Gerwin van Schie


Abstract

In this article, categorisation based on origins is investigated in the Dutch census reports produced between 1899 and 2018. Through this analysis, I will argue that the conceptual pair of autochthon and allochthon (until 2017 used to describe ‘native’ and ‘non-native’ Dutch people) is not a social construct, but rather a sociotechnical construct. Through this analysis, it becomes clear that where the category of the ‘foreigner’ at the beginning of the twentieth century was mainly related to a passport and place of birth, later national origins became a racially tainted concept. Pre-1980, counting machines at CBS only counted people’s information as self-described; since 1980, computers have determined the identity of Dutch citizens. Today, governmental statistics based on the origins of Dutch people are available online for free and can be used by third parties for a variety of applications. Through this historical analysis, it becomes clear that the current practices are not the ‘standard’ way of categorisation, but rather one of the possibilities that are historically, socially and technologically situated.

Keywords: Data; census; discrimination; racism; categorisation

Origin (n.): c.1400, ‘ancestry, race,’ from Latin Originem (nominative origo) ‘a rise, commencement, beginning, source; descent, lineage, birth,’ from stem of oriri ‘arise, rise, get up; appear above the horizon, become visible; be born, be descended, receive life.’

Introduction

As the centre of the Dutch governmental data infrastructure, Statistics Netherlands (Centraal Bureau voor de Statistiek, in short: CBS) is the key source of information for governmental policymaking. An independent administrative body of the Dutch government, CBS ‘performs public service tasks, but operates independently and not under the direct authority of a Dutch ministry’. Its mission is to compile and publish both Dutch and European statistics, including demographic information. An important part of these demographic data is that pertaining to the national origins of Dutch citizens, which is often used as variables when measuring all sorts of social and economic phenomena, ranging from welfare distribution to crime.
In their efforts to increase transparency and create more opportunities for governmental organisations, NGOs and companies to use existing statistical information, many governments around the world have started open data initiatives to make statistics available for the wider public. ‘Open government data’ is understood here as a set of technologies that makes it possible for companies, organisations, journalists and other individuals to conduct research or build applications based on data collected and processed with public money. These data initiatives are symptomatic of the emergence of what has been called, amongst other things, a ‘datafied society’. Datafication can be defined as the process of translating everyday activities and social interactions into quantified information. In accordance with its ‘open data’ policy, CBS made 3,500 datasets available online through its open data portal in 2014. This portal has its own interface in the form of the CBS-Statline website, but also features an application programming interface (API), which allows third parties to directly tap into the CBS system. As a result, statistics relating to the national origins of Dutch citizens are now regularly used as indicators in web applications, such as the Leefbaarometer of the Ministry of the Interior and Kingdom Relations, an interactive online map of the Netherlands that shows the livability of an area down to street level, and the Allochtonenmeter, a commercial application that returns the numbers of Western allochthones (Westerse allochtonen) and non-Western allochthones (niet-Westerse allochtonen) in a postal code area.

Until 2017, autochthons and allochthons were the two labels of the categories through which people were clustered by CBS based on their country of origin. These terms linguistically link origins with birth soil in the form of nationality. The Greek word chtõn means ‘earth’ or ‘soil’ and the prefixes auto and allo refer to ‘same’ and ‘different’ respectively, creating the words autochthon and allochthon, referring to people originating from the same and different ground or soil. In addition, a distinction is made between Western allochthones and non-Western allochthones, as indicative of ‘differences in socio-economic and cultural position’. As noted in a 2016 report by the Scientific Committee for Government Policy (Wetenschappelijke Raad voor Regeringsbeleid), the Netherlands is the only country in the world that employs these particular categories based on national origins for its own citizens. Gloria Wekker reads the inability of the Dutch to see these categorisations as discriminatory and racist as the result of a sense of so-called ‘white innocence’. However, as a ‘régime of truth’ that is not only enacted through governmental policy but also in scientific, political and public discourse, the categorisation system functions as an epistemological tool that also actively shapes reality. With reference to Alain Desrosières, Rottenburg and Merry state that:

collective objects or aggregates of individuals thus can come to ‘hold’, to be accepted as real, at least for a while. [Desrosières] explained, ‘When the actors can rely on objects thus constructed, and these objects resist the tests intended to destroy them, aggregates do exist – at least during the period and in the domain in which these practices and tests succeed.’ In other words, if these aggregates work in an institutionalized practice, they will be accepted as real. It is their institutionalization that makes them real.

In light of the institutional acceptance of constructs such as autochthon and allochthon, I would like to analyse the ontological aspects of the classification system by critically
investigating the sociotechnical origins of race-ethnic classification. In the context of datafication, I understand ontology as a reflection of:

a social construction of reality, defined in the context of a specific epistemic culture as sets of norms, symbols, human interactions, and processes that collectively facilitate the transformation of data into knowledge.\textsuperscript{20}

In this article, I trace the history of the Dutch data ontology based on national origins that has become part of the classification system of CBS, by means of an infrastructural inversion. This methodology is explained by science and technology studies scholar Geoffrey C. Bowker and sociologist Susan Leigh Star as one of recognising ‘the depths of interdependence of technical networks and standards, on the one hand, and the real work of politics and knowledge production on the other’\textsuperscript{21} as a means to not only investigate the political and epistemological underpinnings of the definitions and standards, but also the systems and institutions in which they materialise and come into effect. This methodology, invented for and used in research on knowledge production in science and policy,\textsuperscript{22} starts at the very basis on which a system is built: the categories and definitions that are the foundation of an infrastructure and the historic, cultural and material contexts in which they emerge.

In what follows, I track the generation, curation and distribution of information regarding the origins of people residing in the Netherlands. As sources, I use the census reports produced in the Netherlands between 1899 and 2017— that is, those of 1899, 1909, 1920, 1930, 1947, 1960, 1971, 2001 and 2011.\textsuperscript{23} I focus specifically on the parts that relate to origins in the form of nationality and place of birth, the ways in which their definition has been informed by ethnic and racial discourses in the Netherlands over time, and the systems that are used for counting and calculation. The inclusion of the material means of this ‘making up people’ makes this not only a social history, but a media history as well.\textsuperscript{24} Despite changes in labels over time, all categorisation policies throughout the investigated time period have considered the nationality of citizens or their parents, considering specifically their (former) passports; therefore, I will systematically use the phrase ‘national’ rather than ‘ethnic origins’ in referring to census results – even though ideas on race, ethnicity and nationality are clearly entwined.\textsuperscript{25}

In his seminal work on nationalism, Benedict Anderson highlights the census, together with the map and the museum, as particularly important media in the construction of the modern nation-state. Censuses have historically been an important tool for the government of populations.\textsuperscript{26} These governmentalities are largely informed by governmental categorisation practices imposed on citizens through census practices.\textsuperscript{27} Peter Skerry explains that such categorisation practices make census-taking a political endeavour:

\[ T \]he census is not simply about counting in the abstract, but about counting certain group characteristics whose salience is determined through politics. The census is therefore concerned with establishing the boundaries that define the parts, or groups, that constitute the political community. This is inescapably a political undertaking.\textsuperscript{28}
The politics of the categorisation practices of the Dutch government have already been dealt with in their policy context, both in terms of racial and ethnic connotations and of their performance of inclusion and exclusion practices with regard to Dutch society. With this intervention, I would also like to place the ontology of the governmental information apparatus in its historical context. By discussing technological and societal developments together, I will make an argument for shifting the understanding of race-ethnic categorisation in terms of the conceptual pair of the autochthon/allochthon from social constructs to sociotechnical constructs. These constructs, as epistemological tools, also have performative power: they are capable of shaping the realities of Dutch citizens with different origins in different ways, possibly causing data-driven discrimination and technologically mediated inequalities when used in third-party systems such as the Leefbaarometer or Allochtonenmeter. By employing a historically comparative perspective, I both denaturalise contemporary categorisation practices (by showing that there are many ways of counting people) and contextualise them (by showing that current practices are the result of a long history).

I will perform my analysis by distinguishing between three time periods. First, I will discuss the years 1899–1940, from the foundation of CBS until World War II, which can be characterised by a process going from counting by hand to mechanisation. The second period, 1945–1980, extending from the end of World War II to the end of a period of labor immigration and decolonisation in Suriname and Indonesia, is characterised by a move from mechanised counting to digital counting. The final era, from 1980 onwards, is characterised by a national discussion on and scholarly investigation of immigrant integration and the multicultural society, combined with the increasing datafication of governmental information. The era can be marked by one specific continuity that is particularly relevant for this sociotechnical history: the census has become ‘virtual’. This means that people themselves are no longer counted, but merely their data, taken from bureaucratic systems. The use of the three different periods is mainly intended to structure this article; I am not claiming that there are any ‘hard breaks’ in social or technical development that determine what happens in these periods.

‘Foreigners’ (1899–1940)

CBS was founded in 1899 and tasked with producing independent, objective and scientifically valid statistics. Between 1899 and 1940, the institution was able to largely centralise the production of governmental statistics – a practice that was previously spread over different government departments and municipalities. In addition to institutional reforms, CBS also initiated technological developments in counting and statistics.

In 1899, the census was performed through self-reporting by hand with pen or pencil on a pre-printed form with different paper colours for men, women and children. These forms were then counted by hand. In 1916, CBS was the first Dutch organisation to instal the Hollerith punch-card system in order to effectively count imports and exports, a previously rather time-consuming activity. For the 1920 census, the Hollerith machine was considered, but for
unknown reasons, the ‘classi-compteur’, another mechanical counting device, was chosen for this purpose. The handwritten census forms were counted by women using this classi-compteur. However, from 1930 onwards, the Hollerith system was also used for the census. Census forms, converted to individual punch cards for each person, were handled by CBS employees. These cards were then counted automatically with machines at the CBS building. The results of the censuses were printed in reports with tables in both Dutch and French (a hangover from Napoleonic times).

In the censuses of 1899, 1909, 1920 and 1930, a similar structure of presentation in reports can be seen. Categories relating to national origins are split into two groups (see Figure 1): origins are counted both in terms of nationality (‘nationaliteit’) and in terms of place of birth (‘geboorteplaats’). All nationalities and places of birth are counted individually and their counts are reported in the attachments of the census reports.

The countries that are considered most important are featured in the main body of the report. In 1899, nationalities are not ordered alphabetically in the main report, but by number of people in each category, ranging from almost 32,000 people from Germany to 233 people from Italy. From 1909 onwards, the list is extended with eleven extra countries. With the exception of Austria–Hungary, a dual monarchy that fell apart in 1918, the top five countries remain the same. As of this point in time, the order of the countries and their clustering seems to be based roughly on their proximity to the Netherlands and the number of immigrants present (in that order). Another decisive factor in the order of countries seems to be the way in which the documents are laid out, in that the available space on a printed page (there is only room for 16 countries in total) and the page break in the middle split the table in two sets of eight columns (see Figure 2). The first eight countries are neighbouring countries (Germany, Belgium, Great Britain and Ireland, and France) and other European countries from which larger numbers of people migrated to the Netherlands (Italy, Switzerland, Austria and Hungary). The following eight columns are reserved for the remaining countries, ordered roughly by the number of immigrants (from Russia, the United States, Sweden, Norway, Denmark, Luxembourg, Turkey and, depending on the year, Transvaal Colony or Poland, and Greece or Czechoslovakia). The residual categories remain the same during this period: they contain people from ‘other countries’ (a category denoting the totals of the non-featured countries that are mentioned in the attachments), stateless people (mainly due to World War I and other conflicts in pre-World War II Europe), and people with an ‘unknown’ nationality.

The ‘place of birth’ category in this period of time is rather locally defined. The first two categories refer to people born in the same municipality as where they live, or a different municipality within the same province. Given the fact that, before the 1930s, the Netherlands had a large number of municipalities with a population of fewer than 500 inhabitants, this category concerns rather small-scale entities. The third ‘place of birth’ category, ‘different province in the Kingdom’, likewise refers to places within in the Dutch territories. Technically speaking, this should also include the colonies, since they were part of the Kingdom of the Netherlands at the time; however, this is not what is meant here, as the fourth category focuses on those territories (separating them off). Until the census of 1920, the Dutch Antilles, the Dutch East Indies

<table>
<thead>
<tr>
<th>Census year</th>
<th>1899</th>
<th>1909</th>
<th>1920</th>
<th>1930</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nationality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Netherlands</td>
<td>The Netherlands</td>
<td>The Netherlands</td>
<td>The Netherlands</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Germany</td>
<td>Germany</td>
<td>Germany</td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>Belgium</td>
<td>Belgium</td>
<td>Belgium</td>
<td></td>
</tr>
<tr>
<td>Great Britain and Ireland</td>
<td>Great Britain and Ireland</td>
<td>Great Britain and Ireland</td>
<td>Great Britain and Ireland</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>France</td>
<td>France</td>
<td>France</td>
<td></td>
</tr>
<tr>
<td>Austria-Hungary*</td>
<td>Italy</td>
<td>Italy</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>Switzerland</td>
<td>Switzerland</td>
<td>Switzerland</td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>Austria</td>
<td>Austria</td>
<td>Austria</td>
<td></td>
</tr>
<tr>
<td>Hungary</td>
<td>Hungary</td>
<td>Hungary</td>
<td>Hungary</td>
<td></td>
</tr>
<tr>
<td>Russia</td>
<td>Russia</td>
<td>Russia</td>
<td>Russia</td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>United States</td>
<td>United States</td>
<td>United States</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>Sweden</td>
<td>Sweden</td>
<td>Sweden</td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>Norway</td>
<td>Norway</td>
<td>Norway</td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>Denmark</td>
<td>Denmark</td>
<td>Denmark</td>
<td></td>
</tr>
<tr>
<td>Luxembourg</td>
<td>Luxembourg</td>
<td>Luxembourg</td>
<td>Luxembourg</td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td>Turkey</td>
<td>Turkey</td>
<td>Turkey</td>
<td></td>
</tr>
<tr>
<td>Transvaal Colonyb</td>
<td>Poland</td>
<td>Poland</td>
<td>Poland</td>
<td></td>
</tr>
<tr>
<td>Greece</td>
<td>Czechoslovakiac</td>
<td>Czechoslovakia</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Residual categories</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other countries</td>
<td>other countries</td>
<td>other countries</td>
<td>other countries</td>
<td></td>
</tr>
<tr>
<td>unknown</td>
<td>stateless</td>
<td>stateless</td>
<td>stateless</td>
<td></td>
</tr>
<tr>
<td>stateless</td>
<td>unknown</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
</tr>
<tr>
<td><strong>Place of Birth</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current municipality of residence</td>
<td>Current municipality of residence</td>
<td>Current municipality of residence</td>
<td>Current municipality of residence</td>
<td></td>
</tr>
<tr>
<td>Different municipality in the same province</td>
<td>Different municipality in the same province</td>
<td>Different municipality in the same province</td>
<td>Different municipality in the same province</td>
<td></td>
</tr>
<tr>
<td>Different province in the Kingdom</td>
<td>Different province in the Kingdom</td>
<td>Different province in the Kingdom</td>
<td>Different province in the Kingdom</td>
<td></td>
</tr>
<tr>
<td>In one of the Dutch Colonies</td>
<td>In one of the Dutch Colonies</td>
<td>In one of the Dutch Colonies</td>
<td>Dutch East Indies, Suriname or Curaçao</td>
<td></td>
</tr>
<tr>
<td>A foreign country</td>
<td>A foreign country</td>
<td>A foreign country</td>
<td>A foreign country</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td></td>
</tr>
</tbody>
</table>

a) Austria–Hungary was a dual monarchy that roughly comprised the current territories of Austria, Hungary, Czech Republic, Slovakia and Bosnia Herzegovina between 1867 and 1918.

b) Transvaal-colony was a former colony of Great Britain between 1877 and 1910, part of what is now South Africa. Most likely, this category only refers to white former inhabitants (mostly of Dutch or English descent) of Transvaal Colony, since black people were exempt from the right of citizenship. John Abraham Jacob De Villiers, The Transvaal (London: Chatto & Windus, 1896), 23–24. http://archive.org/details/transvaal00devi.

c) Czechoslovakia roughly comprised the current territories of Czech Republic and Slovakia, between 1918 and 1992.
and Suriname are referred to as ‘the Dutch colonies’. From 1930 onwards, these countries are instead referred to by their names. The final categories are ‘foreign countries’ (meaning all territories outside of the Netherlands and the Dutch colonies) and ‘unknown’ places. At the end of the 1930s, a census was planned for 1940, which was cancelled due to the German occupation in World War II.

The abovementioned time period is characterised by a relatively high measure of autonomy of the people counted, who provided information themselves, and could therefore, theoretically, also provide ‘inaccurate’ information. The definitions regarding who counts as Dutch and who does not are quite inclusive: people born in the Netherlands are considered Dutch, regardless of the origins of their parents. The local nature of the categorisations also seems to indicate that CBS considered migration to be not only a transnational phenomenon. By also taking into account province and municipality of birth, the reports demonstrate that CBS considered ‘internal’ migration to be a significant statistic as well.

**The Invention of the ‘Allochthon’ (1945–1980)**

After World War II, a census was initiated as soon as possible, because the government needed a detailed picture of the state of affairs in order to rebuild the nation. During this 1947 census, CBS still made extensive use of the punch-card system. For the 1960 census, the first electronic statistics machine was used: an IBM 101. Once again, the statistics produced on both occasions were discussed in reports; those reports, however, no longer feature French translations.

In the 1947 and 1960 census reports, the ‘place of birth’ table once more features the options ‘current municipality of residence’ and ‘other municipality in the same province’ (see Figure 3). Compared to the censuses of 1899–1930, the locality level here seems to be upscaled a little,
because the ‘different province in the Kingdom’ category has been split into eleven different categories for each of the provinces in the 1947 census, and into ‘adjacent province’ and ‘another province (not adjacent)’ in the 1960 census. This allows for more detail on migration at the province level. In the main text of the 1960 census report, it is explained that most of the people counted as migrated from ‘foreign countries’ are from ‘Indonesia and other parts of the Empire in the West’. This suggests that people from the Dutch colonies are now included in this category. Finally, the ‘unknown’ category is not mentioned in the 1960 census, possibly because the numbers are very low, and/or because the category is no longer deemed important. It is in the main text of the 1947 census that the word ‘autochthon’ pops up and in the explanation of the definitions of the 1960 census that the words ‘allochthon’ pops up for the first time to describe people born in the same municipality and people born elsewhere (different municipality, province or country).

The table featuring the nationalities of foreign inhabitants of the Netherlands in the census report of 1947 is ordered by continent and has the title: ‘Foreigners by nationality’ (see Figure 4). It starts with all the European countries in alphabetical order. Then, Africa is mentioned, out of which only the South-African Union is highlighted, with the other category being ‘other Africa’. This is followed by the North American countries, the United States and Canada (residual categories here are Central America and South America). The Asian countries featured are China, Turkey and Palestine (with the residual category ‘other Asia’). From Oceania, only Australia and New Zealand are mentioned, and no residual category is stated. The final two, residual categories are ‘unknown’ and ‘stateless’.

In the 1971 census, the table entitled ‘non-Dutch according to counting method, sex and nationality’ arranges countries differently (see Figure 5). First, some of the countries of the former
European Community (a predecessor of the European Union) are listed: Belgium, France, Great Britain, Italy, and West Germany. (The category of ‘other countries from the European Community’ consists of Denmark, Ireland and Luxembourg, as mentioned in footnote 2 of the table in Figure 5). Then, so-called ‘recruitment countries’ (‘wervingslanden’) are mentioned: Greece, Yugoslavia, Portugal, Spain, and Turkey. These countries were targeted by the Dutch government to attract workers in an attempt to fill large numbers of vacancies in specific sectors of the economy. This campaign lasted roughly from 1964 until 1970.47 The African countries mentioned here are Algeria and Tunisia (lumped together) and Morocco, again followed by a residual category: ‘other African countries’. ‘America’ features the United States, Canada and a category of ‘other American countries’. From Asia, China and Indonesia are highlighted, followed by ‘other Asian countries’. Finally, Oceania is not split into individual countries.

The order and structure of this list is rather different than it used to be, reflecting a society that is in motion. Turkey is no longer listed as an Asian country, but as a recruitment country. Interesting is the placement of Algeria, Tunisia and Morocco as African countries, since elsewhere in the same report they are listed as ‘recruitment countries’ as well.48 By the time of the census, Indonesia had gained independence, and therefore is no longer considered a colony here, nor does it count here as part of the Dutch territory. The Netherlands is in the process of decolonisation and labour migration and therefore hosts bigger numbers of people with more than one passport.49 The report therefore explicitly mentions that:


*) Not compiled from the actual census but from Schreven “Documentatierapport volkstelling 1971,” 154. Unknown whether this level of detail was reported in the main report.
People with more than one nationality, one of which is Dutch, are filed under the Netherlands. All other persons with more than one nationality are filed under their first-mentioned nationality. The ‘persons without nationality’ category also includes the people that registered as Maluku Islanders.

‘First mentioned’ in this context means whichever nationality (country of which people have a passport) is filled in first on the census forms. This is a significant detail, since it shows that Dutchness is defined by passport, rather than one’s place of birth. The ‘first-mentioned’ nationality still gives some agency to the people counted, except for when one of the nationalities is the Netherlands. In addition, it explicitly mentions that people with a double passport of which one is Dutch are counted as Dutch, something that is going to change in the next census.

The 1971 census report holds another detail that will become quite important later on: the first two columns with numbers in Figure 5 show the numbers of people (men and women) that were counted the old-fashioned way by CBS employees visiting people at home; the third and

![Figure 4. Nationalities as featured in the 1947 census. Source: CBS, 12e Volkstelling annex woningtelling, 49.](image_url)
fourth columns show the number of people that were counted ‘administratively’, using the municipal administration as source. Prior to the 1971 census, public outrage ensued over the fact that, increasingly, statistical tasks of CBS were performed by a computer, deemed a ‘people-sorting machine’. In addition, the government was planning the introduction of a civil service number (‘Burger Service Nummer’), with which all Dutch citizens would be registered in the Central Persons Administration (‘Centrale Persoonsadministratie’). This idea antagonised many among the Dutch, eliciting memories of World War II, when the German occupier was grateful for the meticulous Dutch administration (that, among other details, contained the addresses and religions of all citizens, including Jews). Resistance to automated administrative systems resulted in a situation whereby a (relatively small) number of people refused to cooperate with the census of 1971. CBS was able to replace their missing counts with counts from the municipal registration database quite accurately. In response to the rising protest among the general public and the ever-greater ease with which the (relatively) accurate municipal registers could be accessed, CBS decided to no longer hold a traditional census in 1981.

In addition to the public outrage concerning privacy, a societal debate emerged concerning the status of immigrants, which would, in time, be referred to as the ‘minorities debate’. A seminal publication that fueled the debate – that lasted until the end of the nineties – was *Allochthonen in Nederland* (‘Allochthons in the Netherlands’) by the sociologist Hilda Verweij-Jonker. In her redefinition of allochthons or ‘foreigners’ (‘buitenlanders’), she mentions three
reasons for her choice of groups included in the report. The discussed groups (1) have come to the Netherlands in large numbers and in a relatively short time, (2) have been subject to some form of government policy to facilitate their arrival in Dutch society, and (3) are clearly physically recognisable because of their skin colour and a language incomprehensible for Dutch people. The third reason makes explicit the racial (skin colour) and ethnic (language) connotations that have since stuck to the construct of allochthon and the particular groups discussed in the book: repatriates (Indonesians), Ambonese (Malukans), Surinamese, Antillians, labour
migrants (mostly Turkish and Moroccan), Chinese, refugees, students from the ‘Third World’ (mostly African and Asian countries). The discourse surrounding some of these groups would later dominate debates about integration and ‘the multicultural society’.

The Datafication of Race-ethnicity (1980–2018)

In the years 1981 and 1991, no census was organised, due to both the aforementioned public resistance and the increased technological ingenuity of the central administration. New, computerised systems, which could count and perform calculations very fast in comparison to the machines of 1971 and earlier, and the greater ease of storing information on magnetic tape made the compilation of statistics more efficient. CBS could now produce year-by-year statistics, which removed the necessity for a census every ten years.

A factor that also affected how demographic information was collected and disseminated was that, from 1985 onwards, people living in the Netherlands for at least five years could file for naturalisation. In addition, children born in the Netherlands gained the right to be registered as Dutch. In the so-called ‘minority debate’, these measures are seen as a way of empowering immigrants and granting them the necessary means for full integration. For statistics, however, this causes a problem, since naturalised immigrants cease to be visible as a group (because they ‘become’ Dutch). In the rather heated debate on the position of migrants in the eighties and nineties, there was a perceived political need for accurate statistics about migrants concerning employment, welfare status and crime. Since nationality was no longer the distinctive characteristic, CBS reverted to the birthplace of individuals’ parents. Due to the use of the municipal registration as the source of information, rather than self-reporting, classifications came to be completely in the hands of CBS from 1981 onwards (as it still is today).

CBS determines the origins of people by first looking at whether or not both their parents have Dutch nationality (see Figure 6). If that is the case, any offspring are considered to have Dutch origins, even if they (the children) are not born in the Netherlands. Offspring of two parents with Dutch origins are considered autochthones. The origins of allochthones are, in principle, derived from their country of birth. If this is the Netherlands, the country of birth of the mother is chosen as a person’s origin; however, if the mother was born in the Netherlands, then

![Decision Tree](image_url)

*Figure 6. Schematic overview of decision tree used for determining a Dutch citizen’s ‘ethnic’ origins.*
the country of birth of the father is used as informing the origin of their children.\textsuperscript{58} Only people with two parents who are born in the Netherlands are considered to be autochthones. This means that in the Netherlands, autochthonous parents can have allochthonous children, but allochthonous parents can only have autochthonous children if they have themselves both been born in the Netherlands. This means that it takes one generation more before the national origins of migrant children are considered to be Dutch. Although the grandchildren of people that migrated to the Netherlands are now officially autochthone, CBS has recently started to monitor the ‘third generation’ as well, since, according to CBS, a migration background is ‘a relevant factor in [one’s] socio-economic development’.\textsuperscript{59} By also producing statistics about third-generation allochthons, CBS is shaping the public discourse in such a way that it seems natural that allochthones never really become Dutch.

Within the category of allochthones, there is a subdivision between Western allochthones and non-Western allochthones (see Figure 7). People from Europe (excluding Turkey), North America, Oceania, Japan and Indonesia are considered Western allochthones. People from Turkey, Africa, Latin America, Asia (excluding Japan), Suriname, the Dutch Antilles and Aruba are considered non-Western allochthones. While in the previous periods there was no distinction between ‘foreigners’, now foreign nations are clustered according to their ‘Westernness’.

People from Suriname, Aruba and the Dutch Antilles (all considered non-Western) have been viewed as problem groups by policymakers, whereas people from Indonesia have supposedly assimilated into Dutch society without any major problems.\textsuperscript{60} This perceived difference has severely shaped the image, constructed within Dutch society, about different groups. The fact that the distinction between Western and non-Western countries is not geographically

<table>
<thead>
<tr>
<th>Clustering of Dutch citizens based on country of origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autochthone</td>
</tr>
<tr>
<td>Western</td>
</tr>
<tr>
<td>The Netherlands</td>
</tr>
<tr>
<td>Europe (but not Turkey)</td>
</tr>
<tr>
<td>Africa</td>
</tr>
<tr>
<td>North America</td>
</tr>
<tr>
<td>Oceania</td>
</tr>
<tr>
<td>Japan</td>
</tr>
<tr>
<td>Indonesia</td>
</tr>
</tbody>
</table>

determined suggests that other factors such as race and ethnicity play a role. Making this system the standard for the creation of all major social and economic statistics about Dutch citizens has resulted in the construction of a race-ethnic hierarchy.

In 2001, for the first time in thirty years, CBS held a register-based census to accommodate requests from the European Union. In the report for the 2001 census, an English-language document, it becomes particularly clear how problematic it is when Dutch citizens with a migration background are consistently referred to as ‘foreigners’. The document states:

In the Netherlands about eighteen percent of the population can be classified as ‘foreigners’. Foreigners are defined as persons of whom at least one parent is born abroad. A distinction is made between first and second generation foreigners. A first generation foreigner is a foreign born person with at least one parent born abroad, a second generation foreigner is a person born in the Netherlands with at least one parent born abroad. Foreigners can be further divided into western and non-western; western foreigners come from Europe (excluding Turkey), North America, Indonesia and Japan, and non-western foreigners come from Turkey, Africa, Latin America and the rest of Asia.

As can be seen in Figure 8, four groups in the ‘foreigners’ category are highlighted: Turks, Moroccans, Surinamese and Antilleans. These are the populations that have been considered ‘problematic’ since the seventies – and are therefore considered non-Western.

---

As noted by Kees Groenendijk:

The typically Dutch distinction between Western and Non-Western allochthones is evidently based on political criteria, namely welfare level, geographical or cultural proximity of the country of origin and assumptions about the problematic character of the group.63

‘The problematic character of the group’, in this context, is an implicit reference to the process of racialisation that has occurred with respect to several of the larger immigrant groups in the Netherlands, most notably the Turkish, Moroccan, Surinamese and Dutch Antilleans. Historians Michael Omi and Howard Winant explain the process of racialisation as ‘the extension of racial meaning to a previously racially unclassified relationship, social practice or group’.64 In this context, it explains how a personal characteristic, such as having particular non-Dutch origins, has become a socially, politically and culturally laden idea. In the 2011 census, the duration of stay is included in the table about immigrants (see Figure 9). Interesting here is that, in the accompanying text, CBS refers to the fact that:

a large majority of immigrants from non-EU/EFTA countries have quite a long migration history. Almost three-quarters stayed in the Netherlands for more than 10 years, and nearly half for longer than 20 years.65

The message here is that most of the people in this category have been ‘around’ for quite a while. By naming the table ‘duration of stay’, it is implied, still, that this group’s stay is not permanent.

<table>
<thead>
<tr>
<th>Country of birth</th>
<th>Total</th>
<th>less than 5 yrs</th>
<th>5-9 yrs</th>
<th>10-19 yrs</th>
<th>20 yrs or longer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x 1,000</td>
<td>%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>122.3</td>
<td>20.9</td>
<td>9.8</td>
<td>17.0</td>
<td>52.3</td>
</tr>
<tr>
<td>Poland</td>
<td>66.6</td>
<td>60.3</td>
<td>19.0</td>
<td>11.2</td>
<td>9.6</td>
</tr>
<tr>
<td>Belgium</td>
<td>50.0</td>
<td>17.5</td>
<td>13.0</td>
<td>18.7</td>
<td>50.8</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>47.2</td>
<td>25.3</td>
<td>14.7</td>
<td>21.6</td>
<td>38.4</td>
</tr>
<tr>
<td>France</td>
<td>23.4</td>
<td>33.1</td>
<td>14.1</td>
<td>19.5</td>
<td>33.3</td>
</tr>
<tr>
<td>Italy</td>
<td>20.8</td>
<td>30.2</td>
<td>12.0</td>
<td>15.8</td>
<td>42.0</td>
</tr>
<tr>
<td>Spain</td>
<td>20.4</td>
<td>30.1</td>
<td>12.0</td>
<td>14.0</td>
<td>44.0</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>15.0</td>
<td>82.0</td>
<td>8.1</td>
<td>6.8</td>
<td>3.2</td>
</tr>
<tr>
<td>Portugal</td>
<td>14.7</td>
<td>34.8</td>
<td>15.6</td>
<td>16.3</td>
<td>33.3</td>
</tr>
<tr>
<td>Romania</td>
<td>12.3</td>
<td>53.9</td>
<td>15.4</td>
<td>18.6</td>
<td>12.2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>468.1</strong></td>
<td><strong>33.6</strong></td>
<td><strong>13.2</strong></td>
<td><strong>16.8</strong></td>
<td><strong>36.3</strong></td>
</tr>
</tbody>
</table>

The 2001 and 2011 censuses were based on workforce data, but CBS is planning to use the database of DUO (the governmental education services, formerly known as ‘IB-groep’) for the 2021 census. This database covers over ten million people, which ensures that fewer estimates have to be made, or that less information has to be obtained from other sources. However, much like in previous censuses, this still involves an automated combining of existing data, requiring a state-determined classification system in which there is no agency for individuals to decide for themselves which identity they feel suits their experiences, a system which is currently in use in the United States.

Taking nationality as a self-evident epistemological basis for government policy and the social sciences has been dubbed a form of ‘methodological nationalism’. Researchers operating within this paradigm ‘assume that countries are the natural units for comparative studies, equate society with the nation-state, and conflate national interests with the purposes of social science’. One of the ways in which the flaws of this epistemology become clear is in the classification of people from areas in the world where borders are contested. The Dutch categorisation system classifies people from Palestine as Israeli and Kurdish Turks as Turkish. People are therefore possibly classified as belonging to the very nation whose prosecution they might have fled. Another pitfall of methodological nationalism is the aforementioned process by which, over time, nationalities become racialised through the assignment of specific connotations to nationality labels. Both Eastern European migrants and migrants from Morocco and Turkey have become known as troublemakers through public policy and discourse. The autochthone/allochthone dichotomy as an epistemology has shaped knowledge production by setting autochthony as the norm. By introducing these terms in its classifications in the 1950s, making them socially meaningful and relevant in the 1960s and 1970s, and setting it as the technological norm in 1981 with the first virtual census, CBS and Dutch policymakers have created their own, historically and geographically situated, norm of ‘autochthon’ against which all ‘Others’ are measured:

Like whiteness, ‘autochthony’ has implicitly (and sometimes explicitly) functioned as the unreflexive norm, a neutral category, a natural fact without a history or relational context. Thus it functions, like whiteness, as a ‘reference category’ against which deviant cultures can be measured, or a cultural ‘whole’ into which minoritized and racialized Others can be reasonably expected to ‘integrate’.

By using autochthony as a ‘reference category’, CBS implicitly makes a value judgement about each category in the system, which therefore becomes a hierarchy. Subsequently, through everyday use and in the popular discourse, the dichotomy of allochthon and autochthon has become:

a racial discourse carried on implicitly in a setting in which the use of the term ‘race’ may be verboten, but where ‘everyone’ knows, and understands, tacitly, the unspoken text.

In the case of racialised discourses imposed on ‘ethnic’ categories in the United States, Dvora Yanow has suggested calling the resulting categories race-ethnic. After several analyses together
with Dutch colleagues, she has come to the conclusion that the same could be said about categories used by the Dutch government in both medical and policy contexts. Critique of the racialisation of the labels used has also been formulated from a postcolonial perspective and in terms of inclusion and exclusion from moral citizenship.

As a result of the pejorative connotations attached to the labels allochthon and autochthon, in 2016, the Netherlands Scientific Council for Government Policy (Wetenschappelijke Raad voor het Regeringsbeleid or WRR) advised the government to abolish the ‘allochthon’ label and replace it with the supposedly more neutral ‘people with a migration background’. The Dutch government subsequently made this change, and as of 2018, official policy documents and statistics no longer contain the tainted words allochthon or autochthon. In this new system, people formerly called autochthon also have a migration background: the Netherlands. However, the categorisation system still functions in the same manner — that is, with forced categories, relying on the distinction between Western and non-Western, and with the parents’ place of birth as a basis.

Conclusion

Through the infrastructural inversion I have undertaken, it becomes clear that definitions concerning people’s origins are not only socially determined, but also heavily rely on the data ontologies through which they are institutionalised. From 1899 until 1940, categorisation was done by relatively time-consuming ways of counting, combined with a rather large amount of agency from the side of the people that are counted. Only a handful of mainly neighbouring countries seem to be of significant importance in the reports. People coming from and living in one of the colonies are still considered Dutch, and foreigners are not yet clustered in ‘Western’ and ‘non-Western’ groups. The period from 1945 to 1980 is characterised by a move from mechanised counting to digital counting. At the same time, there are significant social developments in terms of decolonisation and labour migration. At this point, former colonies become foreign countries, considered Western or non-Western depending on the social position of the relevant groups in Dutch society. To the so-called ‘recruitment countries’ belong those that will later constitute the category of the allochtons. In the case of a double passport, the Dutch passport overrules other passports, which shows that Dutchness prevails.

A national privacy discussion coupled with increasingly accurate bureaucratic systems caused CBS to start counting people ‘administratively’. The final era, from 1980 onwards, is characterised by a national discussion and scientific investigation of immigrant integration and the multicultural society, coupled with an increasing datafication of governmental information. The increasingly automated means of counting in this period was accompanied by the instatement of a state-imposed and automated nationality category. Clustering nationalities on the basis of who is ‘recognisable as a group’ turns an otherwise nationality-based categorisation system into a race-ethnically informed one. In addition, ‘administrative’ counting in a ‘virtual’ census means that people themselves are no longer counted, but rather their data. Now, other passports overrule the Dutch one, and ‘otherness’ prevails.
By discussing social developments as well as analysing the CBS infrastructure, I have shown how the conceptual pair of autochthon and allochthon, and, from 2018 onwards, the categories of people with a Western/non-Western/Dutch migration background, are not only social but also sociotechnical constructs. Now that the CBS website features an API, enabling any organisation, governmental or not, and well-intentioned or not, to create systems and applications that make use of race-ethnically structured data, the race-ethnically informed epistemology quite literally turns into an ontology, for it starts to have an impact on people depending on their position in the hierarchy.

In this article, I have focused primarily on people and data that are part of bureaucratic systems. Further historical research could explore inclusion and exclusion practices of the datasets themselves. Such research could investigate which people were left out and trace how this has changed over time. In addition, further research is needed to determine the risks for inequality and injustice in the Netherlands, as caused by race-ethnically structured data. A start has been made with such research in the Anglo-Saxon world; however, the case discussed here once again confirms that the very Dutch idea that ‘racism is an American problem’ needs an update. When a census report refers to people with Dutch passports as ‘foreigners’ because their parents were born in what is deemed a ‘non-Western’ country, one should be convinced that statistics are potentially highly political, rather than neutral, and that systems based on race-ethnically determined data are capable of creating a race-ethnically determined, datafied society.

Notes

1. From the Online Etymology Dictionary: https://www.etymonline.com/word/origin.
10. See www.leefbaarometer.nl.
11. See www.allochtonenmeter.nl.
22. Also interesting in light of this article is Chapter 6 in Bowker and Star, *Sorting Things Out*, on racial categorisation under apartheid in South Africa.
23. Censuses were conducted once every ten years, with the exception of 1941 (due to World War II), 1981 and 1991 (both due to the fact that it was no longer necessary; see the remainder of this article). The census was reinstalled in 2001 because the European Union expected counts from its member states. Since reports were produced every decade, they also provide insight into their relation to key social, technical and sociotechnical changes over time.
25. CBS currently uses the term ‘ethnic origins’; however, people are always counted as former citizens of nations, not as a member of an ethnic group. This becomes painfully clear in cases of territorial disputes, such as, for example, with Kurdish Turks and Palestinians, who are counted as Turks and Israelis, respectively (see also further on in the main text).
38. The French political influence officially ended in 1813 (see M.J. Van der Burg, “Nederland onder Franse invloed: Cultuurtransfer en staatsvorming in de Napoleontische tijd, 1799–1833” [Dissertation, University of Amsterdam, 2007]), but the cultural influence was noticeable well into the twentieth century.
40. The only exception with regard to the quantities is Poland. It is unclear why it is ranked number 15 in 1920 and 1930 with numbers of people (roughly 2700 in 1920 and 5900 in 1930) that exceed the quantities from all other countries, except Germany, Belgium and Austria.
41. The Dutch Antilles are still part of the Kingdom of the Netherlands as of 2018.
44. Kellenbach, “Mechanisering en automatisering.” 79.
50. CBS, *14e Algemene volkstelling annex woningtelling 28 Februari 1971; People from the Maluku Islands were awaiting the dissolution of Indonesian occupation, which is still yet to happen.
53. Ibid., 412.
60. The reason for this can be found in the different social and economic circumstances under which these groups settled in the Netherlands. People from Indonesia immigrated in the 1950s, at the time of a big economic boom and with plans to stay indefinitely, whereas people from Suriname immigrated at the beginning of the 1980s, in the middle of a recession. Frank Bovenkerk, “Rasdiscriminatie in Nederland,” in *Omdat zij anders zijn: Patronen van rasdiscriminatie in Nederland*, ed. Frank Bovenkerk (Meppel/Amsterdam: Boom Koninklijke Uitgevers, 1978), 13–14.


74. Yanow and Van der Haar, “People out of Place,” 250.


77. See Wekker, *White Innocence*.

78. Not to be confused with legal citizenship. See Schinkel, “From Zoëpolitics to Biopolitics.”

79. Bovens et al., “Migratie en classificatie”.


Biography

Gerwin van Schie obtained his MA in New Media and Digital Culture from Utrecht University in 2013 and worked as a researcher at the Utrecht Data School. He is now a PhD student in Gender Studies at Utrecht University in the NWO-financed project ‘The Datafication of Race and Ethnicity in the Netherlands’. This study investigates how the Dutch state classifies its inhabitants on the basis of their origin and how this information is datafied by CBS (Statistics Netherlands). It will also examine which systems are devised based on this data, how these systems are used by third parties, and how this may encourage data-driven discrimination.