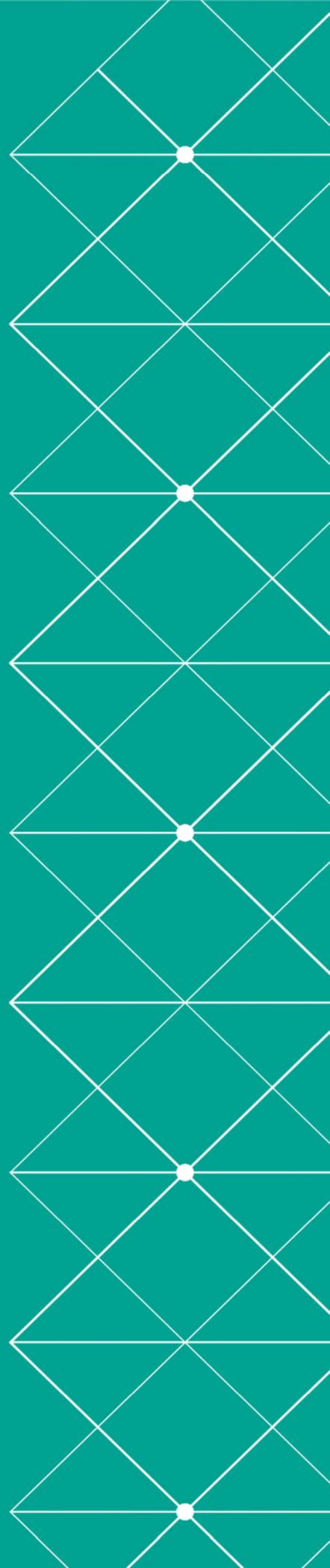# MEASURING CULTURAL AND ETHNIC DIVERSITY

THE EUROPEAN STANDARD CLASSIFICATION OF ETHNIC AND CULTURAL GROUPS

June 2020

Anthony Heath
Centre for Social Investigation

Silke Schneider
GESIS, Mannheim

Abstract

Cultural diversity is an important topic in comparative empirical social science research today. Individuals' socio-cultural and ethnic origins – that are often, but not always, related to their families' migration histories – are a potentially powerful predictor of social attitudes and behaviours. While there are fairly established instruments available for comparatively measuring individuals' migration background or generational status in surveys, this is not the case for individuals' specific socio-cultural and ethnic origins. Therefore, new approaches and instruments for measuring ethnic and cultural origins in (especially cross-national) surveys are needed. While there are numerous statistical classifications available for comparatively coding complex social science concepts such as occupation or education in cross-national data, there is currently no international classification for ethnic and socio-cultural origins. In this working paper, we present such a classification for European purposes, detailing its underlying theoretical concepts, structure and substantive classification criteria. The European Standard Classification of Cultural and Ethnic Groups (ESCEG) and corresponding questionnaire item have been trialled in the European Social Survey (ESS) wave 7 (2014/2015) and have since then been part of the ESS core questionnaire.

Key words: ethnicity, measurement, cultural diversity, migration, survey research

Introduction

As we can witness in the media, public and political debates, cultural diversity is an important topic in most European countries today. The extent to which countries have been successful in integrating migrants and their descendants, and the success or otherwise of policies such as multiculturalism, are high on the political agenda. Differences in attitudes and political aims are also common within nations, often along social, regional, cultural and linguistic lines, and sometimes lead to conflict or demands for autonomy or independence. Monitoring and understanding the extent, or lack, of social and ethnic integration are thus pressing social science issues.

Private bodies and the public sector also increasingly demand data on ethnic and cultural minorities (Abdikeeva 2014; de Voogd 2007). At present a wide range of different methods of identifying and classifying minorities are in use across Europe, with distinct national traditions of measurement. This is a particular problem for comparative research, of which there is a growing volume (see e.g. Crul, Schneider, and Lelie 2012; Heath, Cheung, and Smith 2007; Levels and Dronkers 2008; Morales and Giugni 2011), since the different methods and classifications used cannot be straightforwardly compared. Typically, in survey research, there will be many ethnic groups represented by only a few respondents. Researchers may also be required by the authorities to use broader groups because of disclosure concerns. For comparative research a further major issue is that different ethnic groups predominate in different countries, increasing the need for comparable broader groups in the classification.

Standard statistical classifications are an essential tool for coding cross-national survey data, and data on ethnic and cultural groups is no exception. While there are international standard statistical classifications available for a number of concepts relevant for survey research such as occupation and educational qualifications, some of which are maintained by international

UN-related agencies (see e.g. Hoffmann and Chamie 2002), there is currently no international

standard classification for ethnic and cultural origins.[1] Accordingly, there is no scheme for

aggregating specific ethnic groups into broader groupings comparable to the schemes which exist

for aggregating educational qualifications into education levels. There is not even a standard

geographical scheme for aggregating countries which has gained general assent.[2] This makes it

difficult to compare results from different studies.

There are various reasons for this gap in the classifications landscape. One likely reason is

that any official international classification needs to rely on widely agreed concepts and

definitions. This is very difficult to achieve in the area of ethnic and cultural origins since there is

not even agreement on whether this characteristic should be measured in surveys or official

statistics at all. For example, there has been considerable controversy, notably in France, as to

whether governments should officially collect data on ethnic background (Simon 2012). There

have also been concerns among academics that ethnicity is too complex and fluid a phenomenon

to lend itself to easy measurement (Wimmer 2009). On the former point, we would emphasize

---

[1] To be sure, classifications have been suggested for comparatively measuring individuals' migration
    background (even though there are competing indicators) and generational status in surveys (see e.g.
    Dollmann, Jacob, and Kalter 2014; Maehler et al. 2016). However, this is not the case for individuals'
    specific socio-cultural and ethnic origins, a concept that is not limited to immigrant minorities. Even in
    the case of immigrant minorities, existing measures of migration background are restricted to
    geographical units such as nation states which do not always correspond closely to cultural and ethnic
    origins. For example, migrants from stateless nations such as Kurds and Tamils and diasporic groups
    will be misclassified in most existing schema.

[2] But see the groupings suggested by the UN at http://unstats.un.org/unsd/methods/m49/m49regin.htm.
    There are also several solutions for aggregating countries to the rather simplistic 'global south' (i.e.
    developing or low-income countries) and 'global north' (i.e. developed or high-income countries)
    (International Organization for Migration 2013).

that our proposed classification is designed for research purposes only. This means that no administrative decisions should be influenced or determined by individuals' placement in any of the classification categories.[3] On the second point, we agree with Wimmer that 'one should be careful to avoid the Herderian fallacy of assuming communitarian closure, cultural difference, and shared identity. The study has to ask, rather than assume, whether there is indeed community organization, ethnic closure in networking practices, a shared identity etc.' (Wimmer 2009:265). Another reason for the absence of a standard classification is probably that cultural diversity and immigrant integration, although not new phenomena, have gained public interest in Europe only in the last 10 to 20 years and are thus institutionally not yet well supported, at least in European countries. In contrast, international economic and educational classifications have been established from the 1960s onwards through large bureaucratic organizations such as the International Labour Office (ILO) or UNESCO, who could act as effective custodians of the classifications in their responsibility.

Therefore, new approaches and instruments for measuring ethnic and cultural origins in (especially cross-national) surveys are needed. Specifically, a classification framework for the consistent coding of ethnicity-related data is required. In this paper, we present such a system of classifying cultural and ethnic groups into broader, sociologically-based, groupings for European purposes.

---

[3] Following the Australian Bureau of Statistics, 'the identification of cultural and ethnic groups in the classification, and the way in which they are grouped, does not imply the expression of any opinion on the part of … [the authors] concerning the recognition of any group by governments, organizations or individuals, or the status accorded them. Nor does it imply the expression of an opinion concerning the relative merit or importance of particular cultural and ethnic groups or the people who belong to them.'

Fortunately, the Australian Standard Classification of Cultural and Ethnic Groups 2011 (ASCCEG, Australian Bureau of Statistics 2011) was already available, which the proposed classification builds on. This initial Australian classification was developed through consultations with various stakeholders, such as academics, potential classification users, as well as ethnic and community groups. ASCCEG distinguishes 275 unit groups, and then aggregates these into 28 narrow groups and these are in turn aggregated into nine broad groups. We transferred this classification to match the European context, adjusting both the unit groups identified and the aggregations into broad and narrow groups. In doing this, socio-cultural proximity was given preference over geographical proximity. Also, sub-national divisions within European countries had to be added in order to faithfully code respondents' responses.

The new classification was trialled in the European Social Survey (ESS) round 7 conducted during 2014/2015 (European Social Survey 2018; Heath 2015a). With the ESS central scientific team we developed a question, closely modelled on the measure successfully used by the Australian Bureau of Statistics (ABS), designed to measure ethnic and cultural background. In essence respondents were asked to indicate their ancestry or origins. Country-specific showcards were prepared in consultation with the national coordinators (Heath 2015a, 2017, 2019). Each showcard listed a number of ethnic or national groups, including indigenous minority groups such as Roma, Sami and sub-national cultural groups such as Flemish or Walloons where appropriate and relevant. Our measure, then, is one of cultural and ethnic groups broadly conceived and not a measure of migration background or nationality. Respondents were given the opportunity to write in the names of other groups which were not listed on the showcard, and (as with the Australian model) they were able to indicate two separate groups, for example both Belgian and Flemish. A detailed report evaluating this measure has been published (Heath, Schneider, and Butt 2016). In the current paper, therefore, we do not evaluate the

questionnaire instrument itself but focus on the classification which we have developed for using

the instrument, the European Standard Classification of Cultural and Ethnic Groups (ESCEG, for

the codeframes used in the last ESS rounds see Heath 2015b; Heath, Schneider, and Salini 2017,

2019). After fieldwork in ESS round 7, it was revised in order to improve its accuracy and

applicability. It is this revised version we are presenting here.

The aim of this paper is to describe, and provide the intellectual rationale, for this new

classification of ethnic and cultural groups. The paper starts by delineating the theoretical

concepts underlying the classification and substantive classification criteria and then moves on to

a presentation of its formal structure.

Conceptual basis and classification criteria

Scope of the classification

The ESCEG codifies cultural and ethnic categories (quasi-groups) that are relevant in the

European context. Given the long history of increasingly global migration, its scope is global, i.e.

all cultural and ethnic groups of the world should potentially be classifiable in this framework.

Parallel to the original Australian classification, group size and salience in European countries

was however an important factor in deciding which specific groups to differentiate, and at which

level of the classification. This means that the large numbers of ethnic groups known for many

diverse non-European countries such as China, India or Nigeria are not exhaustively classified

since few of these sub-national ethnic divisions are likely to be salient in the European context. In

the European context, it seems more likely that members of such sub-national groups will

emphasize their national rather than their subnational origins. Our preliminary development work

in Britain, which asked respondents for their own self-descriptions, was in line with this

assumption, as were the write-in responses offered by respondents to the question asked in ESS

round 7. Nevertheless, our classification allows the possibility of adding additional sub-national groups if experience suggests the need for them. (See below on sub-national groups.)

Concepts underlying the classification

The ESCEG framework can be used to classify claims of association[4] related to at least two distinct theoretical concepts: *ethnic identity* and *ethnic origin* (Gans 1979). Ethnic identity refers to the ethnic or cultural group(s) a person most strongly identifies with (subjective belonging or membership). Ethnic origin refers to the ethnic or cultural group(s) a person is thought to descend or originate from (by him/herself or, when looking at perceived ethnic origins, by others). The standard definition of the term "ethnic group" comes from Max Weber. Weber stated that 'we shall call "ethnic groups" those human groups that entertain a subjective belief in their common descent because of similarities of physical type or of customs or both, or because of memories of colonization and migration' (Weber 1978, p. 389). The key component of this definition, as in most sociological approaches to ethnicity, is the central role accorded to subjective identities: whether a particular group of people can be counted as an ethnic or cultural group is a matter for the members of that group to decide, not for outside observers to stipulate on the basis of so-called 'objective' criteria. Ethnicity is essentially self-defined, akin to national identity, party identification, class identity or religious affiliation, whether looking at ethnic identity or ethnic background.

The concepts of ethnic identity and ethnic origin will likely correlate, but they are theoretically independent of one another. In contrast to ethnic identity, origin does not imply

---

[4] Classifying claims of association rather than individuals allows multiple claims of association, so that individuals can have more than one ancestry, or a mixed ethnic identity. This is important to be reflected in survey questionnaires.

identification or close affinity by the individual. It is thus a 'cooler', less fluid and less sensitive

concept, which may be more adequate for measurement in cross-national surveys. While no

attempt is made to trace the historical, let alone genetic origins of individuals (see also Australian

Bureau of Statistics 2011), ethnic origin more closely relates to a family's migration history than

does current ethnic identity. A shared ethnic origin implies a shared cultural heritage, while in

contrast ethnic identity is likely to also reflect current ethnic or cultural integration and

assimilation outcomes. As the American literature has emphasised (Gans 1979), as groups

assimilate across successive generations, they may retain a symbolic ethnicity – a sense of having

a shared origin and heritage, and yet no longer have a strong, bounded sense of ethnic identity.

Different survey questions should be used to assess each of these concepts. Since the ESS used

'ancestry' as the indicator for ethnic origin, it will be used as the example throughout this paper.

The ideas should be transferrable to the concept of ethnic identity though.

Classification criteria[5]

According to the Best Practice Guidelines for developing international statistical classifications

by the United Nations, 'Statistical classifications group and organise information meaningfully

and systematically, usually in exhaustive and structured sets of categories that are defined

according to a set of criteria for similarity' (Hancock 2013). The ESCEG organises ethnic and

cultural groups along socio-cultural rather than just geographic similarity. In many cases, socio-

cultural similarity and geographical propinquity go together because of the diffusion of cultures

to geographically proximate regions, but this is not an invariable rule. So, while geography is an

important indicator of cultural proximity, it may be overruled by sociological criteria. Important

---

[5] Some of our discussion in this section reprises parts of Schneider and Heath (2019).

long-distance trade routes, colonial conquests, and large-scale migration flows (whether forced or in the hope of economic advancement or religious freedom) are three ways in which culture is shaped in ways sometimes unexpected when merely looking at the geographical world map.

However, because ethnic groups can emerge as a result of a variety of different aspects of culture, there is no one single cultural criterion which is a necessary or sufficient condition for the formation of an ethnic group. As Weber emphasised: 'We shall call "ethnic groups" those human groups that entertain a subjective belief in their common descent because of similarities of physical type *or* of customs *or* both, *or* because of memories of colonization and migration' (Weber 1978:389, our italics). Therefore, a range of criteria needs to be considered to achieve a classification decision with respect to any individual ethnic or cultural group. These criteria are presented in the following paragraphs.

Shared history

A long-shared history, which is kept alive (thus constituting a shared heritage), is an important criterion for cultural similarity between ethnic and cultural groups. It often is accompanied by a shared geographic origin (not implying a shared current geographic location). In case there is a shared current geographic location, a cultural and ethnic group typically forms a '*nation*' – a 'cultural artefact of a particular kind' (Anderson 1991:4), an 'imagined political community' (Anderson 1991:6) culminating in a sovereign state. Claims to an own ethnic or national identity and thus new ethnic or national boundaries often go together with the re-writing of historical accounts, or establishing a new ethnic or national history altogether. Nations are thus the core elements of the classification, or could be regarded as the starting point.

However not all ethnic and cultural groups form a nation and not all nations form nation-states (witness stateless nations such as the Kurds, Scots in the UK, or Basques in Spain and

France). Thus our classification is not a classification of nation-states (such exist already, see International Organization for Standardization 2015) *alone*. Instead, it also covers stateless nations as well as non-national cultural and ethnic groups within nation-states or that cross the borders of several nation-states. Many countries show a substantial degree of sub-national ethnic and cultural diversity, which may or may not be related to national aspirations. For example, it is not the same whether a Belgian respondent mentions Flemish or Belgian origins (or identifies as Flemish or Belgian), and it would consequently not be adequate in terms of ethnic and cultural categorization to regard individuals mentioning 'Flemish' as simply 'Belgian'. In addition, some cultural and ethnic groups cross the boundaries of (current) nation states. The concept of 'nation' is thus narrower than the concept of 'cultural and ethnic group', so that nations are one *type* of cultural and ethnic group *amongst others*, and thus only one criterion for identifying cultural and ethnic groups.[6]

Religion

A second core criterion is religion: Religious communities and traditions are a notable element of history that shape family and social customs, much older than the idea of a nation. Some religious communities are (on a global scale) relatively small, socially closed (as indicated by low rates of

---

[6] One could especially criticise the usage of national groups for multi-ethnic countries where national boundaries have largely colonial reasons but do not actually follow language or local ethnic lines. This is certainly true from a local point of view, and we acknowledge that. For a European classification however, this is suitable since migrants from such countries would in Europe still be regarded as rooted in the respective – however artificial – nation state, while the local ethnic group would likely be unknown. The classification also provides codes for some important cross-border groups e.g. in the Sub-Saharan African broad group, and can in principle be extended when needed.

extra-group marriages), culturally homogenous, or otherwise very distinct from others. In this case, religious groups are specifically identified as a cultural and ethnic group in the classification. This is for example true for Jewish, Druze, Parsi and Sikh.[7] For larger religious communities, religious cleavages are sometimes used to distinguish cultural and ethnic groups at higher levels of the classification (e.g. narrow groups 'Mainland and Buddhist South-East Asian' vs. 'Maritime and Muslim South-East Asian'). Many nations are homogenous in religious terms, or strongly dominated by one religious tradition. Others however show internal religious cleavages, which may be an indicator for the relevance of sub-national cultural groups as in the case of Bosnia.

Language

A further criterion is language, which is the most important medium for sustaining and transferring culture (including history, religion and customs) across generations, but also across geographic distances. A shared language (especially sacred language, or proximity with it) and literary tradition is thus another indicator of at least the potential to share some historical roots, religious traditions and customs despite long geographical distance. This is the case for a number of diasporic groups, such as German and Chinese. Different languages in turn are an indicator of different cultural and ethnic groups within a country or nation, as the example of Belgium shows.

---

[7] However, there is also a code 'Muslim nfs', despite the size and heterogeneity of Muslims in the world, which should however only be used when this is the only available information, i.e. when a respondent strongly prefers to name his/her ancestry in religious terms only instead of any national or sub-national group. This pragmatic solution was introduced to accommodate the ESS round 7 show card for Israel. Thus, there are no equivalent codes for 'Christian nfs' or 'Buddhist nfs'.

Social distance

Nationality, religious tradition and language are thus the core indicators for a group's shared history, ethnic origin and culture. Further indicators of similarity of cultural and ethnic groups relate to the social distance between groups, as coined in the sociological concept of panethnicity (Lopez and Espiritu 1990): American and British researchers have found that some ethnic groups have higher rates of friendship, intermarriage, commensality, and collective organization with specific other ethnic groups and do not have random social relations with other minority ethnic groups (Muttarak 2014). Such panethnic groups also tend to have similar political identifications and attitudes for example. A classic example would be Hispanics in the USA, who represent a kind of fusion of many specific Latin American ethnic groups with a common Hispanic heritage.

No classification criterion: Biological ancestry or 'race'

The ESCEG is *not* a classification of racial groups, and usually does not separately identify racial groups within countries. African Americans and Black Canadians are the exceptions, which are justified by the shared history of their ancestors being victims of the slave trade, followed by suppression and discrimination, even long after the abolition of racial segregation. Nevertheless, they clearly belong with the unit groups 'American' and 'Canadian' given their long histories in these two countries and their cultural proximity to other American or Canadian groups.[8]

For immigrant nations such as America, Canada and Australia, one may wonder whether individuals would refer to the origins of their immigrant ancestors, or whether they would

---

[8] Currently no country uses these sub-national categories on their show cards in the ESS, so that they will only be used for post-coding in case a respondent indicated not just 'American' or 'Canadian' but the respective racial group during the survey, i.e. when this is a crucial element in a respondent's self-perception of his/her ancestry.

mention 'American' etc. as their origin, especially if immigration happened many generations ago. The classification and questionnaire items intentionally do not prescribe how far back respondents should go. They will thus themselves select what they know and what is most relevant for them in the given context. This approach is again adopted from the Australian classification: 'the self-perceived group identification approach allows the response 'Australian' and thus allows for the category 'Australian' in the classification. A number of users indicated that the usefulness of the classification would be impaired if it did not allow for the concept of an 'Australian' ethnicity (Australian Bureau of Statistics 2011). One way of thinking of this is that, for some respondents, an Australian heritage will have replaced any pre-migration heritage that their ancestors may have shared (see also Gans 1979). ESCEG is intended to be a classification of ethnic and cultural groups, with a shared cultural heritage, not of biological ancestries.

Structure of the classification

This section explains how the ESCEG is constructed. Statistical classifications consist of 'a set of *discrete, exhaustive and mutually exclusive categories* which can be assigned to one or more variables used in the collection and presentation of data, and which describe the characteristics of a particular population' (Hancock 2013). Given the multiple classification criteria described above, it is difficult to establish classification categories that are entirely mutually exclusive when dealing with ethnic and cultural groups. The classification of some potentially problematic groups is therefore explained in the text below when there would have been obvious alternatives.

The ESCEG has a four-level hierarchical structure, giving it both the depth required for detailed classification and analysis of specific groups, as well as breadth when disregarding highly specific differentiations (see Table 1). Since one (the 'national') level requires more than ten distinctions, the classification uses a 5-digit numerical coding system. As recommended for

hierarchical classifications by UN Statistics, ESCEG runs from the most general and broad

categories at the top level of aggregation (1st digit) to the most detailed level at the bottom level

(5th digit). Each further level thus offers increasingly detailed breakdowns of the broadest groups.

Associated with each code is a short descriptor or label, defining the content of the category, as

well as an explanatory note, which is used to clarify what to include or exclude from a specific

code. An overview of all groups is provided in the appendix. In this version, the classification

consists of 452 detailed groups. In the following, we introduce the building blocks of the

classification. In doing so, we go from more detailed to broad, but presenting the 4th digit level

before the 5th digit level since the latter derives from the former. Only then do we present the 2nd

(narrow group) and 1st (broad group) levels.

Table 1. The ESCEG classification structure: decoding code 11033

digit:	1st	2nd	3rd/4th	5th
group level:	broad	narrow	unit	sub-unit
example:	1	1	03	3
	European	West European	British	Welsh

3rd and 4th digit: national and cross-national minority groups

The third and fourth digit of the classification is made up by categories that most commonly refer

to nation states (e.g. French, Japanese, Kenyan, Chilean, Canadian…), but not exclusively so. As

emphasized above, these core building blocks of the classification also include stateless nations

such as Kurds or Tamils and ethnic minority or regional categories that cross the boundaries of at

least two nation states (e.g. Sami, Roma, Frisian or Silesian). For groups crossing nation-state

boundaries, it is impossible to unambiguously assign them to any *single* national group. Overseas

territories of some nation states (e.g. Bermudan, Aruban) also receive their own category,

because they are specific in that their populations are often ethnically and culturally highly mixed

and contain substantial numbers of indigenous peoples (and, for example in the case of the Caribbean, descendants of slaves) with quite distinct cultural and ethnic roots from the 'national' majority that the territory belongs to. For simplicity, units at this level of the classification will be referred to as 'national' categories anyway.[9]

For understandable reasons, the Australian classification contained many distinct unit groups for different Pacific Islands. At the same time, it lacked some important European distinctions, no doubt because there were relatively few recent migrants to Australia from these backgrounds. A few changes were therefore made to the Australian list of unit groups in the light of what was thought to be most useful for a European context. In doing this, we reviewed published work on ethnic and national divisions within Europe (see e.g. Crul et al. 2012; Heath et al. 2007; Levels and Dronkers 2008; Morales and Giugni 2011), studying which groups were identified, and also consulted experts as well as existing data sources such as the World Factbook (Central Intelligence Agency 2016), which identifies the main ethnic groups within each country. Following this exercise some of the Australian distinctions were dropped, on the assumption that categories such as 'Australasian nec' would suffice. New unit groups were also added, such as Sami (potentially important in Nordic countries) and Quechuan (potentially important in Spain).

Since there are commonly more than nine units within each narrow group (see p.19), the classification requires two digits for the coding of this information. National codes are added to the narrow group the national unit further differentiates. The aim of our final list of unit groups was not to be completely comprehensive but to be pragmatic, including those groups which a reasonable number of respondents in Europe might themselves wish to use. Therefore, there is

---

[9] This is not meant as a sign of support for any potential claims for a sovereign state by those ethnic and cultural groups that cross the boundaries of current states, or overseas territories.

code 99 for cases falling into the respective narrow group but 'not elsewhere classified' (nec) amongst the national groups.

5th digit: Sub-national groups

We made one important addition to the Australian conceptual scheme. As explained in the previous section, the four-digit unit groups mostly refer to national groups. We added a fifth digit for internal differentiation of national groups, e.g. between Flemish and Walloon within the Belgian unit group. Basically, our principle here was that groups which were largely restricted to a single nation state would be identified by a fifth digit within an existing unit group, while groups such as Basques or Frisians which are to be found in two or more neighbouring countries would be assigned a unit group of their own at the 4th digit level (see above). This provides a flexible way of adding new internal distinctions without upsetting the main classification.

The 5th digit level is most important in countries or nations with strong internal ethnic or cultural divisions, which often carry socio-political meaning: Oftentimes, we can expect members of the sub-national categories to be rather unhappy with this state of 'sub-ness' (Anderson 1991:3). The general classification criteria described above (p. 8-12) apply for the decision whether or not and how to make such distinctions. Indicators are, most importantly, the presence of different language groups in one country or strong regional dialects, which often goes along with regional political cleavages resulting in regional political parties or even separatist tendencies. For example, in Spain, Catalan needs to be a possible response option for survey respondents who would refuse to only identify (themselves or their origin) as Spanish, or report Catalan origin as Spanish. While other Spanish regions may be less distinct or politically driven to autonomy, a certain number of them are given equal importance in the classification because they are commonly mentioned by respondents, too. In Belgium, the Flemish and Walloon

16

communities need to be distinguished, and since Brussels does not belong to either of them, it also gets its own category. Major religious cleavages are another source of internal diversity, as in the case of Sikhs in India (see the discussion of classification criteria on p. 10). Often the official national recognition of specific minority groups was helpful in establishing meaningful sub-national (or cross-border) groups.[10]

With sub-national groups that originate from migration but have become established in one or several destination countries, it is difficult to provide clear classification rules. When a group is large and/or has been established for a very long time, as in the case of African Americans and Swedish-speaking Finns, we have decided to give them a sub-national code in the destination country in order to reflect potentially important internal divisions. In contrast, when a group is small and/or has arrived rather recently, as in the case of Ingrians or Estonian Russians, we have decided to give them a sub-national code in the country of origin to better reflect their cultural heritage.

Some national groups have, through clustered (sometimes forced, sometimes voluntary) migration to a number of specific geographic destinations, formed substantial *diaspora communities* over decades or even centuries. These communities have sometimes kept distinct

---

[10] For cases in which we do not know anything but the national category, i.e. 4th digit information, which refers to most entries on show cards referring to national groups, the standard 'national group not further specified' (or nfs) code, ending 0, is used as the default category. The 'nfs' in the label clarifies that, in contrast to the 4th digit level, there might be 'further specified' members of the unit group classified in other 5th digit categories within the same national group. Sub-national categories are then (mostly) ordered alphabetically and counted up from 1 within the national group. For cases in which we do have more information, but of a kind that does not fit into any of the other categories specified at the 5th digit level, a standard 'sub-national group not elsewhere classified' (or nec) code, ending 9, is foreseen. Cross-national minorities, e.g. Basque and Frisian (see section 3.1), do not get further differentiated at the 5th digit level, and just receive a trailing 0 (without 'nfs' in the label).

identities related to their country of origin in their host countries, sometimes even political involvement with their countries of origin, despite the fact that their new area of residence is not bordering with the 'homeland'. Because of, on the one hand, their shared geographic and cultural origin with the national origin group, and on the other hand specific migration history, post-migration cultural change, collective memory of the homeland and sometimes desire to return, we have introduced specific sub-national categories for diasporic groups at the 5th digit level of the classification within the national *origin* group (i.e. German diaspora is a sub-group of German; the Chinese diaspora a sub-group of Chinese; and the Indian diaspora a sub-group of Indian), rather than having many such sub-groups in all countries where they exist.[11] So generally speaking, the cultural origin group is given preference in classification decisions, provided that there is some degree of continuing shared cultural heritage. Note however that many cultural and ethnic groups that are generally not regarded as a diaspora today *originated* from a diaspora, as in the case of English-origin Americans or Australians, or the descendants of Angles and Saxons in the UK. Such groups are recognized to have formed their own culture and eventually nation-states over several centuries and they thus receive their own higher-level codes in the classification.

In some countries, regional or other divisions may exist but not be quite so strong as to be interesting from a substantive point of view. If respondents refer to a specific region (or even city) within those countries anyway, usually by making use of the open response option 'other ancestry', this is made transparent by not assigning them to the standard 'nfs' code, but a designated code. For example, if a respondent writes in 'Bavarian' in the German sample, we

---

[11] However, not all diasporic groups are labelled as such. For example, Jewish are differentiated into Jewish Israeli (for all people in Israel identifying as Jewish or claiming Jewish ancestry) and all 'other' Jewish, which basically refers to the paradigmatic Jewish diaspora.

suggest to use the generic category 8 'German city or region'. Such a code is foreseen for all European national groups, no matter whether there are already specific regional codes or not. (This implies that only up to seven sub-national groups can be distinguished in the classification.)

Since the ESS question allows respondents to give two answers about their ethnic origins, we have not in general added categories for mixed or for hybrid or hyphenated categories.[12] However, we have included a few hyphenated categories such as African American and Swedish-speaking Finnish for long-established groups where there is evidence that they have a distinct cultural heritage in the country of destination. We have also been guided by national practices, as in the case of African American and Hispanic American, so that the classification reflects actual practice rather than imposing a priori distinctions.

2nd digit: narrow groups

At the 2nd classification level, ESCEG specifies 24 narrow groups by summarising 'national' groups regarded as culturally similar. (If no information beyond the 1st digit of the classification (see next section) is available, the standard 'broad group nfs' (not further specified) code, ending 0, is used as the default category, adding another 8 categories to the list of narrow groups.[13])

The list of narrow groups used in the Australian classification needed to be modified in order to make it more useful in a European context. For example, the Australian classification had a separate narrow group for British, another for Irish, but then put all other West European groups into a single narrow group. This would be insufficient – and too Anglo-centric – for use in

---

[12] We should not equate the mention of two ancestries with 'mixed' background, based on intermarriage for example, in the way that concept is used in the official UK classification. Mention of two ancestries could also reflect dual heritages as with Flemish Belgian.

[13] For broad groups that are not commonly used as such, e.g. 'South and South-East Asian', the nfs category will hardly ever be useful but is still included for completeness.

Europe. These two narrow groups were thus moved into the larger group 'West European' as national groups (4th digit).

Sometimes the Australian list of narrow groups appeared to be still based largely on geographical principles with geographically-neighbouring unit groups being placed in the same narrow group, even if they were culturally quite dissimilar. Our principle was to instead aggregate units into culturally-similar narrow and broad groups with some degree of shared heritage. The aim here was to maximise within group cultural homogeneity, which is likely to be most relevant for explanatory purposes, rather than combining disparate ethnic groups which happened to be geographically proximate. For example, 'Mauritian' was included in the 'Southern and Eastern African' narrow group in the ASCCEG, even though most Mauritians are of Indian rather than African heritage, and consequently Hinduism being the major religion in the country and Asian languages being widely used in cultural settings. 'Mauritian' is thus classified as South Asian in ESCEG. Another example is Greenlander, which in ESCEG is classified as belonging to the European broad group due to Danish and Norwegian influence. Also, the ASCCEG narrow group of 'Maritime South East Asian' covers a highly disparate assortment of culturally-different but geographically proximate groups, including Catholic Filipinos, who were ruled by the Spanish for more than 300 years alongside predominantly Buddhist Balinese, Chinese-ancestry Singaporeans, and Muslim Indonesians. In ESCEG, Filipino is thus classified as 'Central American' (which may be surprising) reflecting these historical elements of the heritage, Balinese as 'Mainland and Buddhist South-East Asian', and Singaporean as 'Chinese Asian', leaving a much more homogenous narrow group of 'Maritime and Muslim South-East Asian', including Indonesian. The principle for developing narrow groups therefore was to follow broad geographical principles except where we had evidence, as with the Mauritian and South-East Asian cases, that this was not optimal from a socio-cultural point of view.

Some national groups are difficult to assign to narrow groups. One such case is Russian, which geographically is mostly Asian. Here we followed the Australian example and classified them as (Eastern) European based on criteria of cultural similarity.

There are three narrow groups that are more specific than other narrow groups, actually (mostly) representing national groups that one might rather expect at the 4th digit level: Turkish, Chinese Asian and Jewish. These were regarded as not fitting into any of the other narrow groups, so it was regarded as preferable to have these few national groups at the 2nd digit level already. For Turkish, the third and fourth digits then just add trailing zeros, while for Chinese Asian, there is the national unit group, Chinese, in addition to other Chinese Asian categories (e.g. Singaporean). The 5th digit then follows the standard logic again.

1st digit: broad groups

At the broadest level, ESCEG distinguishes eight cultural and ethnic groups. Together with an administrative code 9 for cases that cannot be classified even into this broad set of ethnic and cultural groups, broad groups are identified through one-digit codes 1-8 (0 is, for technical reasons, not a convenient code at the broadest level of a classification). Adding a one-digit code to the broad group code leads to related narrow group codes (e.g. we find narrow group 11 'West European' nested in the broad group 1 'European').

Even at this level, neither ESCEG nor the original Australian classification use continents, as one would have expected in a geographically based classification. Instead, North African is differentiated from Sub-Saharan African, and East Asian is differentiated from South and South-East Asian, which is again differentiated from Central Asian. This reflects the various cultural criteria described above. ESCEG also differs somewhat from its Australian model at the first digit already. This is partly because, in Europe, other aggregations are desirable than in the Australian

context. Partly however it reflects our preference to more strongly emphasise cultural similarity criteria at this level also. To start with, Oceanian/Australasian and North American were integrated into a new broad largely Anglophone group, while Spanish and Portuguese speaking Latin America was separated from North America. (They formed 'The Americas' in the Australian classification.) As noted above, Latin American has been shown to be a major and distinct panethnic group in the USA.

Next, the different parts of Europe were aggregated into one large European group, also in order to facilitate coding of responses claiming 'European' ethnic origin on its own without further detail. While Europeans do not share one language, they look back on a long common history and high levels of cultural inter-change (often facilitated by European-wide religious traditions such as Catholicism and Protestantism). Finally, 'Central Asian' was moved from the 'South Asian' to the 'North African and Middle Eastern' broad group, and 'South Asian' merged with 'South-East Asians', partly reflecting shared religious traditions.

In terms of how to classify narrow groups within broad groups, one rather difficult case refers to 'Jewish'. Many Jews, especially those not living in Israel, may regard Europe as their closest cultural point of reference because of the long history of Jewish diaspora in Europe. We however decided to follow the Australian classification by classifying them within the 'North African, Middle Eastern and Central Asian' broad group, because this is where Jewish culture originally emerged. Table 2 gives an overview of broad and narrow groups of the classification.

Table 2. Broad and narrow groups, European Standard Classification of Cultural and Ethnic Groups

Broad groups	Narrow groups
1 European	10 European nfs
	11 West European
	12 North European (Nordic)
	13 South European
	14 South-East European
	15 East European
2 North African, Middle Eastern and Central Asian	20 North African, Middle Eastern and Central Asian nfs
	21 Arab
	22 Jewish
	23 Turkish
	24 Iranian and Central Asian
	25 Other North African and Middle Eastern
3 Sub-Saharan African	30 Sub-Saharan African nfs
	31 West and Central African
	32 Africa's Horn
	33 East and South African
4 South and South-East Asian	40 South and South-East Asian nfs
	41 South Asian
	42 Mainland and Buddhist South-East Asian
	43 Maritime and Muslim South-East Asian
5 East Asian	50 East Asian nfs
	51 Chinese Asian
	52 North-East Asian
6 Latin American	60 Latin American nfs
	61 South American
	62 Central American
7 Caribbean	70 Caribbean nfs
	71 English-speaking Caribbean
	72 Non-English-speaking Caribbean
8 North American and Australasian	80 North American and Australasian nfs
	81 North American
	82 Australasian
9 Not classifiable	99 Not classifiable

Summary and outlook

In this paper we have presented the European Standard Classification of Cultural and Ethnic

Groups (ESCEG) to code information on individual's ethnic origins or identities in survey data. It

conceptually and technically builds on the Australian Standard Classification of Cultural and Ethnic Groups (Australian Bureau of Statistics 2011). While in general we followed the Australian classification closely, since it has stood the test of practical usage, we adapted it for European purposes and tried to improve upon the implementation of socio-cultural over geographic criteria. One could debate some of the choices, such as the use of some religious groups like Sikh and Jewish as unit groups and the inclusion of a few hyphenated or hybrid groups such as African American. Their inclusion raises interesting conceptual issues, but unless we had cogent grounds for change, we stayed with the Australian classification.

The aggregation of the building blocks into broad and narrow groups also raises interesting conceptual and empirical issues. Our fundamental criterion for the aggregation is the extent of similarity of the shared cultural heritages, as evidenced for example in patterns of intermarriage. We would thus regard our broad and narrow groups as incipient 'panethnic' groups, drawing on the empirical literature on panethnicity (Lopez and Espiritu 1990; Muttarak 2014). To be sure, as new empirical evidence about inter-ethnic socio-cultural relations and emerging panethnic groupings becomes available, or as cultures and social relationships evolve, our aggregation choices may be found to be defective. Then, adjustments will no doubt need to be made. Providing the 4-digit unit groups remain, different aggregations can be compared both over time and between studies. It is also straightforward to add new fifth digit subnational groups if this becomes appropriate. We have no wish to freeze the classification or to treat ethnic groups as essentialised groups with unchanging characters. As with classifications of occupations, or indeed of nation states, a useful classification has to be a flexible one which can be revised so that it is appropriate to the current character of the phenomenon under scrutiny, while at the same time maximising the possibility of comparisons over time.

Although we hope that the classification can be used in this version for quite some time, usage by the research community will show its usefulness, and point to areas in which the classification needs improving in the future. While it is, for the moment, a classification exclusively used by the European Social Survey, it is hoped that other European data collections may adopt it for classifying ethnic groups across European countries.

As usual with academic classifications, one weakness of this classification is its lack of a custodian, who will make sure it is maintained and updated at certain intervals, and who would support its implementation in different surveys and countries so that the standard is consistently applied. We would therefore suggest to the social scientific community to develop infrastructures allowing the development and maintenance of socially relevant classifications for academic and survey use, rather than limiting itself to (or even blindly adopting) classifications developed in the framework of official statistics.[14]

Acknowledgements

---

[14] In the meantime, when users wish to make changes to the classification for specific purposes, implement it in their own data collection, or even extend it to fit a broader scope than Europe, we would be grateful for them to contact us.

References

Abdikeeva, Alphia. 2014. *Measure, Plan, Act: How Data Collection Can Support Racial Equality*. Brussels: European Network against Racism.

Anderson, Benedict. 1991. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Revised. London: Verso.

Australian Bureau of Statistics. 2011. "Australian Standard Classification of Cultural and Ethnic Groups (ASCCEG)." Retrieved January 15, 2016 (http://www.abs.gov.au/ausstats/abs@.nsf/mf/1249.0).

Central Intelligence Agency. 2016. "The World Factbook." Retrieved February 26, 2016 (https://www.cia.gov/library/publications/the-world-factbook/index.html).

Crul, Maurice, Jens Schneider, and Frans Lelie. 2012. *The European Second Generation Compared : Does the Integration Context Matter?* Amsterdam: Amsterdam University Press.

Dollmann, Jörg, Konstanze Jacob, and Frank Kalter. 2014. *Examining the Diversity of Youth in Europe: A Classification of Generations and Ethnic Origins Using CILS4EU Data (Technical Report)*. 156. Mannheim.

European Social Survey. 2018. *ESS7 - 2014 Documentation Report*. 3.2. Bergen: European Social Survey Data Archive, NSD - Norwegian Centre for Research Data for ESS ERIC.

Gans, Herbert J. 1979. "Symbolic Ethnicity: The Future of Ethnic Groups and Cultures in America." *Ethnic and Racial Studies* 2(1):1–20.

Hancock, Andrew. 2013. *Best Practice Guidelines for Developing International Statistical Classifications. Expert Group Meeting on International Statistical Classifications. New York, 13-15 May 2013*. New York.

Heath, Anthony F. 2015a. "Appendix 11 Ancestry, ESS7-2014 Ed. 1.0." in *ESS-7 2014 Documentation Report.*, edited by European Social Survey. Bergen: European Social Survey Data Archive NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

Heath, Anthony F. 2015b. "European Standard Classification of Cultural and Ethnic Groups (ESCEG) ESS7." Pp. 27–34 in *ESS Round 7 Documentation Report Appendix A6 ed. 2.1-*

Classifications and Coding Standards, ESS7-2014, edited by European Social Survey. Bergen: European Social Survey Data Archive NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

Heath, Anthony F. 2017. "APPENDIX A9: ANCESTRY, ESS8-2016 Ed. 2.0." in *ESS Round 8 Documentation Report*, edited by European Social Survey. Bergen: European Social Survey Data Archive NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

Heath, Anthony F. 2019. "APPENDIX A8: ANCESTRY, ESS9-2018 Ed. 1.0." in *ESS-9 2018 Documentation Report. Edition 1.3*, edited by European Social Survey. Bergen: European Social Survey Data Archive, NSD - Norwegian Centre for Research Data for ESS ERIC.

Heath, Anthony F., Silke L. Schneider, and Sarah Butt. 2016. "Developing a Measure of Socio-Cultural Origins for the European Social Survey." *GESIS Papers* 2016(16).

Heath, Anthony F., Silke L. Schneider, and Luca Salini. 2017. "European Standard Classification of Cultural and Ethnic Groups (ESCEG) ESS8 Ed. 2.1." Pp. 27–37 in *ESS Round 8 Documentation Report Appendix A6 ed. 1.1- Classifications and Coding Standards, ESS8-2016*, edited by European Social Survey. Bergen: European Social Survey Data Archive NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

Heath, Anthony F., Silke L. Schneider, and Luca Salini. 2019. "European Standard Classification of Cultural and Ethnic Groups (ESCEG) ESS9 Ed. 2.0." Pp. 27–37 in *ESS Round 9 Documentation Report Appendix A6 ed. 1.0 - Classifications and Coding Standards, ESS9-2018*, edited by European Social Survey. Bergen: European Social Survey Data Archive NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.

Heath, Anthony Francis, Sin Yi Cheung, and Shawna N. Smith. 2007. *Unequal Chances: Ethnic Minorities in Western Labour Markets*. Vol. 137. Oxford: Oxford University Press.

Hoffmann, Eivind and Mary Chamie. 2002. "Standard Statistical Classifications: Basic Principles." *Statistical Journal of the UN Economic Commission for Europe* 19(4):223–41.

International Organization for Migration. 2013. *World Migration Report 2013. Migrant Well-*

Being and Development.

International Organization for Standardization. 2015. "Country Codes - ISO 3166." Retrieved June 3, 2015 (http://www.iso.org/iso/country_codes).

Levels, Mark and Jaap Dronkers. 2008. "Educational Performance of Native and Immigrant Children from Various Countries of Origin." *Ethnic and Racial Studies* 31(8):1404-1425(22).

Lopez, David and Yen Espiritu. 1990. "Panethnicity in the United States: A Theoretical Framework." *Ethnic and Racial Studies* 13(2):198–224.

Maehler, Débora B., Janna Teltemann, Dominique P. Rauch, and Axinja Hachfeld. 2016. "Die Operationalisierung Des Migrationshintergrunds." Pp. 263–82 in *Methoden der Migrationsforschung: Ein interdisziplinärer Forschungsleitfaden*, edited by D. B. Maehler and U. H. Brinkmann. Wiesbaden: Springer.

Morales, Laura and Marco Giugni, eds. 2011. *Social Capital, Political Participation and Migration in Europe*. Basingstoke: Palgrave Macmillan.

Muttarak, Raya. 2014. "Generation, Ethnic and Religious Diversity in Friendship Choice: Exploring Interethnic Close Ties in Britain." *Ethnic and Racial Studies* 37(1):71–98.

Schneider, Silke L. and Anthony F. Heath. 2019. "Ethnic and Cultural Diversity in Europe: Validating Measures of Ethnic and Cultural Background." *Journal of Ethnic and Migration Studies*.

Simon, Patrick. 2012. "Collecting Ethnic Statistics in Europe: A Review." *Ethnic and Racial Studies* 35(8):1366–91.

de Voogd, Leendert. 2007. "Ethnic and Cultural Surveys - Challanges and Best Practice." *ESOMAR Research World*.

Weber, Max. 1978. "Ethnic Groups." Pp. 385–98 in *Economy and Society: An Outline of Interpretive Sociology*, edited by G. Roth and C. Wittich. Berkeley (CA): University of California Press.

Wimmer, Andreas. 2009. "Herder's Heritage and the Boundary-Making Approach: Studying Ethnicity in Immigrant Societies." *Sociological Theory* 27(3):244–70.

Appendix

Table 3. The European Standard Classification of Cultural and Ethnic Groups, ESS9 v2.0

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
1 European				inc European nfs
	11 West European			
		1101 Austrian	11010 Austrian nfs	
			11018 Austrian city or region nec	
			11019 Austrian nec	
		1102 Belgian	11020 Belgian nfs	
			11021 Flemish	
			11022 Walloon	
			11023 Brussels/Bruxellois	
			11028 Belgian city or region nec	
			11029 Belgian nec	
		1103 British	11030 British nfs	
			11031 English	
			11032 Northern Irish	
			11033 Scottish	
			11034 Ulster	
			11035 Welsh	
			11038 British city or region nec	inc Channel Islander, Cornish, Manx, Gibraltarian
			11039 British nec	
		1104 Dutch	11040 Dutch nfs	
			11048 Dutch city or region nec	excl Frisian
			11049 Dutch nec	
		1105 French	11050 French nfs	
			11051 Breton	
			11052 Corse	
			11058 French city or region nec	inc ch'ti, Normand
			11059 French nec	
		1106 Frisian		
		1107 German	11070 German nfs	inc German-speaking Belgian
			11077 German diaspora/ethnic German	inc Baltic German, Danube Swabian, Volga German, Transylvanian Saxon, often referred to as "Aussiedler" or "Spätaussiedler"; excl Afrikaaner
			11078 German city or region nec	inc Bavarian, Swabian, Saxon, Franks, excl Frisian and Sorb/Wend
			11079 German nec	inc expellees from former German territories such as Sudetendeutsch or Ostpreussisch, excl Silesian

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
		1108 Irish	11080 Irish nfs	
			11081 Irish traveller	
			11088 Irish city or region nec	
			11089 Irish nec	
		1109 Luxembourgish	11090 Luxembourgish nfs	
			11098 Luxembourgish city or region nec	
			11099 Luxembourgish nec	
		1110 Swiss	11100 Swiss nfs	
			11101 Swiss language region nec	inc Swiss-French, Swiss-German, Swiss-Italian, Romansch
			11102 Swiss Canton	
			11108 Swiss city or region nec	
			11109 Swiss nec	excl Yeniche
		1111 Yeniche		
		1199 West European nec		inc Liechtensteiner, Monegasque
	12 North European (Nordic)			
		1201 Danish	12010 Danish nfs	
			12011 Faroese	
			12012 Greenlander	
			12018 Danish city or region nec	
			12019 Danish nec	
		1202 Finnish	12020 Finnish nfs	
			12021 Swedish-speaking Finnish	
			12022 Ingrian	
			12028 Finnish city or region nec	
			12029 Finnish nec	
		1203 Icelandic	12030 Icelandic nfs	
			12038 Icelandic city or region nec	
			12039 Icelandic nec	
		1204 Norwegian	12040 Norwegian nfs	
			12048 Norwegian city or region nec	
			12049 Norwegian nec	
		1205 Sami		
		1206 Swedish	12060 Swedish nfs	
			12068 Swedish city or region nec	
			12069 Swedish nec	
		1299 North European nec		inc Norse, Scandinavian

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
	13 South European			
		1301 Basque		
		1302 Greek Cypriot	13020 Greek Cypriot nfs	
			13028 Greek Cypriot city or region nec	
			13029 Greek Cypriot nec	
		1303 Greek	13030 Greek nfs	
			13038 Greek city or region nec	
			13039 Greek nec	
		1304 Italian	13040 Italian nfs	
			13041 Ladin	
			13042 Sicilian	
			13043 South Tyrolian	
			13048 Italian city or region nec	
			13049 Italian nec	
		1305 Maltese		
		1306 Portuguese	13060 Portuguese nfs	
			13068 Portuguese city or region nec	
			13069 Portuguese nec	
		1307 Spanish	13070 Spanish nfs	
			13071 Andalusian	
			13072 Balearic	
			13073 Canarian	
			13074 Catalan	
			13075 Galician	
			13076 Navarran	
			13077 Valencian	
			13078 Spanish city or region nec	
			13079 Spanish nec	
		1399 South European nec		inc Andorran, Sammarinese
	14 South-East European			
		1401 Albanian	14010 Albanian nfs	
			14018 Albanian city or region nec	
			14019 Albanian nec	
		1402 Bosniak/Bosnian		
		1403 Bulgarian	14030 Bulgarian nfs	
			14038 Bulgarian city or region nec	
			14039 Bulgarian nec	
		1404 Croatian	14040 Croatian nfs	
			14048 Croatian city or region nec	
			14049 Croatian nec	

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
		1405 Kosovar	14050 Kosovar nfs	
			14058 Kosovar city or region nec	
			14059 Kosovar nec	
		1406 Macedonian		
		1407 Moldovan		
		1408 Montenegrin		
		1409 Romanian	14090 Romanian nfs	
			14098 Romanian city or region nec	
			14099 Romanian nec	
		1410 Roma/Gypsy/Sinti		inc Gitanos, Travellers, Kalé and Gens du voyage
		1411 Serbian	14110 Serbian nfs	
			14111 Bosnian Serb	
			14119 Serbian nec	
		1412 Slovene	14120 Slovene nfs	
			14128 Slovenian city or region nec	
			14129 Slovene nec	
		1413 Vlach		
		1414 Ashkali/Balkan Egyptian		
		1415 Gorani		
		1416 Yugoslav		
		1417 Bunjevac		
		1499 South-East European nec		inc "Pomak"
	15 East European			
		1501 Belarusian		
		1502 Czech	15020 Czech nfs	
			15028 Czech city or region nec	inc Bohemian, excl Moravian
			15029 Czech nec	
		1503 Estonian	15030 Estonian nfs	
			15038 Estonian city or region nec	
			15039 Estonian nec	
		1504 Hungarian	15040 Hungarian nfs	
			15048 Hungarian city or region nec	
			15049 Hungarian nec	
		1505 Latvian	15050 Latvian nfs	
			15051 Latgalian	
			15058 Latvian city or region nec	
			15059 Latvian nec	inc Liv
		1506 Lithuanian	15060 Lithuanian nfs	
			15068 Lithuanian city or region nec	
			15069 Lithuanian nec	
		1508 Moravian		

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
		1509 Polish	15090 Polish nfs	
			15091 Lemko	
			15092 Karaim	
			15093 Kashubian	
			15098 Polish city or region nec	
			15099 Polish nec	
		1510 Ruthenian		
		1511 Russian	15110 Russian nfs	
			15111 Estonian Russian	
			15112 Latvian Russian	
			15113 Ukrainian Russian	
			15118 Russian city or region nec	
			15119 Russian nec	
		1512 Silesian		
		1513 Slovak	15130 Slovak nfs	
			15138 Slovak city or region nec	
			15139 Slovak nec	
		1514 Ukrainian	15140 Ukrainian nfs	
			15148 Ukrainian city or region nec	
			15149 Ukrainian nec	
		1515 Sorbian/Wendish		
		1599 East European nec		inc Czechoslovak
2 North African, Middle Eastern and Central Asian				inc North African nfs, Middle Eastern nfs
	21 Arab			
		2101 Algerian		
		2102 Arab Israeli		
		2103 Bahraini		
		2104 Bedouin		
		2105 Egyptian		
		2106 Emirati		
		2107 Iraqi		
		2108 Jordanian		
		2109 Kuwaiti		
		2110 Lebanese		
		2111 Libyan		
		2112 Mauritanian		
		2113 Moroccan		
		2114 Omani		
		2115 Palestinian		
		2116 Qatari		
		2117 Sahrawi/Sahraoui		
		2118 Saudi Arabian		
		2119 Syrian		

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
		2120 Tunisian		
		2121 Yemeni		
		2122 Muslim		
		2199 Arab nec		
	22 Jewish			
		2201 Jewish Israeli	22010 Jewish Israeli nfs	
			22011 Ashkenazi	
			22012 Sephardi	
			22013 Mizrahi	
		2299 Jewish nec	22990 Jewish nec	
	23 Turkish			
		2300 Turkish	23001 Turkish Cypriot	
			23007 Turkish diaspora/ethnic Turk	inc Albanian Turk, Bosnian Turk, Bulgarian Turk, Cretan Turk, Croatian Turk, Dodecanese Turk, Kosovan Turk, Macedonian Turk, Montenegrin Turk, Romanian Turk, Serbian Turk, Slovenian Turk, Western Thrace Turk, excl Turkish Cypriot
			23008 Turkish city or region nec	
			23009 Turkish nec	
	24 Iranian and Central Asian			inc Central Asian nfs, excl Iranian nfs
		2401 Afghan		
		2402 Azeri		
		2403 Hazara		
		2404 Iranian		inc Persian
		2405 Kazakh		
		2406 Kyrgyz		
		2407 Pathan		
		2408 Tajik		
		2409 Tatar		
		2410 Turkmen		
		2411 Uighur		
		2412 Uzbek		
		2413 Bashkir		
		2414 Chuvash		
		2415 Mordva		
		2499 Iranian and Central Asian nec		
	25 Other North African and Middle Eastern			
		2501 Armenian		
		2502 Assyrian/Chaldean		
		2503 Berber		
		2504 Coptic		
		2505 Druze		
		2506 Georgian		

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
		2507 Israeli	25070 Israeli nfs	
			25078 Israeli city or region nec	
			25079 Israeli nec	
		2508 Kurdish		
		2509 Mandaean		
		2510 Avar		
		2511 Chechen		
		2512 Maronite		
		2599 Other North African and Middle Eastern nec		
3 Sub-Saharan African				inc African nfs
	31 West and Central African			inc West African nfs and Central African nfs
		3101 Akan		
		3102 Angolan		
		3103 Beninese		
		3104 Bissau-Guinean		
		3105 Burkinabe		
		3106 Burundian		
		3107 Cabo Verdean		
		3108 Cameroonian		
		3109 Central African		
		3110 Chadian		
		3111 Congolese		
		3112 Ghanaian		
		3113 Guinean		
		3114 Ivorian		
		3115 Liberian		
		3116 Malian		
		3117 Nigerian		
		3118 Rwandan		
		3119 Senegalese		
		3120 Sierra Leonean		
		3121 Yoruban		
		3199 West and Central African nec		
	32 Africa's Horn			
		3201 Afar		
		3202 Amhara		
		3203 Djiboutian		
		3204 Eritrean		
		3205 Ethiopian		
		3206 Fur		inc Darfurian
		3207 Nubian		
		3208 Oromo		
		3209 Somali		
		3210 Sudanese		

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
		3211 Tigre		
		3212 Tigray-Tigrinya		
		3299 Africa's Horn nec		
	33 East and South African			
		3301 Comoran		
		3302 Dinka		
		3303 Kenyan		
		3304 Malawian		
		3305 Masai		
		3306 Motswana		
		3307 Mozambican		
		3308 Nuer		
		3309 Seychellois		
		3310 South African	33100 South African nfs	
			33101 Afrikaaner	
			33102 British South African	
			33109 South African nec	
		3311 Tanzanian		
		3312 Ugandan		
		3313 Zambian		
		3314 Zimbabwean		
		3315 Zulu		
		3316 Malagasy		
		3399 East and South African nec		
4 South and South-East Asian				inc South-East Asian nfs, excl South Asian nfs
	41 South Asian			
		4101 Bangladeshi		
		4102 Bhutanese		
		4103 Indian	41030 Indian nfs	
			41031 Bengali	
			41032 Gujerati	
			41033 Malayali	
			41034 Punjabi	
			41035 Sikh	
			41036 Telegu	
			41037 Indian diaspora	inc Fijian Indian, East Indian Caribbean, Hindustani
			41038 Indian city or region nec	
			41039 Indian nec	
		4104 Kashmiri		inc Azad Kashmiri
		4105 Maldivian		
		4106 Mauritian		
		4107 Nepalese		
		4108 Pakistani	41080 Pakistani nfs	

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
			41081 Sindhi	
			41088 Pakistani city or region nec	
			41089 Pakistani nec	
		4109 Parsi		
		4110 Rohingya		
		4111 Sri Lankan	41110 Sri Lankan nfs	
			41111 Sinhalese	
			41112 Burgher	
			41118 Sri Lankan city or region nec	
			41119 Sri Lankan nec	
		4112 Tamil		
		4199 South Asian nec		
	42 Mainland and Buddhist South-East Asian			
		4201 Balinese		Balinese here rather than in narrow group 43 in view of their Hindu religion
		4202 Burmese		
		4203 Hmong		
		4204 Karen		
		4205 Khmer		
		4206 Lao		
		4207 Mon		
		4208 Thai		
		4209 Vietnamese		
		4299 Mainland and Buddhist South-East Asian nec		
	43 Maritime and Muslim South-East Asian			
		4301 Bruneian		
		4302 Indonesian	43020 Indonesian nfs	
			43021 Acehnese	
			43022 Javanese	
			43023 Madurese	
			43024 Moluccan	
			43025 Sundanese	
			43028 Indonesian city or region nec	
			43029 Indonesian nec	
		4303 Malay	43030 Malay nfs	
			43031 Kadazan	
			43039 Malay nec	
		4304 Timorese		
		4399 Maritime and Muslim South-East Asian nec		

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
5 East Asian				
	51 Chinese Asian			do not use for verbatim answer "Chinese"
		5101 Chinese	51010 Chinese nfs	use for verbatim answer "Chinese"
			51011 Hong Kong Chinese	
			51012 Chinese diaspora	inc Malayan Chinese
			51019 Chinese nec	
		5102 Singaporean Chinese		
		5103 Taiwanese Chinese		
		5199 Chinese Asian nec		
	52 North-East Asian			
		5201 Japanese	52010 Japanese nfs	
			52011 Ainu	
			52012 Burakamin	
			52019 Japanese nec	
		5202 Korean		
		5203 Mongolian		
		5204 Tibetan		
		5299 North-East Asian nec		
6 Latin American				inc Latino, Hispanic
		6001 Amerindian (indigenous peoples of South and Central America nec)		
	61 South American			
		6102 Argentinian		
		6103 Aymara		
		6104 Bolivian		
		6105 Brazilian		
		6106 Chilean		
		6107 Colombian		
		6108 Ecuadorian		
		6109 Paraguayan		
		6110 Peruvian		
		6111 Quechuan		
		6112 Uruguayan		
		6113 Venezuelan		
		6199 South American nec		inc mestizo
	62 Central American			inc Spanish-speaking Caribbean
		6201 Costa Rican		
		6202 Cuban		included here as regarded as culturally Latin American
		6203 Dominican (Dominican Republic)		
		6204 Filipino		don't belong with the East Asians because they are mostly Catholics and their culture is very much influenced by centuries of Spanish influence
		6205 Guatemalan		
		6206 Honduran		

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
		6207 Mayan		
		6208 Mexican		
		6209 Nicaraguan		
		6210 Panamanian		
		6211 Puerto Rican		
		6212 Salvadorean		
		6299 Central American nec		
7 Caribbean				
	71 English-speaking Caribbean			inc West Indian
		7101 Anguillan		
		7102 Antiguan		
		7103 Bahamian		
		7104 Barbadian		
		7105 Barbudan		
		7106 Dominican (Dominica)		island of Dominica, not Dominican Republic
		7107 Guyanese		geographically in South America but English speaking, not Hispanic
		7108 Jamaican		
		7109 Trinidadian		
		7199 English-speaking Caribbean nec		inc Belizean, Virgin Islander, Caribs
	72 French or Dutch-speaking Caribbean			
		7201 Antillais		inc French Guianan and other Caribbean Dom-Tom
		7202 Aruban		
		7203 Curacaoan		
		7204 Haitian		
		7205 Surinamese		
		7299 French or Dutch-speaking Caribbean nec		
8 North American and Australasian				
	81 North American			
		8101 American	81010 American nfs	
			81011 African American	
			81012 Hispanic American	
			81013 Asian American	
			81019 American nec	inc Hawaiian
		8102 Bermudan		
		8103 Canadian	81030 Canadian nfs	
			81031 Quebecois/French Canadian	inc Acadian
			81032 Black Canadian	inc Caribbean Canadian
			81033 Asian Canadian	
			81039 Canadian nec	
		8104 Native North American Indian		inc Inuit, Métis
		8199 North American nec		

Broad groups	Narrow groups	National and cross-border groups	Sub-national groups	Explanatory notes
	82 Australasian			
		8201 Australian	82010 Australian nfs	
			82011 Australian Aboriginal	
			82012 Australian South Sea Islander	
			82013 Torres Strait Islander	
			82019 Australian nec	
		8202 Fijian/iTaukei		
		8203 Maori		
		8204 New Zealander		
		8205 Papua New Guinean		
		8206 Samoan		
		8207 Tahitian		
		8208 Tongan		
		8299 Australasian nec		inc Melanesian, Micronesian and Polynesian
44	444	44444	444444 Not classifiable	
55	555	55555	555555 No further ancestry	
77	777	77777	777777 Refusal	
88	888	88888	888888 Don't know	
99	999	99999	999999 No answer	