

# Chapter 2

## Ethnic Classification in Global Perspective: A Cross-National Survey of the 2000 Census Round

Ann Morning

### 2.1 Introduction

Many if not most countries around the world categorize their inhabitants by race, ethnicity and/or national origins when it comes time to conduct a census. In an unpublished survey of census questionnaires, the United Nations found that 65 % enumerated their populations by national or ethnic group (United Nations Statistics Division 2003). However, this statistic encompasses a wide diversity of approaches to ethnic classification, as evinced by the spectrum of terms employed; ‘race,’ ‘ethnic origin,’ ‘nationality,’ ‘ancestry’ and ‘indigenous,’ ‘tribal’ or ‘aboriginal’ group all serve to draw distinctions within the national population. The picture is further complicated by the ambiguity of the meanings of these terms: what is called ‘race’ in one country might be labelled ‘ethnicity’ in another, while ‘nationality’ means ancestry in some contexts and citizenship in others. Even within the same country, one term can take on several connotations, or several terms may be used interchangeably.

This article surveys the approaches to ethnic enumeration that 141 nations took on their 1995–2004 (or ‘2000 round’) censuses. Using a unique data set compiled by the United Nations Statistical Division, this research identifies several dimensions along which classification practices vary. Specifically, I address three research questions:

1. How widespread is census enumeration by ethnicity, in global terms?
2. Among national censuses that do enumerate by ethnicity, what approaches do they take, in terms of both their question and answer formats?
3. What geographic patterns, if any, do ethnic enumeration practices follow?

---

Abridged from the article published in *Population Research and Policy Review*, 27(2), p. 239–272, 2008.

A. Morning (✉)

Department of Sociology, New York University, New York, NY, USA

e-mail: [ann.morning@nyu.edu](mailto:ann.morning@nyu.edu)

## 2.2 Classification by Ethnicity

This chapter uses a broad definition of ‘ethnic enumeration’ that includes census references to a heterogeneous collection of terms (e.g., ‘ethnic group,’ ‘race,’ ‘people,’ ‘tribe’), which indicate a contemporary yet somewhat inchoate sense of origin-based ‘groupness.’ Despite the fluidity between the conceptual borders of ethnicity, race and nationality, at their cores they share a common connotation of ancestry or ‘community of descent’ (Hollinger 1998). Each concept relies on a different type of proof or manifestation of those shared roots – ethnicity discerns it in cultural practices or beliefs (e.g., dress, language, religion), race in perceived physical traits, and nationality through geographic location – yet they all aim to convey an accounting of origins or ancestry. As a result, in the research to be described I have included all three of these terms – and others – as indicators of one underlying concept of origins. For this umbrella concept I use the label ‘ethnicity’ rather than ‘ancestry,’ however, to emphasize the immediacy that such categories can have when individuals identify themselves. As Alba (1990: 38) points out, ancestry involves beliefs about one’s forebears, while ethnicity is a matter of ‘beliefs directly about oneself.’ He illustrates the difference as being one between the statements, ‘My great-grandparents came from Poland’ (ancestry) versus ‘I am Polish’ (ethnicity).

Identifying a core meaning shared by varied ethnicity-related terms makes possible a global comparative study of ethnic categorization. Previous academic comparisons of census ethnic enumeration have usually included only a few national cases, as part of an intensive examination of the social, historical and political factors behind diverse classificatory regimes (e.g., Kertzer and Arel 2002a; Nobles 2000). And the broader surveys available are generally either regional (e.g., Almey et al. 1992), not based on systematic samples (e.g., Rallu et al. 2004; Statistics Canada and U.S. Census Bureau 1993), or focused on informal conventions rather than official categorization schemes (e.g., Wagley 1965). As a result, no comprehensive international analysis of formal ethnic enumeration approaches precedes this study. One of the fundamental contributions made here is thus an empirical one, in the form of a profile of ethnic enumeration worldwide and typology of such practices.

Providing information about a large sample of contemporary national censuses is also a major step forward for theory-building about the origins of different classificatory systems. Collecting data on the dependent variable of classification type suggests important features to measure and eventually to explain. Rallu et al. (2004) exemplify the possibilities of such an analysis by proposing four types of governmental approach to ethnic enumeration:

1. Enumeration for political control (*compter pour dominer*)
2. Non-enumeration in the name of national integration (*ne pas compter au nom de l'intégration nationale*)
3. Discourse of national hybridity (*compter ou ne pas compter au nom de la mixité*)
4. Enumeration for antidiscrimination (*compter pour justifier l'action positive*)

Rallu et al. (2004) identify colonial census administration with the first category, as well as related examples such as apartheid-era South Africa, the Soviet Union and Rwanda. In these cases, ethnic categories form the basis for exclusionary policies. In the second category, where ethnic categories are rejected in order to promote national unity, western European nations such as France, Germany and Spain are prominent. The third category is largely associated with Latin American countries, where governments take different decisions about whether to enumerate by ethnicity, but a broader discourse praising interethnic mixture or hybridity is not uncommon. The final category is illustrated with examples from Latin America (e.g., Brazil, Colombia) and Asia (China), but the principal cases discussed here are those of England, Canada and the United States, where ethnic census data serve as tools in combating discrimination. Despite the number of regions that Rallu et al. (2004) take into account, however, their conclusions are drawn from a limited set of countries rather than the complete international pool. As a result, the four-part schema they identify might be altered if a wider sample of national censuses were considered.

Another element that is missing from the existing literature on ethnic enumeration is comparative content analysis of the language of census ethnicity items. The studies previously described generally focus on the question of which political motives result in the presence or absence of an ethnic question on a national census. They do not delve into the details of the precise format of the question. But such nuances offer particular applied interest for demographers and other census officials. Maintaining that such technical information is of use for the architects of population censuses, this chapter investigates what terminology is used in different countries (e.g., ‘race’ or ‘nationality’?), how the request for information is framed, and what options are given to respondents in formulating their answer. In this way, the project may suggest alternative approaches to implement when census forms are being redesigned and offer a basis for weighing the relative strengths and weaknesses of diverse formats.

## 2.3 Data and Methodology

As publisher of the annual Demographic Yearbook, the United Nations Statistical Division (UNSD) regularly collects international census information, including both questionnaire forms and data results. For the 2000 round (i.e., censuses conducted from 1995 through 2004), UNSD drew up a list of 231 nations and territories from which to solicit census materials. As of June 2005, this researcher located 141 national questionnaires in the UNSD collection and elsewhere (i.e., from 61 % of the countries listed) and calculated that 30 nations (13 %) had not scheduled a census in that round. Therefore questionnaires were missing from 60 countries (26 % of the original list, or 30 % of the 201 countries expected to have conducted a census within the 2000 round).

The gaps in the resultant database’s coverage of international census-taking were not spread randomly across the globe, as Table 2.1 shows. The nations of Europe were best-represented in the collection, as all of the 2000 census round

**Table 2.1** Countries included in study

	N. America		S. America		Africa		Europe		Asia		Oceania		Total	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Included in study	18	49	11	79	19	34	37	76	37	76	19	76	141	61
Missing questionnaire	17	46	3	21	26	46	0	0	9	18	5	20	60	23
No census planned	2	5	0	0	11	20	12	24	4	8	1	4	30	13
Total	37	100	14	100	56	101	49	100	50	100	25	100	231	100
Region % share in study sample		13		8		13		26		26		13		100
% Region covered		51		79		42		100		80		79		70

Notes:

(1) See Appendix Table A for list of countries comprising each region.

(2) 'No Census Planned' includes both countries that have foregone census enumeration in favour of population registers (this is most often the case in Northern Europe) as well as those that have not scheduled any enumeration for the 2000 round

questionnaires available have been located. Next came Asia (including the Middle East), for which 80 % of the available questionnaires have been obtained, followed by South America and Oceania (79 % each), North America (at 51 %, including Central America and the Caribbean), and Africa (42 %). One effect of this uneven coverage is that African countries, which would make up 22 % of the sample and the second-largest regional bloc after Asia if all its 1995–2004 censuses were included, contribute only 13 % to the final sample of national census questionnaires studied. More generally, the variation in coverage suggests that while the results to be described can be considered a good representation of enumeration in Europe, Asia, South America and Oceania, this is not the case for discussion of North (and Central) America or of Africa. Moreover, the country-level data below do not indicate what percentage of the world's population is covered by the census regimes studied here; findings are not weighted by national population in this inquiry.

Each census form available was checked for questions about respondents' 'race,' 'ethnicity,' 'ancestry,' 'nationality' or 'national origins,' 'indigenous' or 'aboriginal' status – in short, any terminology that indicated group membership based on descent. Although language, religion and legal citizenship questions also appear frequently on national censuses and may be interpreted as reflections of ethnic affiliation, I do not include such indirect references to ancestry. (Consider for example how poor an indicator of ethnicity 'Native English Speaker' status would be in a multicultural society like the United States.) When an ethnicity item as defined above appeared on a census, both the question text and response categories or format were entered verbatim into a database. Translations into English were provided by national census authorities, United Nations staff, the author and others for all but three questionnaires, resulting in a final sample of 138 censuses.

## 2.4 Findings

### 2.4.1 *Frequency of Ethnic Enumeration*

Among the 138 national census questionnaires analyzed, 87 countries or 63 % employed some form of ethnic census classification (see Appendix for complete listing). As Table 2.2 shows, North America, South America and Oceania were the regions with the greatest propensity to include ethnicity on their censuses. While Asia's tendency to enumerate by ethnicity was close to the sample average, both Europe and Africa were much less likely to do so. This regional variation may be explained by Rallu et al's. (2004) hypothesis that concern about the preservation of national unity leads some countries to forgo ethnic enumeration. The tendency toward ethnic counting in the Americas also suggests, however, that societies whose populations are largely descended from relatively recent settlers (voluntary or involuntary) are most likely to characterize their inhabitants in ethnic terms. As Bean and Tienda (1987: 34–35) wrote of the United States, 'an ethnic group is created by the entry of an immigrant group into...society.'

**Table 2.2** Share of countries studied using ethnic enumeration, by region

	N. America		S. America		Africa		Europe		Asia		Oceania		Total	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Enumerating ethnicity	15	83	9	82	8	44	16	44	23	64	16	84	87	63
Total N countries studied in region	18		11		18		36		36		19		138	

## 2.4.2 Census Ethnicity Questions

### 2.4.2.1 Terminology and Geographic Distribution

Not only do nations and regions vary in their censuses' inclusion of ethnicity items, but they also employ widely differing terminology for such questions. In 49 of the 87 cases of ethnic enumeration (56 %), the terms ethnicity or ethnic (or their foreign-language cognates like '*ethnicité*' and '*étnico*') were used. This terminology was found on censuses from every world region. Often the term was combined with others for clarification, as in: 'Caste/Ethnicity' (Nepal); 'cultural and ethnic background' (Channel Islands/Jersey); '*grupo étnico (pueblo)*' (Guatemala); 'Ethnic/Dialect Group' (Singapore); 'Ethnic nationality' (Latvia); and 'race or ethnic group' (Jamaica). Overall, nine different terms or concepts appeared in census ethnicity questions; Table 2.3 lists them in descending order of frequency. The table also distinguishes between 'primary' terms (i.e., first to appear if more than one term is used in one or more questions) and 'secondary,' or following, terms. For example, in the Nepal example above, caste was recorded as the primary term and ethnicity as a secondary term.

As Table 2.3 shows, the second most frequent term after ethnicity was nationality, used by 20 nations (or 23 %). Here nationality denoted origins rather than current legal citizenship status. This distinction was made clear in most cases either by the presence on the census questionnaire of a separate question for citizenship (e.g., Romania, Tajikistan) or by the use of the adjective 'ethnic' to create the term 'ethnic nationality' (Estonia). However, I also include in this category census items that combined ethnicity and nationality by using a single question to identify either citizens' ethnicity or non-citizens' nationality. For example, the Senegalese question ran, '*Ethnie ou nationalité: Inscrivez l'ethnie pour les Sénégalais et la nationalité pour les étrangers*' [Ethnicity or nationality: Write down ethnicity for Senegalese and nationality for foreigners].

References to nationality as ethnic origin came largely from Eastern European nations (e.g., Poland, Romania) and Asian countries of the former Soviet Union such as Tajikistan, Turkmenistan and Uzbekistan (see Table 2.4). This regional concentration reflects a number of historical factors. First, twentieth century (and earlier) movements of both political borders and people in Eastern Europe left groups with allegiances to past or neighbouring governments situated in new or different

**Table 2.3** Terminology of census ethnicity questions

	Number of countries using term as:		Total frequency	
	Primary term	Secondary term	N	%
Ethnicity	45	4	49	56
Nationality	17	3	20	23
Indigenous group/tribe	6	7	13	15
Race	3	10	13	15
Ancestry/descent/origin	3	3	6	7
Cultural group	2	2	4	5
Community/population	3	0	3	3
Caste	2	0	2	2
Colour/phenotype	2	0	2	2

*Notes:*

(1) The number of primary terms does not sum to the full number of countries that enumerated by ethnicity (87) because some censuses either included an ethnicity term in a secondary position only, preceded by terms referring to language or religion, or used no descriptive term at all (e.g., Philippines: 'How does [the person] classify himself/herself?').

(2) The sum of term frequencies exceeds 100 % because some censuses feature more than one term

states (Eberhardt 2003). Second, this reinforced existing Romantic notions of nations as corresponding to ethnic communities of descent (Kertzer and Arel 2002b). Finally, the Soviet Union's practice of identifying distinct nationalities within its borders extended the equation of nationality with ethnic membership (Blum and Gousseff 1996).

Roughly 15 % of the national censuses asked about respondents' indigenous status. These cases came from North America (e.g., Mexico: '*¿[Name] pertenece a algún grupo indígena?*'; [Does [name] belong to an indigenous group?], South America (e.g., Venezuela: '*¿Pertenece usted a algún grupo indígena?*'; [Do you belong to an indigenous group?], Oceania (e.g., Nauru: 'family's local tribe'), and Africa (Kenya: 'Write tribe code for Kenyan Africans'). Indigeneity seems to serve as a marker largely in nations that experienced European colonialism, where it distinguishes populations that ostensibly do not have European ancestry (separating them from mestizos, for example, in Mexico) or who inhabited the territory prior to European settlement. The indigenous status formulation was not found on any European or Asian censuses.

The same number of countries (13, or 15 % of all censuses using some form of ethnic enumeration) asked for respondents' race, but this term was three times more likely to appear as a secondary term than as a primary one. For example, the Brazilian question placed race after colour ('*A sua cor o raça e:*'), and Anguilla used race to modify ethnicity: 'To what ethnic/racial group does [the person] belong?'. Race usage was largely confined to North America (including Central America and the Caribbean), as well as to United States territories in Oceania (American Samoa, Federated States of Micronesia, Guam, Northern Mariana Islands). More specifically, census usage of race is found almost entirely in the former slaveholding societies of the Western Hemisphere and their territories.

**Table 2.4** Census ethnicity terminology by region

Primary or secondary term:	N. America		S. America		Africa		Europe		Asia		Oceania		Total	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Ethnicity	8	53	3	33	4	50	9	56	12	52	13	81	49	56
Nationality	0	0	0	0	2	25	9	56	8	35	1	6	20	23
Indigenous/tribe	2	13	6	67	1	13	0	0	0	0	4	25	13	15
Race	7	47	1	11	1	13	0	0	0	0	4	25	13	15
Countries covered by 4 terms	13	87	8	89	5	63	16	100	20	87	16	100	78	90
No. Countries using some ethnicity term	15		9		8		16		23		16		87	

*Note:*

Percentages do not total to 100, because (a) not all ethnic terms are included; and (b) many countries use more than one ethnic term on their censuses



Of the 13 countries studied that enumerate by race, 11 are either New World former slave societies (United States, Anguilla, Bermuda, Brazil, Jamaica and Saint Lucia) and/or their territories (United States Virgin Islands, Puerto Rico, American Samoa, Guam, and Northern Mariana Islands).

Table 2.4 summarizes the geographic patterns in usage of the four most frequent ethnic terms found on national census questionnaires. Reference to ethnicity is most prevalent in Oceania and least prevalent in South America, whereas nationality is found on more than half of the European censuses but on none in the Americas. Conversely, references to indigenous status or ‘tribe’ reach their peak in South America, but are absent on European and Asian censuses. Similarly, race is not found on European or Asian censuses, but appears on almost half of those used in North America (which includes Central America and the Caribbean). Still, in all regions ethnicity remains the most frequent term used, with the exception of South America, where references to indigenous status appear twice as often as those to ethnicity. Together, the four most frequent terms – ethnicity, nationality, indigenous group and race – appear on 90 % of the censuses that enumerate by ethnicity.

#### 2.4.2.2 The Language of Census Ethnicity Questions: The Subjectivity of Identity

Census ethnicity questions vary considerably not just in their terminology but also in the language they use to elicit respondents’ identities. In particular, census questionnaires differ noticeably in their recognition of ethnicity as a matter of subjective belief, as opposed to objective fact. Twelve (or 14 %) of the 87 countries that practice ethnic enumeration treat it as a subjective facet of identity by asking respondents what they ‘think,’ ‘consider,’ or otherwise believe themselves to be. Examples come from every world region. Saint Lucia’s census asks, ‘To what ethnic group *do you think* [the person] belongs?’ (emphasis added) rather than simply, ‘To what ethnic, racial or national group *does* [the person] belong?’ The same explicitly subjective formulation is found on the census questionnaires of New Caledonia (*‘A laquelle des communautés suivantes estimez-vous appartenir?’* [To which of the following communities *do you think* you belong?]), and Paraguay (*‘¿Se considera perteneciente a una étnia indígena?’*; [*Do you consider yourself* as belonging to an indigenous ethnic group?]), for example (emphases mine).

In addition to the recognition of the subjectivity of identity through references to respondents’ beliefs, these censuses achieve the same end by emphasizing the personal, self-selected aspect of ethnicity; it is what the individual says it is, not the product of an objective external measurement. Accordingly, the individual respondent’s choice is paramount here, as in the Philippines’ question, ‘How does [the person] classify himself/herself?’ or Bermuda’s ‘In your opinion, which of the following best describes your ancestry?’ South Africa’s census asks, ‘How would (the person) describe him/herself in terms of population group?’ while Jamaica asks, ‘To which race or ethnic group would you say you/... belong(s)?’, both questions employing the conditional tense. Deference to the individual’s choice of

self-recognition is found in non-English formulations as well, such as Argentina's '*¿Existe en este hogar alguna persona que se reconozca descendiente o perteneciente a un pueblo indígena?*' [Is there someone in this household who considers him/herself a descendant of or belonging to an indigenous people?], or Suriname's '*Tot welke etnische groep rekent deze persoon zichzelf?*' (With which ethnic group does this person identify him/herself?). Peru's census question even lays out the basis on which individuals might construct their ethnic identity, asking '*¿Por sus antepasados y de acuerdo a sus costumbres Ud. se considera:...*' [Given your ancestors and traditions, you consider yourself...].

Many of these examples also illustrate another strategy of recognizing the subjectivity of identity, and that is the reference to ethnic groups as something with which one is affiliated, as opposed to the more total ethnicity as something that one is. The difference between an essential being ethnic and a constructed belonging to an ethnicity can be illustrated by juxtaposing the question 'What is your ethnic group?' (United Kingdom) against 'To what ethnic group do you belong?' (Guyana). The difference is subtle, yet it marks a distinction between a more essentialist concept of ethnicity as objectively given, and a more constructionist understanding of ethnicity as socially and thus subjectively developed. In addition to the 14 % of the national censuses studied that presented ethnicity as subjective in the ways previously described, another 21 % (18 countries) used the concept of belonging (*appartenir* in French, *pertenecer* in Spanish) in the formulation of their ethnicity question. Again, this approach was found on censuses from every world region.

It is clear however that in the majority of cases, census ethnicity questions were brief and direct, simply treating ethnicity as an objective individual characteristic to be reported. Some did not in fact include a question, merely a title (e.g., 'Ethnic Group,' Bulgaria). However, it should be noted that three national censuses from Eastern Europe indicated that it was not obligatory to respond to the ethnicity question, ostensibly due to its sensitive nature. Croatia's census notes 'person is not obliged to commit himself/herself,' Slovenia's reads, 'You don't have to answer this question if you don't wish to,' and Hungary adds, 'Answering the following questions is not compulsory!'

### 2.4.3 Answering the Ethnicity Question

#### 2.4.3.1 Response Formats

Turning now to the structuring of response options for ethnicity questions, the national censuses studied employed three types of answer format:

1. Closed-ended responses (e.g., category checkboxes; code lists)
2. Closed-ended with open-ended 'Other' option (i.e., permitting the respondent to write in a group name not included on the list presented)
3. Open-ended (i.e., write-in blanks)

The three approaches were used in nearly equal proportions among the 87 countries employing ethnic enumeration: 32 (37 %) used the entirely closed-ended approach, 28 (32 %) the mixed approach, and 27 (31 %) permitted respondents to write in whatever ethnic identity they chose.

The closed-ended approach generally took two forms: either a limited number of checkbox category options, or the request to select a code from a list of ethnic groups assigned to codes. The former strategy can be found, for example, on the Brazilian census, which gave respondents five options to choose from to identify their 'colour or race': (1) *Branca* (white); (2) *Preta* (black or dark brown); (3) *Parda* (brown or light brown); (4) *Amarela* (yellow); (5) *Indigena* (indigenous). This listing of five categories is a relatively brief one; another such example is Romania's series of 'nationality' answers: (1) Romanian; (2) Hungarian; (3) Gypsy/Roma; (4) German and (5) Other. At the other end of the spectrum, Guatemala offered a list of 22 indigenous groups plus Garifuna and Ladino, and Argentina and Paraguay each presented a list of 17 indigenous groups for selection by the respondent. However, the second type of closed-ended format – the linking of ethnic groups to code numbers – permitted respondents to select from an even longer list of choices; Laos offered 48 such code options. Other countries to use the code-list strategy were Ghana, Kenya, Malaysia, the Philippines and India.

An even wider range of responses was possible on the censuses that featured the combination of closed-ended categories with a fill-in blank for the 'Other' option alone. After giving respondents six options to choose from – Estonian, Ukrainian, Finnish, Russian, Belorussian and Latvian – the Estonian census requested that individuals choosing the seventh 'Other' box write in their specific 'ethnic nationality.' In Mongolia, respondents either identified with the Khalkh option or wrote in their ethnicity. Singapore listed 13 possibilities for 'ethnic/dialect group' – Hokkien, Teochew, Cantonese, Hakka (Khek), Hainanese, Malay, Boyanese, Javanese, Tamil, Filipino, Thai, Japanese and Eurasian – before requesting specification from anyone selecting the last, 'Others' option.

In the last, entirely open-ended strategy, respondents were simply asked to 'write in' (Senegal) or 'provide the name of' (China) their ethnic group. This approach may not always offer the respondent as much latitude as it appears, however. In nations where one's ethnic affiliation is firmly fixed in other official records (e.g., mandatory identity documents), individuals may not choose freely from an unlimited range of identities so much as they reproduce the label that has already been assigned to them by state bureaucracies.

Although the sample of censuses studied was fairly evenly divided across the three types of ethnic response format, each world region generally favoured one approach more than the others. Table 2.5 shows that in South America and Africa, the closed-ended approach was taken by about two thirds of the national censuses, whereas roughly the same share in Europe used the mixed approach, and about two thirds of Asian censuses relied on the open-ended strategy.

In addition to geographic distribution, census ethnicity response formats also vary depending on whether the terminology in use is ethnicity, nationality, indigenous status/tribe, or race (see Table 2.6). In particular, questions on nationality are

**Table 2.5** Census ethnicity response formats by region

	N. America		S. America		Africa		Europe		Asia		Oceania		Total	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Primary term only:														
Closed-ended	7	47	6	67	5	63	2	13	6	26	6	38	32	37
Closed w/ 'other' write-in option	6	40	3	33	1	13	11	69	2	9	5	31	28	32
Open-ended	2	13	0	0	2	25	3	19	15	65	5	31	27	31
Total	15	100	9	100	8	101	16	101	23	100	16	100	87	100

*Note:*

Each country is represented only once here even if its census includes more than one question on ethnicity. In that case, only the first question is classified here

**Table 2.6** Census ethnicity response formats by question type

	Ethnicity		Nationality		Indigenous/ Tribe		Race	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Primary term only:								
Closed-ended	17	38	2	12	4	67	2	67
Closed w/ 'other' write-in option	14	31	6	35	1	16	1	33
Open-ended	14	31	9	53	1	16	0	0
Total	45	100	17	100	6	99	3	100

*Note:*

Only 71 countries, rather than the full 87 that enumerate by ethnicity, are included in this table because it is limited to census questionnaires whose primary ethnicity term is one of the four most frequent terms: ethnicity, nationality, indigenous/tribe, or race. See Table 2.3 for the breakdown of ethnicity terms by primary and secondary status

most likely to permit some kind of write-in response, while those inquiring about indigenous status and race are the least likely to do so. The first finding may reflect the expectation that fairly few national origins are likely to be elicited and thus an open-ended approach is not likely to become unwieldy. The second finding may reflect governmental tendencies to develop official lists of indigenous and racial groups that are formally recognized by the state, coupled with a sense of necessity to assign all respondents to such predetermined indigenous or racial groups. In addition, popular conceptions of these identities may depict them as involving a limited number of categories (such as 'black,' 'white,' and 'yellow' colour groupings) or even simple dichotomies (e.g., indigenous versus non-indigenous).

### 2.4.3.2 Response Options

Census response formats for ethnicity vary in other ways worth noting:

(a.) Mixed or Combined Categories. Several census questionnaires permit the respondent to identify with more than one ethnicity. This flexibility takes three forms. First, some censuses allow the respondent to check off more than one category (e.g., Channel Islands – Jersey; Canada; New Zealand; United States; U.S. Virgin Islands). Other census questionnaires offer a generic mixed-ethnicity response option (e.g., 'Mixed': Channel Islands – Jersey, Saint Lucia, Anguilla, Guyana, Zimbabwe, Trinidad and Tobago, Jamaica, Mozambique, Solomon Islands, Suriname; 'Mestizo': Belize, Peru; 'Coloured' in South Africa). Finally, some censuses specify exact combinations of interest, for example: 'White and Black Caribbean,' 'White and Black African,' etc. in the United Kingdom; 'Black and White,' 'Black and Other,' etc. in Bermuda; 'Part Cook Island Maori,' Cook Islands; 'Eurasian,' Singapore; 'Part Ni-Vanuatu,' Vanuatu; 'Part Tokelauan/Samoan,' 'Part Tokelauan/Tuvaluan,' etc., Tokelau; 'Part Tongan,' Tonga; and 'Part Tuvalu' in Tuvalu.

(b.) Overlap between ethnic, national, language and other response categories. The conceptual proximity between such concepts as ethnicity and nationality is

illustrated once again by some censuses' use of the same set of response categories to serve as answers to distinct questions on ethnicity, nationality, or language. For example, the Bermudan census response category 'Asian' can be selected when responding either to the race or the 'ancestry' question. An even more striking example comes from Hungary, where the same detailed list of categories serves as the response options to three separate questions (one each for nationality, culture and language). The options are: Bulgarian; Gipsy (Roma); Beas; Romani; Greek; Croatian; Polish; German; Armenian; Roumanian; Ruthenian; Serbian; Slovakian; Slovenian; Ukrainian; Hungarian, and 'Do not wish to answer.' Moldova also uses the same responses for three questions (one each on citizenship, nationality and language), while Estonia and Poland use the same categories for their citizenship and ethnic nationality questions, and Latvia, Romania, and Turkmenistan use the same response options for nationality and language questions.

It is also worth recalling that even when only one ethnicity question appears on a census with one set of response options, the answer categories themselves may reference multiple concepts such as race and nationality. The United States' race question, which includes answers like 'white' and 'black' alongside national or ethnic designations like 'Korean' and 'Japanese,' provides a good example. Similarly, Saint Lucia and Guyana's ethnicity options include races like 'black' and 'white' alongside national designations like 'Chinese' and 'Portuguese.'

Nationality and ethnicity are also intertwined on censuses that use a single question to ask respondents for ethnicity if they are citizens, but for something else if they are foreigners. For example, Indonesia requests, 'If the respondent is a foreigner, please specify his/her citizenship and if the respondent is an Indonesian, please specify his/her ethnicity.' Kenya's ethnicity question reads, 'Write tribe code for Kenyan Africans and country of origin for other Kenyans and non-Kenyans.' Zambia's ethnicity question instructs, 'If Zambian enter ethnic grouping, if not mark major racial group.' And Iraq's census asks only Iraqis to answer the ethnicity question.

Perhaps the simplest cases of conceptual overlap occur, however, on censuses that combine multiple terms in the same item, such as the conflation of ethnicity and race in the Solomon Islands' question: 'Ethnicity. What race do you belong to? Melanesian, Polynesian, Micronesian, Chinese, European, other or mixed?'

(c.) Use of examples. National censuses vary considerably in the extent to which they employ examples to facilitate response to their ethnicity questions. Given typical space constraints, this strategy is not widespread; instead, the list of checkbox response options may serve as the principal illustration of the objective of the question. For example, the Philippine presentation of examples before its closed-ended code-list question is unusual: 'How does [the person] classify himself/herself? Is he/she an Ibaloi, Kankanaey, Mangyan, Manobo, Chinese, Ilocano or what?' Instead, examples are more likely to be employed when the answer format calls for an open-ended write-in response; it is in this context, for example, that Fiji offers respondents the examples 'Chinese, European, Fijian, Indian, part European, Rotuman, Tongan, etc.' The U.S. Pacific territories do the same for their 'ethnic origin or race' write-in item.

In summary, both the amount of latitude that census respondents enjoy when answering an ethnicity question and the amount of guidance or clarification they are given vary widely across the international spectrum.

## 2.5 Conclusions

### 2.5.1 *Summary of Findings*

Although widespread, ethnic enumeration is not a universal feature of national censuses; 63 % of the censuses studied here included some type of ethnicity question. In nearly half of these cases, 'ethnicity' was the term used, but significant numbers of censuses inquired about 'nationality,' 'indigenous status,' and 'race.' Each of these terms tended to be associated with a particular type of response format: questions about indigenous status were most likely to entail a closed-ended response format (checkboxes or code lists), whereas nationality questions were the most likely to permit open-ended responses (i.e., fill-in blanks). National census practices also varied in terms of their allowance of multiple-group reporting and use of examples.

The large number of questionnaires studied here (138 in total, with 87 employing ethnic enumeration) permits the exploration of geographic patterns in census practices. Based on this sample, it appears that nations in the Americas and in Oceania are most likely to enumerate by ethnicity, while those in Europe and Africa are the least likely. Among the countries that do practice census ethnic classification, the term 'nationality' is most likely to be used in Eastern Europe and the former Soviet Union, while 'indigenous status' is most likely to be a concern in the Americas, as is 'race.'

### 2.5.2 *Evaluating Ethnic Enumeration*

In addition to the empirical, theoretical and applied contributions to be made to existing research on ethnic classification (see Morning 2008), the findings reported here are relevant to debates about the formulation, feasibility and desirability of both census ethnic enumeration and international guidelines concerning it. Any proposal for new enumeration strategies, however, must reckon with the fact that census construction is not merely an exercise in survey design; it is fundamentally a political process, where state and group interests and ideology thoroughly inform the final census product (Anderson 1988; Kertzer and Arel 2002a; Nobles 2000; Skerry 2000). The United States in particular offers a long record of instances in which official racial classification has been shaped by forces other than methodological concerns (Lee 1993; Morning 2003; Wolfe 2001). The current format that

distinguishes Hispanics as an ethnic group but not a race; the inclusion of multiple sub-categories of the 'Asian' race option; and the retention of a 'Some other race' response are just a few examples of census features championed by political actors.

Consequently, it is not enough to appeal to methodological principles of logic, consistency, parsimony or clarity – nor to international precedent – when calling for change in census questionnaires. Political interpretation and agendas around the census must also be taken into account. More specifically, potential revisions that are suggested by cross-national comparison must address the policy concerns and motivations that shaped the current questionnaire. Are these political exigencies still salient or have they diminished in importance? Does the proposed revision solve or exacerbate the social problem in question, or do neither? Will the suggested change have other benefits or costs? How do they compare to the benefits and costs of the existing arrangement? Although survey design problems such as inconsistency or lack of clarity may not seem pressing enough to overhaul longstanding census items, we should not overlook the fact that they entail real costs: confusion, non-response, offense and lack of representation are just a few. In other words, the kinds of census design flaws that cross-national comparison reveals are most likely to be addressed if their implications for data quality are translated into the political language of social costs and benefits that has always shaped national census-taking.

International guidelines for the conduct of population censuses must also take both design imperatives and policy motivations into account. The most widely-applicable guidance is the United Nations Statistics Division's (1998) Principles and Recommendations for Population and Housing Censuses (Revision 1). In its discussion of ethnic enumeration, this document stresses the practical difficulty of proposing a common, cross-national approach to ethnic enumeration:

The national and/or ethnic groups of the population about which information is needed in different countries are dependent upon national circumstances. Some of the bases upon which ethnic groups are identified are ethnic nationality (in other words country or area of origin as distinct from citizenship or country of legal nationality), race, colour, language, religion, customs of dress or eating, tribe or various combinations of these characteristics. In addition, some of the terms used, such as 'race', 'origin' and 'tribe', have a number of different connotations. The definitions and criteria applied by each country investigating ethnic characteristics of the population must therefore be determined by the groups that it desires to identify. By the very nature of the subject, these groups will vary widely from country to country; thus, no internationally relevant criteria can be recommended. (p. 72)

Despite the United Nations' conclusion that 'no internationally relevant criteria can be recommended,' given the many ways that ethnicity is operationalized around the world (i.e., with measures such as language or dress), this analysis has revealed a great deal of commonality in official approaches to ethnic enumeration. And despite national variety in the groups recognized or the ethnicity terminology used, a broad class of ethnicity questions targeting communities of descent can be identified. Diversity in indicators of ethnicity – which as the U.N. rightly notes, are context-driven – does not preclude recognizing and analyzing them as reflections of a shared fundamental concept. Despite the different formulations used, such as 'race' or 'nationality,' their shared reference to communities of descent justifies both



academic and policy interpretation of them as comparable categorization schemes. Just as different countries might define ‘family’ membership differently, we can recognize that their varied enumeration approaches target an underlying, shared concept of kinship – and suggest census guidelines accordingly. In short, these findings challenge the United Nations conclusion that international guidance on ethnic enumeration is not possible.

The feasibility of proposing international guidelines on ethnic enumeration is an entirely separate matter, however, from the question of what recommendations should be made, including first and foremost any guidance about whether ethnicity should be a census item at all. The debate about the desirability of formal ethnic classification is a political one – and it is important and timely. In the United States, some public figures have called for the removal of racial categories from official state-level records, believing that government policies should not be informed by data on race (Morning and Sabbagh 2005). In some European countries, France in particular, the potential introduction of official ethnic classification has been hotly debated (Blum 2002; Simon and Stavo-Debaugé 2004). While supporters believe such categories are necessary to identify and combat discrimination, opponents fear that government adoption of such a classification scheme would divide the nation, stigmatize some groups, and generally bolster concepts of difference that have been closely associated with prejudice. Given such concerns, Zuberi’s (2005) admonition that ethnic categories not be used on censuses without a clear objective, and one that will not harm those groups traditionally stigmatized by such classifications, is essential. But as the French case illustrates, it can be difficult to ascertain the pros and cons of ethnic enumeration, as its likely impact may be highly contested. While the presentation of results on global classification practices cannot answer the normative questions posed here, empirical findings on the reach and uses of such categorization schemes should nonetheless be a meaningful resource that informs the important debate over whether populations should be enumerated by ethnicity at all.

**Acknowledgements** The author warmly thanks the following people and institutions for their contributions: Kevin Deardorff (U.S. Census Bureau); United Nations Statistical Division (Department of Economic and Social Affairs), Demographic and Social Statistics Branch (particularly Mary Chamie, Jeremiah Banda, Yacob Zewoldi, Margaret Mbogoni, Lisa Morrison-Puckett and intern Julia Alemany); International Programs Center, U.S. Census Bureau; Adrian Hayes (Australian National University); Caroline Persell and Sylvia Simson (New York University); Leslie Stone (Inter-American Development Bank); Gerald Haberkorn (Secretariat of the Pacific Community); Patrick Corr (Australian Bureau of Statistics); and anonymous reviewers. I also wish to thank the attendees at the following presentations of this research: U.S. Census Bureau Migration Speaker Series; Population Association of America; International Union for the Scientific Study of Population; American Sociological Association; and the Demographic and Social Statistics Branch (United Nations) Speaker Series. The initial version of this research was funded by the U.S. Census Bureau Immigration Statistics Branch. However, the conclusions – and the shortcomings – are solely those of the author.

**Open Access** This chapter is distributed under the terms of the Creative Commons Attribution Noncommercial License, which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## Appendix: Countries Included in Regional Groupings

Organizing scheme borrowed from United Nations Statistical Division

North America	South America	Africa	Europe	Asia	Oceania
<b>Anguilla*</b>	<b>Argentina*</b>	Algeria	Albania*	Afghanistan	<b>American Samoa*</b>
Antigua and Barbuda	<b>Bolivia*</b>	Angola	Andorra	<b>Armenia*</b>	<b>Australia*</b>
Aruba	<b>Brazil*</b>	Benin	Austria*	<b>Azerbaijan*</b>	<b>Cook Islands*</b>
Bahamas*	<b>Chile*</b>	Botswana*	Belarus*	Bahrain*	<b>Fiji*</b>
Barbados	Colombia	Burkina Faso	Belgium*	Bangladesh	French Polynesia*
<b>Belize*</b>	Ecuador	Burundi	Bosnia and Herzegovina	Bhutan	<b>Guam*</b>
<b>Bermuda*</b>	Falkland Islands (Malvinas)	Cameroon	<b>Bulgaria*</b>	Brunei Darussalam	<b>Kiribati*</b>
British Virgin Islands	French Guiana*	Cape Verde*	Channel Islands (Guernsey)*	Cambodia*	Marshall Islands
<b>Canada*</b>	<b>Guyana*</b>	Central African Republic	<b>Channel Islands (Jersey)*</b>	<b>China*</b>	<b>Micronesia (Federated States of)*</b>
Cayman Islands	<b>Paraguay*</b>	Chad	<b>Croatia*</b>	<b>Cyprus*</b>	<b>Nauru*</b>
<b>Costa Rica*</b>	<b>Peru*</b>	Comoros	Czech Republic*	Democratic People's Republic of Korea	<b>New Caledonia*</b>
Cuba	<b>Suriname*</b>	Congo	Denmark	East Timor*	<b>New Zealand*</b>
Dominica	Uruguay*	Cote d'Ivoire	<b>Estonia*</b>	<b>Georgia*</b>	Niue
Dominican Republic	<b>Venezuela*</b>	Democratic Republic of the Congo	Faeroe Islands	<b>Hong Kong*</b>	Norfolk Island
El Salvador		Djibouti	Finland*	<b>India*</b>	<b>Northern Mariana Islands*</b>
Greenland		Egypt*	France*	<b>Indonesia*</b>	Palau
Grenada		Equatorial Guinea	Germany	Iran	Papua New Guinea*
Guadeloupe		Eritrea	Gibraltar	<b>Iraq*</b>	Pitcairn
<b>Guatemala*</b>		Ethiopia	Greece*	Israel*	Samoa

(continued)

North America	South America	Africa	Europe	Asia	Oceania
Haiti*		Gabon	Holy See	Japan*	<b>Solomon Islands*</b>
<b>Honduras*</b>		Gambia	<b>Hungary*</b>	Jordan	<b>Tokelau*</b>
<b>Jamaica*</b>		<b>Ghana*</b>	Iceland	<b>Kazakhstan*</b>	<b>Tonga*</b>
Martinique		Guinea*	Ireland*	Kuwait*	<b>Tuvalu*</b>
<b>Mexico*</b>		Guinea-Bissau	Isle of Man*	<b>Kyrgyzstan*</b>	<b>Vanuatu*</b>
Montserrat		<b>Kenya*</b>	Italy*	<b>Lao People's Dem. Republic*</b>	Wallis and Futuna Islands*
Netherlands Antilles		Lesotho*	<b>Latvia*</b>	Lebanon	
Nicaragua*		Liberia	Liechtenstein*	<b>Macao*</b>	
<b>Panama*</b>		Libyan Arab Jamahiriya	<b>Lithuania*</b>	<b>Malaysia*</b>	
<b>Puerto Rico*</b>		Madagascar	Luxembourg*	Maldives*	
Saint Kitts and Nevis		Malawi*	Malta*	<b>Mongolia*</b>	
<b>Saint Lucia*</b>		Mali	Monaco*	Myanmar	
Saint Pierre and Miquelon		Mauritania	Netherlands	<b>Nepal*</b>	
Saint Vincent and the Grenadines		<b>Mauritius*</b>	Norway*	Occupied Palestinian Territory*	
<b>Trinidad and Tobago*</b>		Morocco*	<b>Poland*</b>	Oman	
Turks and Caicos Islands		<b>Mozambique*</b>	Portugal*	Pakistan*	
<b>United States*</b>		Namibia*	<b>Republic of Moldova*</b>	<b>Philippines*</b>	
<b>U.S. Virgin Islands*</b>		Niger	<b>Romania*</b>	Qatar	
		Nigeria	<b>Russian Federation*</b>	Republic of Korea*	
		Réunion	San Marino	Saudi Arabia	
		Rwanda	Slovakia	<b>Singapore*</b>	
		Saint Helena	<b>Slovenia*</b>	<b>Sri Lanka*</b>	
		Sao Tome and Principe	Spain*	Syrian Arab Republic	
		<b>Senegal*</b>	Svalbard and Jan Mayen Islands	<b>Tajikistan*</b>	
		Seychelles*	Sweden	Thailand*	
		Sierra Leone	Switzerland*	Turkey*	

(continued)

North America	South America	Africa	Europe	Asia	Oceania
		Somalia	<b>Former Yugoslav Republic of Macedonia*</b>	<b>Turkmenistan*</b>	
		<b>South Africa*</b>	<b>Ukraine*</b>	United Arab Emirates	
		Sudan	<b>United Kingdom*</b>	<b>Uzbekistan*</b>	
		Swaziland*	<b>Yugoslavia*</b>	<b>Vietnam*</b>	
		Togo		Yemen*	
		Tunisia			
		Uganda			
		United Rep. of Tanzania*			
		Western Sahara			
		<b>Zambia*</b>			
		<b>Zimbabwe*</b>			

Countries marked with an asterisk \* are those whose censuses from the 1995–2004 period were used for this study; countries in bold include an ethnicity question on the census

## References

- Alba, R. (1990). *Ethnic identity: The transformation of white America*. New Haven: Yale University Press.
- Almey, M., Pryor, E.T. & White, P.M. (1992, May 23). National census measures of ethnicity in the Americas. In *Proceedings of conference on the peopling of the Americas*. Vera Cruz, Mexico: International Union for the Scientific Study of Populations.
- Anderson, M. J. (1988). *The American census: A social history*. New Haven: Yale University Press.
- Bean, F. D., & Tienda, M. (1987). *The Hispanic population of the United States*. New York: Russell Sage.
- Blum, A. (2002). Resistance to identity categorization. In D. I. Kertzer & D. Arel (Eds.), *Census and identity: The politics of race, ethnicity, and language in national censuses in France* (pp. 121–147). Cambridge: Cambridge University Press.
- Blum, A., & Gousseff, C. (1996). Statistiques ethniques et nationales dans l'empire Russe et en URSS. In J. L. Rallu & Y. Courbage (Eds.), *Démographie et Ethnicité*. Paris: INED.
- Eberhardt, P. (2003). *Ethnic groups and population changes in twentieth-century Central-Eastern Europe: History, data, and analysis*. Armonk: M.E. Sharpe.
- Hollinger, D. A. (1998). National culture and communities of descent. *Reviews in American History*, 26(1), 312–328.
- Kertzer, D. I., & Arel, D. (2002a). Census and identity: The politics of race, ethnicity, and language. In D. I. Kertzer & D. Arel (Eds.), *National censuses*. Cambridge: Cambridge University Press.
- Kertzer, D. I., & Arel, D. (2002b). Censuses, identity formation, and the struggle for political power. In *Census and identity: The politics of race, ethnicity, and language in national censuses* (pp. 1–42). Cambridge: Cambridge University Press.

- Lee, S. M. (1993). Racial classifications in the U.S. Census: 1890–1990. *Ethnic and Racial Studies*, 16(1), 75–94.
- Morning, A. (2003). New faces, old faces: Counting the multiracial population past and present. In H. DeBose & L. Winters (Eds.), *New faces in a changing America: Multiracial identity in the 21st century* (pp. 41–67). Thousand Oaks: Sage.
- Morning, A. (2008). Ethnic classification in global perspective: A cross-national survey of the 2000 census round. *Population Research and Policy Review*, 27(2), 239–272.
- Morning, A., & Sabbagh, D. (2005). From sword to plowshare: Using race for discrimination and antidiscrimination in the United States. *International Social Science Journal*, 57(183), 57–73.
- Nobles, M. (2000). *Shades of citizenship: Race and the census in modern politics*. Stanford: Stanford University Press.
- Rallu, J.-L., Piché, V., & Simon, P. (2004). Démographie et ethnicité: Une relation ambiguë. In G. Caselli, J. Vallin, & G. Wunsch (Eds.), *Démographie: Analyse et synthèse* (pp. 481–515). Paris: Institut National d'Etudes Démographiques.
- Simon, P., & Stavo-Debaugé, J. (2004). Les politiques anti-discrimination et les statistiques: Paramètres d'une incohérence. *Sociétés Contemporaines*, 53, 57–84.
- Skerry, P. (2000). *Counting on the census? Race, group identity, and the evasion of politics*. Washington, DC: Brookings Institution Press.
- Statistics Canada and U.S. Census Bureau. (1993). *Challenges of measuring an ethnic world: Science, politics, and reality*. Washington, DC: U.S. Government Printing Office.
- United Nations. (1998). *Principles and recommendations for population and housing censuses (ST/ESA/STAT/SER.M/67/Rev.1)*. Department of Economic and Social Affairs, Statistics Division.
- United Nations Statistics Division. (2003). *Ethnicity: A review of data collection and dissemination* (Unpublished document). Demographic and Social Statistics Branch, United Nations Statistics Division.
- Wagley, C. (1965). On the concept of social race in the Americas. In D. B. Heath & R. N. Adams (Eds.), *Contemporary cultures and societies in Latin America* (pp. 531–545). New York: Random House.
- Wolfe, P. (2001). Land, labor, and difference: Elementary structures of race. *The American Historical Review*, 106(3), 866–905.
- Zuberi, T. (2005, June 9). The limits of racial and ethnic data. In *Demographic and social statistics branch seminar*. New York: United Nations Statistics Division.