



Reading Colonial Census Categories Against the Grain of Community Self-Description: A Methodological Framework

Ngozi Mthembu^{1*}, Karine Gebremedhin², Chioma Pedersen³

¹Department of Social Sciences, Transylvanian Institute of Social Sciences, Georgia

²Department of Social Sciences, Baikal Center for Social Sciences, New Zealand

³Department of Social Sciences, Loire Research Institute of Social Sciences, Belgium

Article Info

Volume 1, Issue 1, 2026

doi: [10.62715/ijhcs.2026.v1.i1.pp13-18](https://doi.org/10.62715/ijhcs.2026.v1.i1.pp13-18)

Received: 18 September 2025

Accepted: 08 November 2025

Published: 10 January 2026

Abstract

Colonial censuses sorted populations into castes, tribes, races and religions, and historians have shown that these schemes shaped the societies they claimed to describe. That insight has produced two opposed habits of reading. One treats census categories as inventions and discards the tables; the other mines the tables for numbers and ignores how they were made. This article proposes a framework for a third course, in which a category is read against what the people placed in it said about themselves. The framework has six steps: reconstruct the enumeration as a procedure, trace each category across successive censuses, read the reports for signs of friction, assemble an independent body of community self-description, compare the two in a four-way grid, and interpret the differences with attention to how classification and self-understanding acted on each other over time. The steps are illustrated with reference to the well-documented censuses of British India and Malaya, using published scholarship only. The framework is a proposal for researchers working on any colonial enumeration. It has not been applied to a fresh body of material here, and its limits, above all the unequal survival of community voices, are stated.

Keywords: Colonial census, Classification, Caste, Self-description, Archival method, South Asia

© 2026 Ngozi Mthembu et al. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction

The census is among the most used and most distrusted sources for the history of colonised societies. It offers what little else does: a count of everyone, repeated at intervals, broken down by place. It also carries the assumptions of the officials who designed it. Since the essay by Cohn (1987) on the census and objectification in South Asia, historians have understood that asking every person to state a caste, tribe or religion, and then publishing the totals, did more than record identities. It fixed names, drew boundaries and gave groups a number to argue with.

The argument has been taken in a strong form, that colonial knowledge made caste into the central institution of Indian society (Dirks, 2001), and it has been qualified by scholars who point to enumeration

* Corresponding author: Ngozi Mthembu, Department of Social Sciences, Transylvanian Institute of Social Sciences, Georgia. E-mail: ngozi.mthembu@transylvanian.edu

© 2026 Ngozi Mthembu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

and group identity before British rule (Guha, 2003; Peabody, 2001). Researchers who open a census volume today inherit both positions and often little guidance on what to do with the page in front of them.

This article sets out a framework for reading census categories against the self-descriptions of the communities concerned. It is methodological: it proposes a sequence of steps and the reasoning for each, drawing its illustrations from published work. It reports no new archival findings.

2. Background

Three strands of scholarship inform the design: work on what enumeration does to the people counted, work on how official archives may be read for the voices of the governed, and work on the social effects of classification in general.

2.1. Enumeration and its effects

Anderson (1991) observed that the colonial census allowed no fractions and no blanks: everyone had one place, and only one, in the grid. Appadurai (1993) described how number became a language of rule in British India, and Kaviraj (1992) contrasted the fuzzy communities of earlier periods, whose members did not know how many they were, with the enumerated communities of modern politics. The Indian case is the best documented. Provincial counts in the 1860s and early 1870s were followed from 1881 by a synchronous census of the whole of British India taken every ten years, and caste was tabulated in detail until 1931 (Cohn, 1987). Comparative studies have since shown the same dynamics in censuses across the world (Kertzer and Arel, 2002).

2.2. Reading along and against the grain

To read against the grain is to use a document for purposes its authors did not intend. Guha (1983) showed how reports written to justify the suppression of peasant risings could be turned to reveal the reasoning of the insurgents. Spivak (1988) warned that such recovery has limits, since the governed appear in official prose only as the governors understood them. Stoler (2009) added a prior requirement: before reading against an archive one must read along it, learning its conventions, its anxieties and the questions it could not settle. The framework below adopts that order.

2.3. Classification and feedback

Hacking (1986) argued that categories of people differ from categories of things because people respond to being classified. A label creates new ways to be a person, those labelled adapt to or resist it, and the category changes in turn. Bowker and Star (1999) drew attention to what every scheme leaves over: the residual box marked "other", where whatever does not fit is put out of sight. Both ideas apply directly to a census, which is revised every decade in the light of the last count and of the protests it provoked.

3. The Proposed Framework

The framework has six steps. The first three concern the census itself; the fourth assembles a separate body of evidence; the fifth and sixth bring the two together.

3.1. Step 1: Reconstruct the enumeration as a procedure

A published table is the end of a chain. Before it came a schedule with columns and headings, printed instructions to enumerators, a house-to-house visit in which a householder often answered for everyone under the roof, and an office where clerks sorted the returns into the approved list. The researcher should establish each link: what question was asked and in which language, who was likely to have answered, what the enumerator was told to do with an unfamiliar reply, and who decided how returns were grouped. The administrative reports that accompany most census volumes describe these arrangements. Collections such as that edited by Barrier (1981) show how much the answers to these questions varied between provinces and decades.

3.2. Step 2: Trace each category through successive censuses

Categories have histories. A name may appear, divide, be absorbed into another or vanish, and its position in the scheme may shift. Hirschman (1987) followed the classifications used in the censuses of colonial and independent Malaya and showed how lists of peoples hardened over time into a small set of races. In India the commissioner of the 1901 census, Risley (1908), sought to arrange castes in order of social precedence, an experiment later censuses retreated from. Setting the lists side by side shows which categories were stable and which were being worked on, and the latter are where reading against the grain is most productive.

3.3. Step 3: Read the reports for friction

Census commissioners wrote at length about their difficulties, and those passages are the most valuable in the volumes. They complain of vague or unclassifiable returns, of people who gave an occupation or a sect when asked for a caste, of groups returning themselves under new and grander names, and of petitions pressing for a change of title or rank. Each complaint marks a point where the scheme met a self-description it could not absorb. The researcher should extract these passages systematically, together with footnotes to tables, notes on reclassification and the size of residual headings, and treat them as an index of disagreement.

3.4. Step 4: Assemble community self-description

The second body of evidence must come from outside the census, or it will merely echo it. Table 1 lists the main kinds of source and the caution each requires. Petitions and memorials addressed to census officials are the most direct, and were numerous in India after 1901, when caste associations organised to contest the ranks assigned to them (Carroll, 1978). Caste and community histories, association journals, the vernacular press, genealogies, temple and shrine records and, for recent periods, oral tradition extend the range. Histories of single communities show how such materials can be combined (Bayly, 1999).

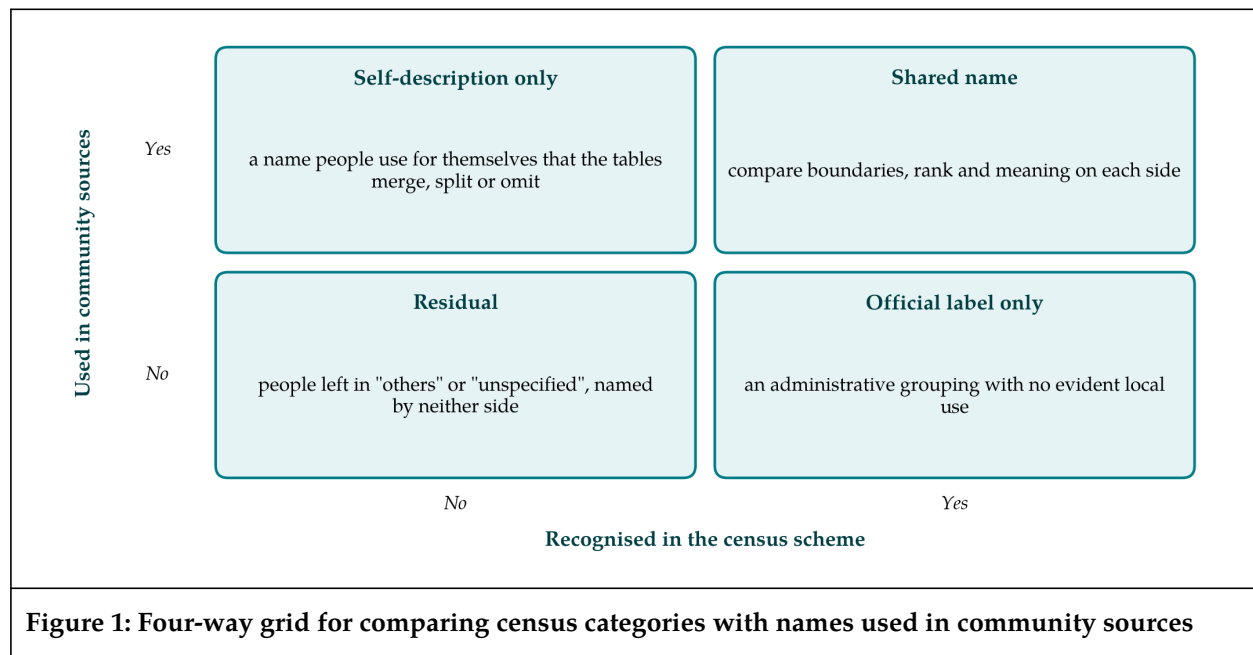
Table 1: Sources of community self-description and the cautions that apply to each		
Source	What it offers	Caution
Petitions to census officials	Explicit claims about name, origin and rank	Framed in the terms officials would accept
Association journals and histories	Sustained argument about who belongs	Voice of educated, organised leaders
Vernacular newspapers	Debate at the time of each count	Uneven survival; urban readership
Genealogies and origin narratives	Terms of kinship, lineage and place	Often compiled or revised late
Temple, shrine and guild records	Membership and precedence in practice	Local in scope; access may be restricted
Oral tradition and testimony	Names and distinctions in everyday speech	Shaped by later politics of identity

None of these is an unmediated voice. A petition to the census commissioner is written to persuade him, and a caste history printed in the 1920s may already draw its vocabulary from earlier census reports. That circularity is part of what the framework sets out to trace.

3.5. Step 5: Compare in a four-way grid

With both bodies of evidence in hand, each name is placed in a grid according to whether it occurs in the census scheme, in community sources, in both or in neither. Figure 1 shows the four positions. A shared name invites comparison of what each side meant by it. A name found only in community sources points to people the tables merged with others or split apart. An official label with no local use suggests an administrative convenience. The fourth cell holds those entered as "others" or "unspecified", whom

neither the census nor the surviving community sources name, and whose presence should at least be noted.



3.6. Step 6: Interpret the differences over time

The grid is a snapshot, and the last step sets it in motion. For each category the researcher asks how the position changed from one census to the next and what moved it: an official decision, a campaign by an association, a change in the law that attached benefits or disabilities to a label. Here the idea of feedback between classification and the classified earns its place (Hacking, 1986). The history of caste in India offers many instances in which groups adopted new names, claimed descent from ancient orders or reformed their customs in ways that the census both prompted and recorded (Bayly, 1999). The interpretation should also remain open to continuities with earlier practice, since communities had been listed, ranked and taxed by pre-colonial states and brought their own understandings to the colonial count (Peabody, 2001).

4. Applying the Framework

The framework is meant to scale. A dissertation might follow one district through five censuses; an article might take one contested category through two. In either case the work is best begun with the reports and schedules of a single census read whole, before any table is used for its figures. Researchers working on Africa, Southeast Asia or the Caribbean will find different vocabularies, with tribe, race or nationality in the place that caste holds in the Indian volumes, but the same chain of schedule, instruction, return and table, and often the same pattern of protest (Kertzer and Arel, 2002). The numbers themselves need not be thrown away. Once the making of a category is understood, its totals can be used with stated reservations.

A simple working record helps. For each category the researcher keeps a sheet that notes the wording of the question in each census, the instruction given to enumerators, the heading under which returns were tabulated, the passages of the report that discuss it, and the community sources that use or reject the name. Laid side by side, such sheets make change over time visible and show at a glance where evidence from one side is missing. They also discipline the temptation to quote a vivid petition as if it settled what a whole community thought.

5. Limitations and What Remains to Be Shown

The framework is proposed and has not been validated here by application to new material. Its chief weakness is built into the evidence. Self-description survives mainly from groups with literate leaders,

printing presses and the confidence to petition, and these were seldom the poorest. The grid will therefore be fullest for upwardly mobile communities and emptiest for those who most needed a hearing, and a researcher who forgets this will mistake the voice of an association for the voice of a people (Spivak, 1988).

Two further cautions apply. Communities were never unanimous, and "self-description" may conceal disputes between sections, generations and sexes within a group. And the method is laborious: reading reports in full and searching vernacular print takes far longer than extracting a table. Whether the gain in understanding repays that labour across different regions and periods is a question that only repeated use can settle.

6. Conclusion

Census categories were neither pure inventions nor neutral descriptions. They were produced in an exchange, unequal but real, between officials who needed a tidy list and people who had their own names for themselves. The six steps proposed here are a way of recovering that exchange from the census and from sources outside it, and of using colonial counts critically without either trusting or dismissing them.

Conflict of interest

To be declared by the author(s) before publication.

Funding

To be declared by the author(s) before publication.

References

- Anderson, B. (1991). *Imagined Communities: Reflections on the Origin and Spread of Nationalism* (revised edn). Verso, London.
- Appadurai, A. (1993). *Number in the colonial imagination*. In: Breckenridge, C.A. and van der Veer, P. (eds), *Orientalism and the Postcolonial Predicament: Perspectives on South Asia*. University of Pennsylvania Press, Philadelphia.
- Barrier, N.G. (ed.) (1981). *The Census in British India: New Perspectives*. Manohar, New Delhi.
- Bayly, S. (1999). *Caste, Society and Politics in India from the Eighteenth Century to the Modern Age*. Cambridge University Press, Cambridge.
- Bowker, G.C. and Star, S.L. (1999). *Sorting Things Out: Classification and Its Consequences*. MIT Press, Cambridge, MA.
- Carroll, L. (1978). *Colonial perceptions of Indian society and the emergence of caste(s) associations*. *Journal of Asian Studies*, 37(2): 233-250.
- Cohn, B.S. (1987). *The census, social structure and objectification in South Asia*. In: *An Anthropologist among the Historians and Other Essays*. Oxford University Press, Delhi.
- Dirks, N.B. (2001). *Castes of Mind: Colonialism and the Making of Modern India*. Princeton University Press, Princeton.
- Guha, R. (1983). *The prose of counter-insurgency*. In: Guha, R. (ed.), *Subaltern Studies II: Writings on South Asian History and Society*. Oxford University Press, Delhi.
- Guha, S. (2003). *The politics of identity and enumeration in India c. 1600-1990*. *Comparative Studies in Society and History*, 45(1): 148-167.
- Hacking, I. (1986). *Making up people*. In: Heller, T.C., Sosna, M. and Wellbery, D.E. (eds), *Reconstructing Individualism: Autonomy, Individuality, and the Self in Western Thought*. Stanford University Press, Stanford.
- Hirschman, C. (1987). *The meaning and measurement of ethnicity in Malaysia: An analysis of census*

- classifications. *Journal of Asian Studies*, 46(3): 555-582.
- Kaviraj, S. (1992). *The imaginary institution of India*. In: Chatterjee, P. and Pandey, G. (eds), *Subaltern Studies VII: Writings on South Asian History and Society*. Oxford University Press, Delhi.
- Kertzer, D.I. and Arel, D. (eds) (2002). *Census and Identity: The Politics of Race, Ethnicity, and Language in National Censuses*. Cambridge University Press, Cambridge.
- Peabody, N. (2001). *Cents, sense, census: Human inventories in late precolonial and early colonial India*. *Comparative Studies in Society and History*, 43(4): 819-850.
- Risley, H.H. (1908). *The People of India*. Thacker, Spink and Co., Calcutta.
- Spivak, G.C. (1988). *Can the subaltern speak?* In: Nelson, C. and Grossberg, L. (eds), *Marxism and the Interpretation of Culture*. University of Illinois Press, Urbana.
- Stoler, A.L. (2009). *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense*. Princeton University Press, Princeton.