations, in contradistinction to terms of ethnic self-identification which are immediate and idiosyncratic yet are potentiated by these wider processes. Indeed, a focus on ethnicity as personal identity in analytical studies may obscure the influence of these broader processes of discrimination and disadvantage so relevant to health. This may happen when persons choose terms of identity which do not provide clues to their phenotype (appearance) or where classifications vary between populations to mask this.

For the purposes of social surveys, ethnic monitoring, and other such general contexts, classifications built

on a broad conceptual base (rather than a narrow and exclusive focus on identity) may be needed. However, one of the drawbacks of using categorisations built on discourses relating to the consequences of historical processes of colonialism, such as those in the Census, is that they tend to carry strong racial connotations, based on the binary of white and black groups and of 'pure' races. The state's justification appears to be that these dimensions of similarity/difference are of most use because they are prominent or easily identifiable, obvious to the perceiver, and the basis for colour discrimination, a justification that is shared by the social cognition tradition in psychology (Allport 1954, Hamilton and Trolier 1986, Fiske and Neuberg 1990). Ethnic minorities concealed in the white group, such as the Irish, the Turks, Greeks, and Cypriots do not easily fit into such categorisations. Moreover, most of these categorisations require respondents to identify with a single group and fail to meet the needs of those with more complex ethnicity, including the large and increasing number of persons of mixed parentage, estimated to comprise 11.5 per cent of the ethnic minority population according to pooled Labour Force Survey data for 1989-91 (OPCS, 1992). Anderson (1991), for example, has asserted that: 'The function of the census is that everyone is in it, and that everyone has one - and only one - extremely clear place. No fractions'.

The danger of the routine use of such Census categories in studies of disadvantage mediated by discrimination and other processes is the reification of group labels and the support their usage lends to the view that groups like 'Indian' or 'Black African' are homogeneous wholes. Rudat (1994) and Elam and Chinouya (2000) have observed that Black Africans living in the UK are an extremely heterogeneous group. The resulting representations of such homogenisation accentuate within-category similarities and between-category dissimilarities and such distortions are carried into analytical studies that use Census categorisation for denominators. The resulting accounts contain biased perceptions of inter-group characteristics and can have dramatic consequences for excluded groups.

Based on the fact of colour-discrimination, many official collections and social surveys now ask people to assign themselves to such ethnic group classifications on the basis of a question combining colour and

country/region of origin. The purpose of such questions is to identify members of what have been termed 'visible minorities' (Pendakur 1993), 'communities (or people) of colour' (Braithwaite, Bianchi, *et al.* 1994; Yuen, 1997, for arguments against such usage), and 'ethnic or racial group' (the 1991 Great Britain Census). Indeed, the heavy emphasis on race by the 1991 Great Britain Census has been constructed by the Canadian sociologist Isajiw (1993) as an attempt 'to enforce the external boundaries given by the dominant group...done in the name of combatting discrimination and prejudice', although, of course, race in some contexts can also be an internal boundary, as demonstrated by the way 'black' has been reclaimed in both North America and Britain as a positive identity. Similarly, Ang-Lygate (1997) sees these categories as legitimised by European processes of colonization: 'Far from being a progressive step towards eliminating ethnic discrimination...this kind of structured racialization actually highlights the racist legacy of a neo-colonial society obsessed with categorising those who are perceived to be immigrant outsiders'.

The role of internal, identity-generated forces and their outcome in the way individuals assign themselves into an ethnic group does have merit as an alternative focus in some particular types of study, including those seeking to explain ethnic inequalities in health. This approach focuses on the active and constructive differentiation from others on the basis of continuities (such as history, culture, and parents' immigration identity), in contradistinction to definitions given by the dominant group or agencies of the state, 'difference-as-diversity' rather than 'difference-as-opposition' (Ang-Lygate 1997). This 'internal definition', one's sense of belonging to a group, provides a meaningful approach in the context of the growing interest in the relationship between levels of social cohesiveness in a society and inequalities in health. One body of research has focused on variations in risk of adverse health outcomes in minority groups with respect to their relative size or density locally, an example of ecological effect modification. Several studies have suggested that the risk of mental illness and suicide varies inversely with ethnic density in local areas (Wechsler and Pugh 1967, Rabkin 1979, Halpern 1993, Neeleman and Wessely 1999, Halpern and Nazroo 2000). The literature on psychological adjustment provides similar findings. For example, in a study of racial socialization and racial identity attitudes amongst adolescents of mixed African-Caribbean/White

parentage, Fatimilehin (1999) tentatively concludes that dual identification and multicultural environments may be associated with a positive sense of racial identity. Tizard and Phoenix (1993) showed that mixed parentage teenagers who attended multiracial schools were more likely to have a positive racial identity.

These research findings feed into the wider literature on increased social cohesion (and its protective effects) in neighbourhoods with large and relatively homogeneous minority populations (Smaje 1995a). They underline the need for some measure of affiliation to a group or 'involvement in and commitment to a self-conscious *community* of some kind' (Ballard 1996) to capture this sense of community belongingness. This approach will require research on how participation in or affiliation with these groups of communal ethnic loyalty can be best represented in terms of identifiers. Clearly, the way collectivities define themselves may be highly context-dependent. For example, while the shared experience of discrimination may create commonalities in some settings, some groups may become united as a single political force to gain official recognition by the state or to obtain access to entitlement programmes and other state resources. Others may form local alliances around single-issue objectives. Further, the state's role as official labeller may forge or foster group identities through the process of administrative allocation. Petersen (1987) even suggests that 'few things facilitate a category's coalescence into a group so readily as its designation by an official body'.

(iii) Varying methods of assignment and allocation: Self-assigned ethnicity, obtained by self-reporting or direct interview, is now the basis for the collection of nearly all ethnic group data, including that in censuses. However, it may be difficult to monitor when the sampling units are households and the data is collected by self report for household members, some censuses and census testing programmes reporting that around half of questions are filled in by someone else (Cornish 1993) or that respondents 'mainly' choose to answer on behalf of other household members (Rainford 1997). How the ethnicity of children is assigned in this context may be consequential. Research by Khoo (1991) has shown that parents are uncertain about giving their children's ancestry when the parents have differing or multiple ancestry.

Hodes, Creamer, *et al.* (1998) found differences in how parents described adolescents, especially in relation to peer culture, underlying the need to specify whether adolescents or parents should respond to ethnicity questions.

Observer-assigned ethnicity is no longer regarded as an acceptable method of assignment and has long been eschewed by government surveys and censuses. It is clearly inappropriate for conceptualisations of ethnic group and is regarded as an unacceptable method for healthy adults by nearly all general practice patients (Pringle and Rothera 1996). There remain only a few contexts in which such assignment can be justified from a conceptual viewpoint. They include the need to obtain the wider society's perception of ethnicity, as exemplified by Singh, Croudance *et al.* (1998)'s study of risk of compulsory admission, in which a measure 'specifically of *socially perceived ethnicity*' or the way ethnicity is constructed in everyday social interaction was required. The NHS Executive judged proxy reporting of ethnic group in hospital inpatient settings to be acceptable only for the very young and permanently confused psychiatric patients (NHS Executive 1994). Thus, the inclusion of observer-assignment in recent guidelines on ethnicity, race, and culture (Anonymous 1996) is problematic.

In addition to assignment, other contextual factors may influence responses to ethnicity questions and the quality of interview and questionnaire data. A number of investigators emphasise the importance of using interviewers of the same ethnic group (Warnecke, Johnson, *et al.* 1997, Modood, Berthoud, *et al.* 1997) to minimise the cultural distance between respondent and interviewer. In the 4th National Survey 87 per cent of all respondents in the minority sample were interviewed by a member 'of the same broad ethnic group as themselves', for which purpose 'a Pakistani being interviewed by an Indian would count as a broad match, but not if the interviewer was a Caribbean or a Chinese'. For the 1st and 2nd BMEG surveys and the health and lifestyle survey of the Chinese, language- rather than ethnic-matching was undertaken where the need to conduct interviews in the respondents' own language was identified. The strategy of 'ethnic matching' is controversial. Firstly, if researchers match on the basis of broad, socially-perceived 'race' groups like

'Black' or 'Asian', this strategy must be justified (for example, on grounds of access, language, or identity), but it does not constitute ethnic matching and, arguably, reproduces a racist taxonomy. Secondly, the justification for ethnic matching is based by advocates on an edict of epistemological privilege which holds that a single 'truth' or 'reality' exists for people of a particular ethnic group, that respondents prefer this method, and that, as a result, more sensitive and accurate information is collected.

However, the frequent assumption that ethnically homophilous interviews generate the highest response quality, a perspective that has been investigated with respect to other characteristics such as gender (Huddy *et al.* 1997), age and social position, is not supported empirically. Although there is a scarcity of published research that focuses on this issue in the context of ethnic relations in Britain, a general finding in the wider research literature is that survey respondents significantly bias their responses to items pertaining to the interviewer's culture (Shuman and Converse 1971, Hatchett and Shuman 1975/76, Campbell 1981, Cotter, Cohen, *et al.* 1982, Reese, Danielson *et al.* 1986, Webster 1996), the interview effects in some cases being large enough to justify matching interviewer and respondent with respect to ethnic background. Race-of-interviewer effects in telephone surveys have been found to be similar in magnitude to those reported in personal surveys (Cotton, Cohen *et al.* 1982, Reese, Danielson, *et al.* 1986). However, few such effects are found for non-ethnically/racially sensitive questions (Kish 1962, Welch *et al.* 1973, Schaeffer 1980, Weeks and Moore 1981, Groves and Magilavy 1986, and Webster 1996).

In the intensive interviewing setting of qualitative research, too, the assumption made by researchers until recently (Groves 1989) that matching interviewers' and respondents' gender, race/ethnicity, and age would help establish the minimum rapport necessary for validity has been questioned. Harkess and Warren (1993) argue that the effects of familiarity upon validity frequently appear problematic and that it might be more important to scrutinise the kind of personal knowledge interview respondents have of one another in terms of effects on validity of results. Similarly, in a survey of first time mothers where ethnic matching was not used, 71 per cent of respondents said that they had no preference for the ethnicity of the interviewer (Phoenix 1994). In intensive interviews with 100 West Indian families, Rutter, Yule, *et al.* (1974) also

found no evidence that the rapport established between interviewees and white interviewers was less good than that with black interviewers, yielding data of similar validity and quality. Indeed, black interviewers encountered a 22.8 per cent refusal rate, compared with 2.9 per cent for white interviewers. Ethnic matching may be justified in some contexts, notably, where issues relating to racism and discrimination are being discussed (Douglas 1998). However, the argument for ethnic matching as a *general* research strategy based on validity appears unsustainable and there are other drawbacks, including the fact that such an approach risks marginalisation of black researchers (Rhodes 1994).

Assignment of ethnicity by *validated name analysis* has been used to construct sampling frames in a wide variety of contexts (Balarajan, Adelstein, *et al.* 1984, Ecob and Williams 1991, Hoare, Thomas, *et al.* 1994, and Martineau and White 1998). The method has been demonstrated to be reliable through ascertainment of sensitivity, specificity, and predictive values in studies of South Asians (Nicoll, Bassett *et al.* 1994, Kolev, House *et al.* 1996), as Hindu and Sikh names are easily identifiable, and groupings based on religion and region of origin can be defined by this method. However, it is not usually possible to assign people to country of origin using Muslim names (Martineau and White 1998) since many are not specific to a certain country. Naming algorithms have been used to identify members of the Chinese group (Choi, Hanley, *et al.* 1993, Harland, White, *et al.* 1997, Freeth, Elliott *et al.* 1997, Pitson, Sproston, *et al.* 2000) but up to a fifth of sample members in a British Columbia study were not distinguished by this method (Coldman, Braun, *et al.* 1988). Name information has also been used to identify the Jewish population in Britain (Miller, Schmool, *et al.* 1996) and persons of patrilineal Irish descent in Scotland (Abbotts, Williams, *et al.* 1998). However, there are drawbacks to validated name analysis. When compared with a reference standard, a computer programme for the identification of names of South Asian origin was found to have a sensitivity of 90.5 per cent but a high false positive rate (36.8 per cent), making the programme alone not an adequate single identification strategy (Cummins, Winter, *et al.* 1999). For studies that require Census denominator data for the construction of rates, the method is unsatisfactory because of the differences in allocation methods. Also, increasing rates of intermarriage between some of

the ethnic groups, for example, between Chinese women and White men (of 324 Chinese women in unions in the 1991 Census 1 per cent household SAR, 25 per cent were married or cohabiting with a white partner), creates selection biases (against the women and offspring of such partnerships) in the outcome of naming methods, especially for women who adopt Anglicized first names. While this method of external allocation may be useful for selected groups where alternative data sources are unavailable or incomplete and of poor quality (such as hospital inpatient ethnic monitoring data), for the exploitation of existing datasets, or for sample accrual in areas with a small ethnic minority population (rarely included in studies), without follow-up it is not a satisfactory measure of ethnic identity.

Given that ethnicity is highly sensitive to contextual factors, importance must be accorded to accurate reporting of the method of acquiring ethnicity data, including mode of administration (self-administered or interviewer-administered self-report, observer rating, proxy report by relatives or others, analysis of names, or a combination of methods), organisational issues including use and form of ethnic matching, interview formats (such as face-to-face, computer-assisted, and telephone), as well as information on the conceptual basis (such as ethnic origin), classification, and question instruction used. Where appropriate tests of reliability should be undertaken to establish that ethnic group data is comparable, including *test-retest* with the same subjects and *inter-observer (interviewer) reproducibility*. However, a recent review of seventy studies of ethnic identity (Phinney 1990) found that less than a fifth reported the reliability of measures, which was often so low as to challenge findings, and in none was there evidence of test-retest reliability.

How the problems associated with existing approaches can be circumvented

(i) *An 'ethnic' rather than 'racial' approach:* Debates around ethnic group categorisation suggest the need for a different approach, based on self-identified ethnic rather than racial categories. The term 'black' for some groups has emerged as a positive self-identification that demarks a set of culture patterns and values

that delineate internal boundaries. In Britain 'Black British' has gained currency as an identity for some members, some 60,000 persons in the 1991 Census in Great Britain describing themselves in free text as 'Black-British' (about a quarter of those selecting 'Black-Other'), foregoing the category options of 'Black-Caribbean' and 'Black-African'. Similarly, in a recent health survey¹ 17 per cent of 722 African-Caribbean respondents (36 per cent in the population aged 16-29) described themselves in an open response question as 'Black British', compared with 51 per cent as 'Afro-Caribbean', and 27 per cent 'West Indian'. In the Fourth National Survey (Nazroo 1997) 10 per cent of respondents with ethnic family origins in the Caribbean described their ethnic group as 'Black British'. Similarly, of 196 Africans 46 per cent described themselves as 'Black African' but 18 per cent selected 'Nigerian' and 12 per cent 'Ghanaian' in open response. Ethnic differences within the white group also require acknowledgement through the inclusion of categories like 'Irish', 'Scottish', 'Welsh', 'English', 'Turkish', 'Cypriot', 'Polish', and 'Jewish', descriptions readily used in open-response questions.

While the way ethnicity is measured will depend upon the research questions asked, the use of categories based on ethnicity or internal identification can lay claim to conceptual integrity and sensitivity to the way members of the different ethnic groups describe themselves. This approach has received some acknowledgement in the revised ethnic group question for the England and Wales 2001 Census, in which the five main ethnic group options ('White', 'Mixed', 'Asian or Asian British', 'Black or Black British', and 'Chinese and Other ethnic group') are broken down into several 'cultural background' groups (Her Majesty's Government 1999). The acknowledgement of the growing use of the terms 'Black British' and 'Asian British' is desirable, albeit incorporated into the group labels of the Black and Asian ethnic groups to enable ONS to identify the origins of these groups in Africa or the Caribbean and the South Asian countries. The needs of the substantial number of people who wrote in a mixed origins description in the 1991 Census are now met by the inclusion of the subdivided 'Mixed' group. However, the limited breakdown of the white group into 'British', 'Irish', and 'Any other White background' does not satisfactorily address the 'myth of homogeneity' of this group.

While the abandonment of race and the adoption of ethnicity works satisfactorily in Britain, largely because of the 'racialisation' of the concept of ethnicity here as demonstrated in the censuses, the situation in the USA (and the Caribbean) is different. The US censuses have traditionally collected data on ethnic origin/ancestry *and* race/national origin groups. However, with respect to the collection of more specific data for 'ethnic health' research, a substantial body of opinion favours the continued use of socially defined race categories as well as ethnicity. The view is that the privileging of cultural differences by the use of ethnicity diverts analytical attention from those issues of structural constraint and power that so shape the experiences, socio-economic position, and social relations of 'racial' groups. Harrison (1994) has written: 'Given the burden of US racism, the racialized ethnicity of, for instance, black Americans cannot be erased by the popular force of mere semantic ethnicization, the upgrading of black social identity by the "African American" self-definition'. Smith (1996) makes a similar critique of the use of the term 'ethnic' to discuss problems of poverty and inequality in the United States, arguing that it 'tells us nothing about the real processes involved or about the value commitments that underlie various courses of political action in the modern world'.

(ii) *Ethnic categorisation that reflects diversity:* The ethnic diversity of Britain today is highly complex and the broad and crude categories used in the Census may be working to conceal such diversity. There may now be a need for different types of information that reflect a move away from broad categorisations. For some groups, such as those originating in South Asia, religion is a primary identifier and, arguably, more useful than the 1991 Census categorisation by a South Asian *national* identity (Indian, Pakistani, Bangladeshi). However, the fact that religious identifications are expressed within the context of ethno-national identifications (for example, a 'Hinduism' exclusive of Islam as definitive of Indian identification) suggests the need for religious identifiers in addition to, rather than instead of, national ones. This point is also illustrated by the emergence of a nascent Kashmiri Muslim identity.

Jacobson (1997) presents the interesting finding that amongst young British Pakistanis religion is a more significant source of social identity than ethnicity, manifest in the *particularistic* reference of ethnicity against the *universal* relevance of Islam and the fact that the social boundaries which encompass experiences of religious identity are 'pervasive and clear-cut in comparison to increasingly permeable ethnic boundaries'. Another study (Bhopal 1998) shows that South Asian women's experiences of religion are heterogeneous and structured by marital status and education, one group wishing to maintain and reinforce their religious identity while single and/or highly educated women decide not to participate in religious and traditional practices. An investigation of the acculturation processes of young people whose parents were migrants from the Indian Subcontinent (Ghuman 1997) reveals important differences by gender and religious group. Religion and region of origin - expressed in the main ethno-religious groups of Pakistani Muslims, Bangladeshi Muslims, Punjabi Sikhs, Gujarati Hindus, and Indian Muslims - are relevant to many aspects of health related behaviour, including alcohol use and diet (Balarajan, Adelstein, *et al.* 1984), smoking (Nazroo 1997), the utilisation of services (for example, attendance levels in hospital outpatient settings) (Gatrad 1997), and health experience (Nazroo 1998a).

Given the lack of a religion question in most Government social surveys and the 1991 Census - but its inclusion in ONS's 1997 Census Test and in the 2001 Censuses for England and Wales (Her Majesty's Government 1999) and Scotland (Scottish Executive 2000) - and also of an Arab or Middle East category, the Muslim community in Britain is largely invisible in official statistics and survey findings in spite of the group's experience of 'cultural-racism' (Modood, Beishon *et al.* 1994) and its growing numbers in absolute and relative terms (Anwar 1993), Muslims now accounting for about half of the population in groups other than white. Religion and ethnicity are also relevant to the identification of the Jewish community in Britain, recent survey research (Miller, Schmool *et al.* 1996, Goldberg and Kosmin 1997) emphasising the need for both religious and ethnic identifiers for this now widely dispersed group in Britain, a substantial proportion of whom are assimilated, outmarried, or distanced from the community, to meet the needs of users. 'Jewish' was identified by category or example as an 'ethnic or cultural group' in

the 1991 and 1996 Canadian censuses, emerged as a well defined group in the 1991 Australian ancestry census tests, and a case was made for its inclusion in the Great Britain 2001 Census (Schmool 1998). The Irish are another group strongly differentiated by religion and evidence that religion may be salient in structuring health differences amongst the Irish in Britain (Abbotts, Williams, *et al.* 1997) underlies the need to identify the religious composition of the group.

Finally, studies may require the collection of information on culture as a collective attribute or on individual dimensions such as language, diet, and customs as studies have shown that there is no simple link between minority cultural behaviour and minority ethnic identities. While studies of Irish adolescents in England (Ullah 1987) and of Armenian Americans (Der-Karabetian 1980) revealed a close relationship between ethnic self-definition and indices of ethnic group behaviour, Garcia (1982) found a negative relationship between ethnic self-identification and preference for various ethnic practices. Also, Hutnik (1991) found that of the young British Indians who said that they thought of themselves as Indians, many were not Indian in their cultural behaviour (as defined in the study) while some who were culturally Indian did not think of themselves as Indians. Modood, Beishon *et al.* (1994) conclude: 'There is...a genuine gap between self-identification and cultural practice. An ethnic identity...cannot simply be read off from a description of distinctive cultures'. This finding emphasises the importance of the conceptual distinction drawn by Jenkins (1996) between 'nominal identity' (the name of an identity) and 'virtual identity' (the experience of the identity, what it means to bear it): the virtual identity can change while the name stays the same, the experience may stay relatively stable while the name changes, or both can change.

(iii) Addressing the multidimensional nature of ethnicity: The multidimensional nature of ethnicity is accepted by most researchers who have attempted to define it. Equally, its capture through measurement on several dimensions, as well as by a single 'global' measure (such as the use of mutually exclusive group labels in an ethnic group classification), has been advocated. This approach appears to have developed from an interest in the de-centering of ethnicity and the development of the concept of "new ethnicities"

(Hall 1992a, 1992b), the constituents of ethnicity, which has arrived in the wake of postmodernism with its rejection of organic unity and the espousal of the fragmentary. The internal composition of these diverse and complex ethnicities is the result of processes such as 'hybridisation' and cultural 'fusion'. Concerns about the reification of 'standard' ethnic categories and the growing alignment of definitions of ethnicity with policy needs have also focused interest on the separate dimensions of the concept. These approaches have challenged the idea that ethnicity can be measured in a unitary way. Both single measures and multidimensional categorisation have a place in studies of the sociology of health and illness and in the wider field of medicine, the choice of method depending on the focus, context, and scale of the investigation.

McKenzie and Crowcroft (1996) and the *BMJ* (Anonymous 1996) advocate the collection for individuals and groups of patients of descriptive data on ethnicity specific to the hypothesis in question or a range of information, minimally comprising *genetic differences, ethnicity (self- and observer-assigned), country or area of birth (the subject's own, or parents' or grandparents' if applicable), years in country of residence, and religion*. Such a recommendation is a sensible and useful approach to many kinds of ethnicity research. This dataset would be of value, for example, when the need is for sociologically rich subgroup information required for the analysis of health risks. As the authors of the guidelines acknowledge, other domains of culture, such as language and diet, may be important to collect in particular studies, and also measures of socio-economic status as a confounding factor (but also meriting detailed investigation in their own right). Similarly, Ahmad and Sheldon (1993) and Bradby (1995) have argued that ethnicity data collection be focused on the particular dimensions that are important for service provision. Indeed, most recent health and lifestyle surveys have collected several domains of ethnicity: unprompted race or ethnic origin, country of birth, parents' or grandparents' country of birth, language(s) spoken at home, and religion in the first BEMG survey; these excluding languages spoken at home but including diet in the second BEMG survey; ethnic origin, main language spoken at home, country of birth, and diet in the Chinese health and lifestyle survey; and ethnic group, family origin, country of birth, and language in the Fourth National Survey.

In adopting a multidimensional measurement approach, however, a number of problems may arise. In many settings including official censuses, government social surveys and public health surveillance, it may not always be practical to collect multidimensional ethnicity data. *Respondent burden* - the time, energy and other demands placed on those to whom the question set is administered (Sharp and Frankel 1983) - and *administrative burden* are important considerations in collecting the set recommended by McKenzie and Crowcroft (1996), especially in treatment and care settings. Given that there is little documented evidence about such burden for questions on 'religion' and other items, or the cumulative effect of a combination of such questions on length of interview and associated fatigue factors, the feasibility of collecting these data needs to be addressed.

While asking about ethnic origin in surveys does not appear to affect the response rates (Windsor 1992), with respect to religion, there are specific concerns which may need to be examined and explored. They include how the classifications relate to the self-understandings of the people concerned, given the empirically supported distinction between such dimensions of religion as believing, belonging, and practising; the way in which religious adherence is often understood to be a very personal matter; some theological objections to the collection of such information; and fears of misuse of the data. However, in Canada, albeit a country with a long tradition of collecting data on religion in its censuses (and in the General Social Survey), the non-response rate in the 1988 National Census Test was only 2.0 per cent, comparable with that observed in the 1991 Census (1.5 per cent), and the invalid response rate just 0.1 per cent. In the 1997 British Census Test the completion rate for the question on religion (91.6 per cent) was similar to ethnic group (92.2 per cent) and country of birth (92.3 per cent) (ONS 1998). Country of birth as the sole indicator of ethnicity has recognised shortcomings (Smaje 1995b) but is useful as a supplementary question to establish generational status. Other items, such as birthplace of parents or grandparents, may be unsuitable for some respondents through lack of knowledge or difficulty of recall. In the 1998 Canadian NCT the non-response rate (6.5 per cent) to the question on birthplace of parents was higher than that for

similar questions - birthplace of respondents (1.6 per cent) and ethnic origin (4.5 per cent) - but invalid responses were minimal and similar for father (0.3 per cent) and mother (0.2 per cent). However, the non-response rate for birthplace of grandparents was about 11.5 per cent, although invalid responses were again small (about 1 per cent for each of the four grandparents). More information is needed on the acceptability of asking about a range of personal identifiers of the type suggested by McKenzie and Crowcroft (1996), by indicating the level of missing and invalid data in surveys, and the reasons for both, before such comprehensive description can be endorsed as a methodological *modus operandi*.

The multidimensional approach to measurement is eschewed by Witzig (1996), who argues that the independent determination of every variable that makes up ethnicity, 'thus rendering the global concept of ethnicity obsolete for the health management of individual persons', is too simplistic. Rather, he sees ethnicity as an important part of self-identity, thus acknowledging a long tradition by critical social theorists of developing the 'group identity' approach to ethnicity. While the decomposition of the concept of ethnicity is appropriate in some contexts and circumvents to some extent the need to develop appropriate group labels, it shifts the focus away from social identification and, consequently, also the exploration of what group identities mean in terms of the beliefs, attitudes, and practices of respondents who use them. McKenzie (1998) has recently argued for a similar cautious approach, contending that it is not always desirable or even possible to 'unpack' culture or ethnicity since they are context driven social and psychological concepts and it is difficult to know how each factor affects the whole.

Newly emerging issues in measuring ethnic aspects of health

(i) *Mutually exclusive categories versus flexibility in self-ascription*: Notwithstanding its evocative title, Bradby's contribution makes no mention of persons of 'mixed' race/ethnicity as one genre of non-standard reporting in the ethnic group question on census forms and survey schedules. In both Great Britain and the USA an increasing number of persons choose not to fit into one of the standard pre-

designated ethnic group categories. In the 1991 Census almost three-quarters of a million people, about 25 per cent of those in ethnic group categories other than White, passed over these categories and utilised the 'free text' fields to give their own description, including 230,000 who declared as being of a 'mixed' origin (Aspinall 1995).

Indeed, despite this important shift, the mixed origin population is frequently ignored in the literature on the collection of ethnicity data. The choice and number of racial/ethnic categories and assignment of persons of mixed or multiracial background have been identified as key quality improvement issues with respect to demographic data (Fiscella, Franks, *et al.* 2000, Dakis and Rubin 1998). The growing number of persons identifying mixed origins renders this position unsatisfactory and requires a change in the way ethnic group categories are constructed. The mixed origin population was estimated at 309,000 persons (1989-91) (OPCS 1992), 11.5 percent of all persons in minority ethnic groups. Responses to the family origins and group membership questions in the Fourth National Survey suggest mixed ethnicity in 1.2 per cent of the White, 8.2 per cent of the Caribbean, 2.5 per cent of the Indian and African Asian, 0.6 per cent of the Pakistani, 0.7 per cent of the Bangladeshi, and 7.5 per cent of the Chinese ethnic family origins category (Nazroo 1997). The historical bias has been toward mutually exclusive single race/ethnic group categorisation, those not wishing to be forced into one of these categories being consigned to 'Other' groups.

This bias is beginning to be broken down in questions on ancestry or ethnic origin. The 1993 and 1996 Canadian Test/Census open response questions on ethnic origin invited respondents to 'specify as many...groups as applicable', 36 per cent of the population giving multiple responses (2 or more ethnic origins reported) in both (28 per cent in the 1986 and 29 per cent in the 1991 Censuses), the proportion being much lower in groups (Asian, Caribbean, African, Arab) who had settled in Canada more recently. In the 1990 US Census open-response ancestry question, 33.5 per cent of the population reported 2 or more ancestries. While the 1980 and 1990 US Census questions on 'race' did not allow for multiple reporting in

spite of the very high levels of intermarriage in some groups for whom a single answer is inappropriate, the 1996 Canadian Census population group question on 'visible minorities' did allow respondents to specify more than one group, 1.9 per cent of the population reporting on this question giving multiple visible minority groups. The United States has now revised Statistical Directive 15 permitting persons of 'mixed race' to tick as many categories as apply. Similarly, the US 2000 Census permits respondents to mark one or more races amongst 12 pre-designated and 3 free-text categories. Clearly, intermarriage is leading to a growing number of persons having multiple ethnic and racial origins, especially amongst members of groups who have been settled for several generations. Questions must allow for hybridized identities representing allegiances to multiple groups, the challenge being the identification of a method that does not imply an outcome from two putatively 'pure' categories.

How such questions are asked depends on the concept of interest. To obtain the 'full size' of the mixed population, that is, all persons of mixed parentage, an operational definition of the type proposed by Berthoud is required. However, research shows that many young persons of mixed parentage prefer to identify as Black or with some other group (Phinney and Alipuria 1996, Fatimilehin 1999, Tizard and Phoenix 1993), so self-perceived ethnic group would provide a different measure of this population. In Britain the option to give those of mixed origins a separate pre-designated ethnic group category has been agreed for the 2001 Census. Although the cultural background options offered in England and Wales are 'mixed race' ('White and Black Caribbean', 'White and Black African', 'White and Asian', and 'Any other mixed background' (free-text)) rather than mixed ethnicity (e.g., 'Italian and other white background'; 'Nigerian and West Indian'), this question format is clearly preferable to multi-ticking of categories which can lead to confusion between identities that *may* signify complex ethnicity (e.g. Indo-Caribbean, Anglo-Indian) and those of mixed parentage origins. Important new analytical work using census microdata, notably the Public Use Samples in the USA (Hwang, Saenz, and Aguirre 1994) and the Samples of Anonymised Records in Great Britain (Holdsworth and Dale 1995, Dale and Holdsworth 1995, Berrington 1996), and other survey datasets (Berrington 1994) is providing significant advances in our understanding

of the patterns of intermarriage and their socio-economic context which, in turn, throws light on the breakdown of separate ethnic group identities and helps to explain the high utilisation of the free-text fields in the 1991 Census to declare mixed origin descriptions.

Other non-standard responses include groups that were concealed in the White category in the 1991 Census. For example, persons wishing to identify as Irish in the final 'Any other ethnic group' free text field had to forego the first precoded category of 'White', a primary identifier for many of them, the estimated 11,000 who did so substantially undercounting the size of this group in Great Britain (OPCS/GRO(S) 1994). Significantly, an estimated further 20,000 people in Greater London ignored the question instruction and ticked the 'White' box on the Census form and also wrote in 'Irish' (these people being coded as White in the full classification). The need for an 'Irish' category has been acknowledged in the 2001 Census in England and Wales, albeit unsatisfactorily juxtaposed with a 'White...British' category, and an Irish category has now been included in the question for Scotland, following a debate in the Scottish Parliament on the question in the Census white paper (Scottish Executive 2000).

Indeed, the need for flexibility in ethnic categorisation is recognised in the 2001 Census question by the inclusion of a free text option in the cultural background options of *each* of the five main ethnic groups. The use of two-tier classifications, embodying the main population groups in the upper and ethnic origin categories in the lower tier (with the option of multi-ticking in the latter) would also give respondents flexibility in self-description, enabling, for example, persons to identify as 'Black' and 'Irish' or 'Black Caribbean' and 'British'. The utility of this option was examined in the 2001 Census development programme but set aside for one which breaks down the main groups. However, a question that cues people for their cultural background using categories located within broad 'race' groups cannot satisfactorily accommodate such backgrounds when they cross 'race' boundaries.

(ii) *Fixed identity categories versus open response*: Recent substantive analyses by Lieberman (1985), Lieberman and Waters (1993), Alba (1990), and Modood, Beishon, *et al.* (1994), pointing to considerable

flux and fluidity in the notion of ethnicity, have shaped conceptualisations. The current challenge is to find ways of measuring ethnic identity in cross-sectional surveys, given that it is labile and fluid. Options include the use of more pre-designated and free-text options in classifications, the use of informants in cognitive settings to explore issues of terminology, and the monitoring of open-ended responses as a way of determining which group labels to employ and when and how to change them (Aspinall 1997). This should also protect researchers against *ad hoc* decision-making in the selection of ethnic group categories, including 'inventing ethnic groups' (Bhopal 1997) and the use of arbitrary classifications.

Postmodernist thinkers have taken the dynamic and ever shifting nature of people's identities as reason to challenge the use of categorisation in itself (Pfeffer 1998), arguing that the protean nature of identities requires an approach that explores the endless processes which shape identity. Indeed, a growing body of sociological literature using innovative methodologies, such as autobiographical and life-writing genres, has demonstrated ways of unravelling personal and social constructions of identity, including resistance to fixed identity categories. The various formulations of 'métissage' (Lionnet 1989, Glissant 1989) describe practices that interweave multiple and composite identities and subject positions drawn from personal experience and are particularly applicable to the experiences of exile and the situation of migrant groups. The significance of métissage lies in its ability to address the fluidity, complexity, and mixture of multiple ethnic, class, and gendered identities. Yet such approaches, by eluding categorisation and by virtue of the use of textual practices to delineate ethnic identities, are difficult to re-integrate into the survey setting. They sidestep to some extent the intractable problem of capturing 'mixed' ethnicity by retreating from any generalisation.

The shortcomings of fixed identity categories have, however, been acknowledged by the Census agencies in the USA and Canada who have taken the courageous step of abandoning them altogether for ethnic origin, using instead open response questions with examples. In much health-related research the impracticalities of a free response - where, for example, the analytical usefulness of the data must be

maintained across administrative records and surveys - means that categorisation into a limited set of choices must take place. However, it is important that, at the same time, longitudinal studies are undertaken to test developmental theories, to follow identity paths, and to show fluidity and change. Frable (1997) calls for new theoretical conceptions that reflect dual or multiple social group memberships, incorporating gender, race, ethnicity, sexuality, and class. Integrating the insights and findings of these approaches into health research is essential if the appropriate vocabulary is to be developed that will enable people's complex multiple social identities to be described.

(iii) Establishing meaning and achieving precision in the measurement of social class, socioeconomic status, and health: While the focus of this paper is the collection of ethnicity data, it would be unsatisfactory to exclude mention of measurement of social class, socioeconomic status, and health. There are two issues that require exploration. Firstly, the predominant model in research is to treat social class and culture or ethnicity as analytically different things. This has resulted in a superficial treatment of social class as a theoretical category of analysis, usually reducing it to socioeconomic group as a descriptive category of analysis and mechanically 'controlling' for it using one or more measures like education, income, and/or occupation. Smaje (1996) argues that such treatment 'can have the unfortunate effect of directing attention away from a key analytical question, namely the nature of the relationship between ethnicity, socio-economic status and health'. Ethnicity or culture is the lived experience of people and is shaped by their location in the class and socioeconomic stratification systems. Thus, research should be aimed at establishing the complex connections between culture, behaviour, and people's class location. This might encompass, for example, consideration of the relationship between migration history and the ownership of economic and 'human capital', the way different social stratification systems determine attitudes to labour market participation, the relationship between class location and people's life chances, and the experience of reception by the host country with respect to racial discrimination, economic exclusion, and differences in opportunities to translate human capital (such as education) into stable employment. Smaje (1996) also makes the important point that, at a conceptual level, it may be difficult to

disentangle the social definition of ethnic 'groups' from socioeconomic 'groups'. He argues that the conventional controlling for socioeconomic factors tends to 'desocialise' and 'de-historicise' the role of material factors and, therefore, 'elide(s) important questions about the disparate ways in which socio-economic status may be produced and reproduced in different ethnic groups'.

Secondly, most studies on ethnic inequalities in health have either ignored socio-economic status altogether or regarded it as a confounding factor, frequently using measures that are imprecise or unsatisfactory. Townsend and Davidson (1982) and Nazroo (1998a) have warned of the inflating of occupational status when use is made of occupation as recorded on death certificates and evidence of misclassification has been reported in the United States (McLaughlin and Mehl 1991). Nazroo (1997, 1998a) has also documented the limitations of traditional class groupings, including their lack of internal homogeneity, and argued that other factors, such as standard of living (a more direct reflection of the material circumstances of respondents), may be better measures of socio-economic status. In the United States the educational achievement of the decedent on the death certificate, introduced to provide a proxy measure of socio-economic status, has been shown to have substantial validity, including across ethnicity-gender groups (Rosamond, Tyroler, *et al.* 1997).

In the United States, Kaufman, Cooper *et al.* (1997) have demonstrated how the validity of studies that control for confounding can be threatened by the presence of different sources of residual confounding, including categorisation of socioeconomic status variables (often leaving considerable variation in risk within groups); measurement error in socioeconomic variables and their surrogates, including that arising from poor operational specification; use of aggregated socioeconomic status measures, for example, values for post/ZIP codes or census areas/tracts; and socioeconomic indicators that are not commensurate between the selected ethnic groups. They show from examples that the effect of such sources is to bias interpretation of data toward the conclusion of independent racial/ethnic effects. Winkleby (1997) also concludes from a review of studies that examine the impact of ethnic and socioeconomic status on cardiovascular disease

that the effects of ethnicity can be overstated and the effects of socioeconomic status underestimated. More work is needed that explores the relationship between health endpoints, ethnic categorisations, and socioeconomic status (including the validity of capture by scalars like education and occupational status) with respect to differences *between* ethnic groups and the nature of the processes involved.

Similar issues are raised with respect to the measurement of health and disease, in particular, the varying meanings that illness has in different cultures and between cultures in the same society. Ethnicity and culture may affect assessments of subjective experience like social handicap and quality-of-life domains such as pain (Kleinman 1986, Bowling 1994), and norms may differ across cultural groups for the restrictions and limitations of illness (Kleinman 1986, Bowling 1994). While many health instruments are adapted for use across cultures through a process of translation, back-translation, review, and modification, they may still not be applicable. This may be particularly true for instruments designed to identify mental illness, evidence suggesting, for example, that they may perform differently for Asians (Nazroo 1998b). Indeed, the 'standard' psychometric measurement used in the validation of instruments for cross-cultural use may mask linguistic and cultural effects (Warnecke, Farrans, *et al.* 1996). Care is also required in devising question wording in surveys of persons of varying racial and ethnic backgrounds, given reported variation related to ethnicity in the way respondents define physical activity, form judgements (including tendency to choose extreme response categories), edit answers, and in willingness to disclose socially undesirable behaviour or overreport socially desirable behaviour (Warnecke, Johnson, *et al.* 1997).

Conclusions and research priorities

The availability of a growing knowledge base on issues of terminology and classifications, deriving from current practice in census and other government agencies and the research traditions of social scientists and demographers, justifies a more rigorous approach to issues of ethnic category selection and the development of classifications. This can only successfully be undertaken through an appreciation of the

conceptual basis of terminology, including the distinction between, and mutual implication of, internal and external definitions, and a stronger development of the theoretical understandings of ethnicity. Work is also needed on how best ethnicity should be conceptualised and measured in the different approaches to explaining ethnic inequalities in health, including epidemiological approaches, structural approaches that focus on material disadvantage, and those that use ethnic identity itself.

Moreover, the challenges presented by the changing character of the population in terms of ethnic group composition, inter-ethnic marriage, and family forms - and the recognition that ethnic identity is dynamic - set a priority on the development of new measurement tools and analytical approaches to investigate ethnic identity through the life course, its intersection with gender, class, and sexuality, the multi-dimensional nature of the concept, and its frequent resistance to easy categorisation. The evidence base suggests there is no true measure of ethnicity that can be applied in a wide variety of contexts and consequently no way that it can be fixed or easily measured. Rather, its contingent, complex and labile nature demands that the means of measurement should be related to the purpose of the research.

Acknowledgements

I am most grateful to the reviewers of this article for providing detailed, informative, and insightful comments on the manuscript throughout the editorial stage.

Notes

1 I am very grateful to Christine Callum for supplying me with customised tabulations from the Health Education Authority's 1992 Black and Minority Ethnic Groups Health and Lifestyle Survey.

Bibliography

Abbotts, J., Williams, R.. and Davey-Smith, G. (1998) Mortality in men of Irish heritage in West Scotland, *Public Health*, 112 (4), 229-32.

Abbotts, J., Williams, R., Ford, G., Hunt, K. and West, P. (1997) Morbidity and Irish Catholic descent in Britain: an ethnic and religious minority 150 years on, *Social Science and Medicine*, 45(1), 3-14.

Ahmad, W.I.U. and Sheldon, T.A. (1993) 'Race' and Statistics. In Hammersley, M. (ed) *Social Research: Philosophy, Politics and Practice*. London: Sage.

Alba, R.D. (1990) *Ethnic Identity: The Transformation of White America*. New Haven: Yale University Press.

Allport, G.W. (1954) *The Nature of Prejudice*. Cambridge and Reading, MA: Addison-Wesley.

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

Ang-Lygate, M. (1997) Charting the spaces of (un)location. On theorizing diaspora. In Mirza, H.S. (ed) *Black British Feminism: A Reader*. Routledge: London.

Anonymous (1996) Style matters. Ethnicity, race, and culture: guidelines for research, audit, and publication, *British Medical Journal*, 312, 1094.

Anwar, M. (1993) *Muslims in Britain: 1991 Census and Other Statistical Sources*. Birmingham: Centre for the Study of Islam and Christian-Muslim Relations.

Aspinall, P.J. (1995) Department of Health's requirement for mandatory collection of data on ethnic group of inpatients, *British Medical Journal*, 311, 1006-9.

Aspinall, P.J. (1997) The conceptual basis of ethnic group terminology and classifications, *Social Science and Medicine*, 45 (5), 689-98.

Australian Bureau of Statistics (1994) *Testing of Ethnic Origin Questions for the 1996 Census. Census Working Paper 94/4*. Brisbane: ABS.

Balarajan, R., Adelstein, A.M., Bulusu, L. and Shukla, V. (1984) Patterns of mortality among migrants to England and Wales from the Indian subcontinent, *BMJ*, 289, 1185-7.

Ballard, R. (1996) Negotiating race and ethnicity: Exploring the implications of the 1991 Census, *Patterns of Prejudice*, 30 (3), 3-33.

Berrington, A. (1994) Marriage and family formation among the white and ethnic minority population in Great Britain, *Ethnic and Racial Studies*, 17 (3), 517-46.

Berrington, A. (1996) Marriage patterns and inter-ethnic unions. In Coleman D., Salt J. (eds) *Ethnicity in the 1991 Census. Vol. 1. Demographic characteristics of the ethnic minority populations*. London: HMSO.

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

Bhopal, K. (1998) South Asian Women in East London: religious experience and diversity, *Journal of Gender Studies*, 7 (2), 143-56.

Bhopal, R. (1997) Is research into ethnicity and health racist, unsound or important science, *BMJ*, 314, 1751-6.

Bhopal, R., Kohli, H. and Rankin, J. (1997) Editors' Practice and Views on Terminology in Ethnicity and Health Research, *Ethnicity and Health*, 2(3), 223-27.

Bourne, J. (1985) *Racism and Jewish Feminism*. London: Institute of Race Relations.

Bowling, A. (1994) Beliefs about illness causation among Turkish and White British people living in a deprived inner London district, *Health Education Research*, 9, 355-64.

Bradby, H. (1995) Ethnicity: not a black and white issue. A research note, *Sociology of Health and Illness*, 17 (3), 405-17.

Braithwaite, R.L., Bianchi, C. and Taylor, S.E. (1994) Ethnographic approach to community organization and health empowerment, *Health Education Quarterly*, 21 (3), 407-16.

Campbell, B.A. (1981) Race-of-interviewer effects among Southern adolescents, *Public Opinion Quarterly*, 45(2), 231-44.

Chapple, J. and Anionwu, E.N. (1998) Genetic services. In Rawaf, S. and Bahl, V. (eds) *Assessing health needs of people from minority ethnic groups*. London: Royal College of Physicians of London & Faculty of Public Health Medicine.

Choi, B.C., Hanley, A.J., Holowaty, E.J. and Dale, D. (1993) Use of surnames to identify individuals of Chinese ancestry, *American Journal of Epidemiology*, 138 (9), 723-34.

Coldman, A.J., Braun, T., and Gallagher, R.P. (1988) The classification of ethnic status using name information, *J Epidemiol Community Health*, 42, 390-5.

Coleman, D. and Salt, J. (eds) (1996) *Ethnicity in the 1991 Census, Vol. 1. Demographic characteristics of the ethnic minority populations*. London: The Stationery Office.

Cornish, J. (1993) National experiences in the measurement of ethnicity: Australia. In *Challenges of measuring an ethnic world: science, politics and reality. Proceedings of the Joint Canada-United States Conference on the Measurement of Ethnicity*. Washington, DC: US Government Printing Office.

Cotter, P.R., Cohen, J., and Coulter, P.B. (1982) Race-of-interviewer effects in telephone interviews, *Public Opinion Quarterly*, 46, 278-84.

Cross, W. (1978) The Thomas and Cross models of psychological nigrescence. A literature review, *Journal of Black Psychology*, 4, 13-31.

Cummins, C., Winter, H., Cheng K.-K., Maric, R., Silcocks, P. and Varghese, C. (1999) An assessment of the Nam Pehchan computer program for the identification of names of south Asian ethnic origin, *Journal of Public Health Medicine*, 21(4), 401-6.

Dakis, P. and Rubin, L. (1998) Obstruction of valid race/ethnicity data acquisition by current data collection instruments, *Meth Inform Med*, 37, 188-91.

Dale, A. and Holdsworth, C. (1995) The construction of ethnicity in the 1991 British Census: evidence from the microdata. In *Proceedings of the Annual Research Conference of the American Bureau of the Census*. US: Department of Commerce.

Der-Karabetian, A. (1980) Relation of two cultural identities of Armenian-Americans, *Psychological Reports*, 47, 123-28.

Douglas, J. (1998) Developing appropriate research methodologies with black and minority ethnic communities. Part 1: reflections on the research process, *Health Education Journal*, 57, 329-38.

Dunn, J. and Fahy, T.A. (1990) Police admissions to a psychiatric hospital: Demographic and clinical differences between ethnic groups, *British Journal of Psychiatry*, 156, 373-78.

Ecob, R. and Williams, R. (1991) Sampling Asian minorities to assess health and welfare, *Journal of Epidemiology and Community Health*, 45, 93-101.

Elam, G. and Chinouya, M. (2000) *Feasibility study for health surveys among Black African populations living in the UK: Stage 2 - Diversity among Black African communities*. London: National Centre for Social Research.

Entwistle, D.R. and Astone, N.M. (1994) Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status, *Child Development*, 65(6), 1521-40.

Farley, R. (1991) The new census question about ancestry: What did it tell us? *Demography*, 28(3), 411-29.

Fatimilehin, I.A. (1999) Of Jewel Heritage: racial socialization and racial identity attitudes amongst adolescents of mixed African-Caribbean/White parentage, *Journal of Adolescence*, 22, 303-18.

Fiscella, K., Franks, P., Gold, M.R. and Clancy, C.M. (2000) Inequality in Quality. Addressing Socioeconomic, Racial, and Ethnic Disparities in Health Care, *JAMA*, 283(19), 2579-84.

Fiske, S.T. and Neuberg, S.L. (1990) A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation, *Advances in Experimental Social Psychology*, 23, 1-74.

Frable, D.E.S. (1997) Gender, racial, ethnic, sexual, and class identities, *Annual Review of Psychology*, 48, 139-62.

Freeth, S., Elliott, D., Lound, C., and Dobbs, J. (1997) *Strategies for sampling the Chinese population in*

England. London: Social Survey Division, Office for National Statistics.

Garcia, J. (1982) Ethnicity and Chicanos: Measurement of ethnic identification, identity, and consciousness, *Hispanic Journal of Behavioural Sciences*, 4, 295-314.

Gatrad, A.R. (1997) Comparison of Asian and English non-attenders at a hospital outpatient department, *Archives of Disease in Childhood*, 77(5), 423-6.

Ghuman, P.A.S. (1997) Assimilation or integration? A study of Asian adolescents, *Educational Research*, 39 (1), 23-35.

Gilroy, P. (1987) *There Ain't No Black in the Union Jack*. London: Hutchinson.

Gilroy, P. (1993) *The Black Atlantic: Modernity and Double Consciousness*. London: Verso.

Glissant, E. (1989) *Caribbean discourse*. Charlottesville: University Press of Virginia.

Goldberg, J. and Kosmin, B.A. (1997) *The social attitudes of unmarried young Jews in contemporary Britain. JPR Report No. 4*. London: Institute of Jewish Policy Research.

Groves, R.J. (1989). *Survey Errors and Survey Costs*. New York: Wiley.

Groves, R.M. and Magilavy, L.J. (1986) Measuring and explaining interview effects in centralized telephone surveys, *Public Opinion Quarterly*, 50, 251-66.

Hall, S. (1992a) The question of cultural identity. In Hall, S., Held, D. and McGrew, T. (eds) *Modernity and its Futures*. Cambridge: Polity Press.

Hall, S. (1992b) New ethnicities. In Donald, J. and Rattansi, A. (eds) *"Race", Culture and Difference*. London: Sage.

Halpern, D. (1993) Minorities and mental health, *Social Science and Medicine*, 36, 597-607.

Halpern, D. and Nazroo, J. (2000) The ethnic diversity effect: Results from a national community survey of England and Wales, *International Journal of Social Psychiatry*, 46(1), 34-46.

Hamilton, D.L. and Trolie, T.K. (1986) Stereotypes and stereotyping: An overview of the cognitive approach. In Hamilton, D.L. (ed) *Cognitive Processes in Stereotyping and Intergroup Behaviour*. Hillsdale, N.J.: Erlbaum.

Harding, S. and Balarajan, R. (1996) Patterns of mortality in second generation Irish living in England and Wales: longitudinal study, *British Medical Journal*, 312, 1389-92.

Harkess, S. and Warren, C.A. (1993) The social relations of intensive interviewing: Constellations of strangeness and science, *Sociological Methods and Research*, 21(3), 317-39.

Harland, J.O., White, M., Bhopal, R.S., Raybould, S., Unwin, N.C. and Alberti, K.G.M.M. (1997) Identifying Chinese populations in the UK for epidemiological research: experience of a name analysis of the FHSA register, *Public Health*, 111, 331-37.

Harrison, F.V. (1994) Racial and gender inequalities in health and health care, *Medical Anthropology Quarterly*, 8(1), 90-5.

Hatchett, S. and Shuman, H. (1975/76) White respondents and race-of-interviewer effects, *Public Opinion Quarterly*, 39, 523-8.

Her Majesty's Government (1999) *The 2001 Census of Population. Cm. 4253*. London: The Stationery Office.

Hoare, T., Thomas, C. and Biggs, A. (1994) Can the uptake of breast screening by 'Asian' women be increased? A randomised controlled trial of a linkworker intervention, *J Public Health Med*, 16, 179-85.

Hodes, M., Creamer, J. and Woolley, J. (1998) Cultural meanings of ethnic categories, *Psychiatric Bulletin*, 22 (1), 20-4.

Holdsworth, C. and Dale, A. (1995) *Ethnic homogeneity and family formation: evidence from the 1991 Household SAR. Occasional Paper 7*. Manchester: Census Microdata Unit.

Huddy, L., Billig, J., Bracciodieta, J., Hoeffler, L., Moynihan, P.J. and Pugliani, P. (1997) The effect of interviewer gender on the survey response, *Political Behaviour*, 19(3), 197-220.

Hutnik, N. (1991) *Ethnic Minority Identity, A Social Psychological Perspective*. Oxford: Clarendon Press.

Hwang, S.S., Saenz, R. and Aguirre, B.E. (1994) Structural and individual determinants of outmarriage among Chinese-Americans, Filipino-Americans, and Japanese-Americans in California, *Sociological Enquiry*, 64(4), 396-414.

Ifekwunigwe, J.O. (1997) Diaspora's daughters, Africa's orphans? On lineage, authenticity and 'mixed race' identity. In Mizra, H.S. (ed) *Black British Feminism: A reader*. Routledge: London and New York.

Isajiw, W.W. (1993) Definitions and Dimensions of Ethnicity: A theoretical framework. In Statistics Canada and Bureau of the Census. *Challenges of Measuring an Ethnic World: Proceedings of the Joint Canada-United States Conference on the Measurement of Ethnicity, 1992*. Washington: US Government Printing Office.

Jacobson, J. (1997) Religion and ethnicity: Dual and alternative sources of identity among young British Pakistanis, *Ethnic and Racial Studies*, 20 (2), 238-56.

Jenkins, R. (1994) Rethinking ethnicity: identity, categorization and power, *Ethnic and Racial Studies*, 17(2), 197-223.

Jenkins, R. (1996) *Social Identity*. London: Routledge.

Jenkins, R.R. and Parron, D. (1995) Guidelines for adolescent health research: issues of race and class, *Journal of Adolescent Health*, 17 (5), 314-22.

Johnson, M.R.D., Owen, D. and Blackburn, C. (2000) *Black and minority ethnic groups in England: the second health and lifestyles survey*. London: Health Education Authority.

Kaufman, J.S., Cooper, R.S. and McGee, D.L. (1997) Socioeconomic status and health in blacks and whites: The problem of residual confounding and the resiliency of race, *Epidemiology*, 8(6), 621-28.

Khoo, S.-E. (1991) Consistency of Ancestry Reporting Between Parents and Children in the 1986 Census, *Journal of the Australian Population Association*, 8, 129-139.

Kish, L. (1962) Studies of interviewer variance for attitudinal variables, *Journal of the American Statistical Association*, 57, 91-115.

Kljakovic, M. (1993) Is it easy collecting ethnicity data in general practice? *New Zealand Medical Journal*, 106, 103-4.

Kleinman, A. (1986) Culture, the quality of life and cancer pain: Anthropological and cross-cultural perspectives. In Ventifrida, V., van Dam, F.S.A.M., Yancik, R. and Tamburini, M. (eds) *Assessment of Quality of Life and Cancer Treatment: Proceedings of the International Workshop on Quality of Life Assessment and Cancer Treatment, Milan, December 1985*. Amsterdam: Excerpta Medica.

Kolev, S.T., House, I., Bell, G., Shaw, D. and Murray, V. (1996) Concentration of blood lead and ethnicity in the United Kingdom, *Occupational and Environmental Medicine*, 53(12), 841-3.

Kumari, K. (1998) The birth of new Asian ethnicities, *Issues in Social Work Education*, 18(2), 25-30.

Lieberson, S. (1985) Unhyphenated Whites in the United States, *Ethnic and Racial Studies*, 8, 159-80.

Lieberson, S. (1993) The Enumeration of Ethnic and Racial Groups in the Census: Some Devilish Principles. In *Challenges of measuring an ethnic world: science, politics and reality. Proceedings of the Joint Canada-United States Conference on the Measurement of Ethnicity*. Washington, DC: US Government Printing Office.

Lieberson, S. and Waters, M.C. (1993) The ethnic responses of Whites - what causes their instability, simplification, and inconsistency, *Social Forces*, 72 (2), 421-50.

Lionnet, F. (1989) *Autobiographical voices: Race, gender, self-portraiture*. Ithaca, NY: Cornell University Press.

Martineau, A. and White, M. (1998) What's not in a name. The accuracy of using names to ascribe religious and geographic origin in a British population, *J Epidemiol Community Health*, 52, 336-7.

McKenzie, K. (1998) Ethnicity in psychiatric epidemiology, *British Journal of Psychiatry*, 172(3), 278.

McKenzie, K. and Crowcroft, N.S. (1996) Describing race, ethnicity, and culture in medical research, *British Medical Journal*, 312, 1054.

McKenzie, K., van Os, J., Fahy, T., *et al.* (1995) Psychosis with good prognosis in Afro-Caribbean people now living in the United Kingdom, *British Medical Journal*, 311, 325-8.

McLaughlin, J. and Mehl, E. (1991) A comparison of occupational data from death certificates and interviews, *Am J Ind Med*, 20, 335-42.

Miller, S., Schmool, M. and Lerman, A. (1996) *Social and political attitudes of British Jews: some key findings of the JPR survey. JPR report No. 1.* London: Institute for Jewish Policy Research.

Modood, T., Beishon, S. and Virdee, S. (1994) *Changing Ethnic Identities.* London: Policy Studies Institute.

Modood, T., Berthoud, R., Lakey, J., Nazroo, J., Smith, P., Virdee, S. and Beishon, S. (1997) *Ethnic minorities in Britain: diversity and disadvantage.* London: Policy Studies Institute.

Nazroo, J.Y. (1997) *The Health of Britain's Ethnic Minorities: Findings from a National Survey.* London: Policy Studies Institute.

Nazroo, J.Y. (1998a) Genetic, cultural or socio-economic vulnerability? Explaining ethnic inequalities in health, *Sociology of Health and Illness*, 20(5), 710-30.

Nazroo, J.Y. (1998b) Rethinking the relationship between ethnicity and mental health: the British Fourth National Survey of Ethnic Minorities, *Soc Psychiatry Psychiatr Epidemiol*, 33, 145-8.

Neeleman, J. and Wessely, S. (1999) Ethnic minority suicide: a small area geographical study in south London, *Psychological Medicine*, 29, 429-36.

NHS Executive (1994) *Ethnic group data collection (DSC Notice: 54/94/P40).* Leeds: NHS Executive.

Nicoll, A., Bassett, K. and Ulijaszek, S.J. (1986) What's in a name? Accuracy of using surnames and forenames in ascribing Asian ethnic identity in English populations, *J Epidemiol Community Health*, 40, 364-8.

Okazaki, S. and Sue, S. (1995) Methodological issues in assessment research with ethnic minorities, *Psychological Assessment*, 7 (3), 367-75.

ONS (SSD) (1996) *1991 Census Validation Survey: Quality Report*. London: The Stationery Office.

ONS (1998) *Evaluation of the Main Objectives of the 1997 Census Test*. London: ONS.

OPCS (1992) *Labour Force Survey 1990 and 1991. Series LFS No. 9*. London: HMSO.

OPCS/GRO(S) (1994) *1991 Census Supplement to Report on Ethnic Group and Country of Birth (Table A)*. London: OPCS.

Pendakur, R. (1993) Visible minority as a redefinition of race. In *International Population Conference/Congres International de la Population: Montreal 1993, Vol. 3*. Liege, Belgium: International Union for the Scientific Study of Population [IUSSP].

Peterson, W. (1987) Politics and the measurement of ethnicity. In Alonso, W. and Starr, P. (eds) *The Politics of Numbers*. New York: Russell Sage Foundation.

Pfeffer, N. (1998) Theories of race, ethnicity and culture, *British Medical Journal*, 317, 1381-4.

Phinney, J. (1989) Stages of ethnic identity in minority group adolescents, *Journal of Early Adolescence*, 9, 34-49.

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

Phinney, J.S. and Alipuria, L.L. (1996) At the interface of cultures: multiethnic/multiracial high school and college students, *Journal of Social Psychology*, 136(2), 139-58.

Phoenix, A. (1994) Practising feminist research: the intersection of gender and 'race' in the research process. In Maynard, M. and Purvis, J. (eds) *Researching women's lives from a feminist perspective*. London: Taylor and Francis.

Pitson, L., Sproston, K., Erens, B. and Whitfield, G. (2000) *Health survey of the Chinese population in England: technical report*. London: SCPR.

Pringle, M. and Rothera, I. (1995) *Ethnic group data collection in primary care: Problems and solutions*. Nottingham: University of Nottingham Medical School.

Pringle, M. and Rothera, I. (1996) Practicality of recording patient ethnicity in general practice: descriptive intervention study and attitude survey, *BMJ*, 312, 1080-2.

Rabkin, J.G. (1979) Ethnic density and psychiatric hospitalization: hazards of minority status, *American Journal of Psychiatry*, 136, 1562-66.

Rainford, L. (1997) *2001 Census testing programme: Report on the ethnic group and religion question test carried out in March 1997*. London: Office for National Statistics.

Rankin, J. and Bhopal, R. (1999) Current census categories are not a good match for identity, *BMJ*, 318, 1696.

Ratcliffe, P. (ed.) (1996) *Ethnicity in the 1991 Census, Vol. 3. Social geography and ethnicity in Great Britain: geographical spread, spatial concentration, and internal migration*. London: The Stationery Office.

Reese, S.D., Danielson, W.A., Shoemaker, P.J., Chang, T.-K. and Hsu, H.-L. (1986) Ethnicity-of-interviewer effects among Mexican-Americans and Anglos, *Public Opinion Quarterly*, 50,563-72.

Reijneveld, S.A. (1998) Reported health, lifestyles, and the use of health care of first generation immigrants in the Netherlands: do socioeconomic factors explain their adverse position? *J Epidemiol Community Health*, 52, 298-304.

Renaud, V. and Costa, R. (1994) *National Census Test: Ethnic Origin. Report No. 16*. Ottawa: Statistics Canada (Housing, Family and Social Statistics Division).

Rhodes, P.J. (1994) Race-of-interviewer effects: a brief comment, *Sociology*, 28(2), 547-58.

Rosamond, W.D., Tyroler, H.A., Chambless, L.E., Folsom, A. R., Cooper, L. and Conwill, D. (1997) Educational achievement recorded on certificates of death compared with self-report, *Epidemiology*, 8(2), 202-4.

Rudat, K. (1994) *Black and Minority Ethnic Groups in England: Health and Lifestyles*. London: Health Education Authority.

Rutter, M., Yule, W., Berger, M., Yule, B., Morton, J. and Bagley, C. (1974) Children of West Indian Immigrants - 1. Rates of behavioural deviance and of psychiatric disorder, *Journal of Child Psychology and Psychiatry*, 15, 241-62.

Schaeffer, N.C. (1980) Evaluating race of interviewer effects in a national survey, *Sociological Methods and Research*, 8, 400-19.

Schmool, M. (1998) The ethnic question on the British census: A Jewish perspective, *Patterns of Prejudice*, 32(2), 65-71.

Shuman, H. and Converse, J.M. (1971) The effects of black and white interviewers on black responses, *Public Opinion Quarterly*, 35, 44-68.

Scottish Executive (2000) *Details of the 2001 Census Amendment Order*. News Release SE 1210/2000.

Senior, P.A. and Bhopal, R. (1994) Ethnicity as a variable in epidemiological research, *British Medical Journal*, 309, 327-30.

Sharp, L.M. and Frankel, J. (1983) Respondent burden - A test of some common assumptions, *Public Opinion Quarterly*, 47 (1), 36-53.

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), 141-63.

Singh, S.P. (1997) Ethnicity in psychiatric epidemiology: need for precision, *British Journal of Psychiatry*, 171, 305-8.

Singh, S.P., Croudance, T., Beck, A., and Harrison, G. (1998) Perceived ethnicity and the risk of compulsory admission, *Soc Psychiatry Psychiatr Epidemiol*, 33, 39-44.

Smaje, C. (1995a) Ethnic residential concentration and health: evidence for a positive effect? *Policy and Politics*, 23, 251-69.

Smaje, C. (1995b) *Health, "Race" and Ethnicity: Making Sense of the Evidence*. London: King's Fund Institute.

Smaje, C. (1996) The ethnic patterning of health: new directions for theory and research, *Sociology of Health and Illness*, 18 (2), 139-71.

Smith, R. (1996) On the disutility of the notion of 'ethnic group'. In Smith, R. (ed) *The Matrifocal Family: Power, Pluralism, and Politics*. London: Routledge.

Sproston, K., Pitson, L., Whitfield, G. and Walker, E. (1999) *Health and lifestyles of the Chinese population in England*. London: Health Education Authority.

Statistics Canada (1996) *1996 Census Guide and reasons why the questions are asked*. Statistics Canada.

Thompson, B. and Tyagi, S. (1996) *Names we call home: Autobiography on racial identity*. New York and London: Routledge.

Tizard, B. and Phoenix, A. (1993) *Black, White, or Mixed Race? Race and Racism in the Lives of Young People of Mixed Parentage*. London and New York: Routledge.

Townsend, P. and Davidson, N. (1982) *Inequalities in Health (the Black Report)*. Middlesex: Penguin.

Tucker, C., McKay, R., Kojetin, B., Harrison, R., de la Puente, M., Stinson, L., and Robinson, E. (1996) *Testing Methods of Collecting Racial and Ethnic Information: Results of the Current Population Survey Supplement on Race and Ethnicity. Statistical Note No. 40*. Washington: Bureau of Labor Statistics.

Ullah, P. (1987) Self-definition and psychological group formation in an ethnic minority, *British Journal of Social Psychology*, 26, 17-23.

Warnecke, R.B., Johnson, T.P., Chavez, N., Sudman, S., *et al.* (1997) Improving question wording in surveys of culturally diverse populations, *Annals of Epidemiology*, 7(5), 334-42.

Warnecke, R..B., Ferrans, C.E., Johnson, T.P., *et al.* (1996) Measuring quality of life in culturally diverse populations, *Monogr. Natl. Cancer Inst.*, 20, 29-39.

Webster, C. (1996) Hispanic and Anglo interviewer and respondent ethnicity and gender: The impact on survey response quality, *Journal of Marketing Research*, 33(1), 62-72.

Weeks, M.F. and Moore, R.P. (1981) Ethnicity-of-interviewer effects on ethnic respondents, *Public Opinion Quarterly*, 45, 245-9.

Welch, S., Comer, J. and Steinman, M. (1973) Interviewing in a Mexican-American Community: An Investigation of Some Potential Sources of Bias, *Public Opinion Quarterly*, 37, 115-26.

Weschler, H. and Pugh, T.F. (1967) Fit of individual and community characteristics and rates of psychiatric hospitalization, *American Journal of Sociology*, 73, 331-8.

Windsor, J. (1992) What can you ask about? The effect on response to a postal screen of asking about two potentially sensitive questions, *Journal of Epidemiology and Community Health*, 46, 83-5.

Winkleby, M.A. (1997) Accelerating cardiovascular risk factor change in ethnic minority and low socioeconomic groups, *Annals of Epidemiology*, 7(7S), S96-S103.

Witzig, R. (1996) The medicalization of race: scientific legitimization of a flawed concept, *Annals of Internal Medicine*, 125 (8), 675-9.

Yuen, E. (1997) Social Movements, Identity Politics and the Genealogy of the Term "People of Color", *New Political Science*, 38-9, 97-107.

Zeuner, D., Ades, A.E., Karnon, J., Brown, J., Dezateux, C. and Anionwu, E.N. (1999) Antenatal and neonatal haemoglobinopathy screening in the UK: review and economic analysis, *Health Technol Assess*, 3(11).