

Clusteren op basis van cultuur

Etno-linguïstische indelingen van herkomstlanden als alternatieven voor westers/niet-westers in integratieonderzoek

Auteurs: Roel Jennissen, Mark Bovens en Godfried Engbersen

Working Paper nummer 48

Oktober 2021

Colofon

De serie 'Working Papers' omvat studies die in het kader van de werkzaamheden van de WRR tot stand zijn gekomen. De verantwoordelijkheid voor de inhoud en de ingenomen standpunten berust bij de auteurs. Een overzicht van alle Working papers is te vinden op www.wrr.nl.

Wetenschappelijke Raad voor het Regeringsbeleid

Buitenhof 34
Postbus 20004
2500 EA Den Haag
070-356 46 00
info@wrr.nl
wrr.nl

Working Paper nummer 48
ISBN: 978-90-83201-21-4

© Wetenschappelijke Raad voor het Regeringsbeleid,
Den Haag 2021

De inhoud van deze publicatie mag (gedeeltelijk) worden gebruikt en overgenomen voor niet-commerciële doeleinden. De inhoud mag daarbij niet veranderen. Citaten moeten altijd aangegeven zijn.

Inhoudsopgave

Ten geleide	4
Clusteren op basis van cultuur	5
Relevante eisen bij de classificatie van herkomstgroepen.....	6
Classificeren op basis van culturele achtergrond bij integratievraagstukken.....	8
Een etno-linguïstische indeling van herkomstlanden in twaalf clusters	11
Naar meer behapbare clusters.....	17
Conclusie.....	22
Literatuur.....	24
Bijlage 1. De indeling van herkomstlanden in geaggregeerde groepen op basis van culturele verwantschap (taal en religie)	28

Ten geleide

Het WRR Working Paper nr. 48 *Clusteren op basis van cultuur: etno-linguïstische indelingen van herkomstlanden als alternatieven voor westers/niet-westers in integratieonderzoek* is geschreven door Roel Jennissen, lid van de wetenschappelijke staf van de WRR, en de raadsleden Mark Bovens en Godfried Engbersen. In dit Working Paper dragen de auteurs een alternatieve indeling van herkomstlanden aan voor integratieonderzoek, in plaats van de standaardindeling in westers en niet-westers.

Deze deelstudie is uitgevoerd binnen het project ‘Migratiediversiteit’. De centrale vraag bij dit project is: hoe kan de overheid beter omgaan met de toenemende verscheidenheid naar herkomst in gemeenten en regio’s?

De WRR bedankt de stafleden Meike Bokhorst, Josta de Hoog en Will Tiemeijer voor hun waardevolle opmerkingen. Dit geldt ook voor Saskia Bonjour (Universiteit van Amsterdam), Mérove Gijsberts (SCP) en Helga de Valk (NIDI) die een eerdere versie van dit Working Paper van schriftelijk commentaar voorzagen.

De serie ‘Working Papers’ omvat studies die in het kader van de werkzaamheden van de WRR tot stand zijn gekomen. De verantwoordelijkheid voor de inhoud berust bij de auteurs.

Prof. mr. J.E.J. (Corien) Prins
Voorzitter WRR

Prof. dr. F.W.A. (Frans) Brom
Secretaris WRR

Clusteren op basis van cultuur

Nederland is uitgegroeid tot een migratiesamenleving die een grote aantrekkingskracht uitoefent op migranten uit alle delen van de wereld.¹ Een belangrijke vraag is hoe deze migranten en hun kinderen hun weg weten te vinden in de belangrijkste instituties van de Nederlandse samenleving en met welke problemen ze daarbij te maken hebben. Onderzoek hiernaar staat te boek als ‘integratieonderzoek’. Een voorbeeld daarvan is het *Jaarrapport Integratie* van het CBS, dat ontwikkelingen op het gebied van demografie, onderwijs, arbeid, inkomen en uitkeringen, criminaliteit, gezondheid, en sociale en maatschappelijke participatie in kaart brengt.²

In ons land wonen tegenwoordig migranten uit meer dan 200 herkomstlanden.³ Het is ondoenlijk om in het integratieonderzoek al die herkomstlanden afzonderlijk op te nemen in de analyses. Daarom worden groepen van herkomstlanden vaak samengevoegd tot clusters op basis van gedeelde kenmerken die relevant kunnen zijn voor het verklaren van participatie en integratie.

Een veel gebruikte clustering in het integratieonderzoek is die in westerse tegenover niet-westerse herkomstlanden. Deze tweedeling is in 1999 door het CBS geïntroduceerd. De gedachte hierachter was dat er verschillen zijn in de sociaal-economische en culturele positie tussen migranten uit westerse en niet-westerse landen die relevant zijn voor hun mate van integratie. In een eerdere publicatie hebben we laten zien dat deze tweedeling wetenschappelijk niet is onderbouwd. Het is niet duidelijk op welke theoretische gronden het ene land als westers geldt en het andere land als niet-westers. Vanwege de grote variëteit aan herkomstlanden is deze standaardtweedeling ook niet erg informatief meer. De verschillen binnen de twee clusters zijn soms net zo groot als de verschillen tussen de twee clusters.

¹ WRR 2021.

² Zie bijvoorbeeld De Mooij et al. 2020.

³ Volgens de Verenigde Naties (VN) waren er in 2019 slechts 193 plus 2 (Vaticaanstad en de Palestijnse Autoriteit) landen op de wereld. Het herkomstland van individuen is echter gebaseerd op het geboorteland of het geboorteland van een ouder van een persoon. Daarom kunnen in Nederland verblijvende personen herkomstlanden hebben die niet meer bestaan zoals Joegoslavië, Tsjechoslowakije, de Sovjet-Unie en zelfs Oostenrijk-Hongarije. Dit verklaart waarom er meer herkomstlanden zijn.

Daarom biedt deze simpele tweedeling geen goed zicht meer op integratie. Daarnaast roept de labeling westers versus niet-westers ongewenste associaties op.⁴

Hoe krijgen we dan wel een goed zicht op integratie? In dit paper presenteren wij alternatieven voor de tweedeling westers/niet-westers die gebruikt kunnen worden in het onderzoek en het beleid op het gebied van integratie en samenlevingsvraagstukken. Onze belangrijkste uitgangspunten zijn:

1. Een clustering van herkomstlanden op basis van welvaart is niet nodig omdat er veel relevante economische achtergrondvariabelen beschikbaar zijn.
2. Een clustering op basis van culturele afstand tot de Nederlandse samenleving kan wel relevant zijn voor integratieonderzoek en kan niet goed worden vervangen door andere variabelen.
3. Daarbij dient de clustering meervoudig te zijn en moeten tweedelingen zoveel mogelijk worden vermeden.

In dit paper presenteren we een meervoudige indeling van herkomstlanden die is gebaseerd op de culturele afstand tot de Nederlandse samenleving. De achterliggende gedachte is dat culturele factoren – naast sociaal-economische factoren – relevant kunnen zijn bij vraagstukken rond de integratie van migranten en hun kinderen. Deze indeling bouwt voort op een clustering die is ontwikkeld in de WRR-verkenning *De nieuwe verscheidenheid*. De nieuwe indeling onderscheidt zes, en soms zeven, verschillende clusters van groepen. Daarmee wordt zowel recht gedaan aan de complexe werkelijkheid als aan de eisen dat een indeling behapbaar moet zijn en de tand destijds moet kunnen doorstaan. Zijn tweedelingen dan helemaal uit den boze? Wij zullen beargumenteren dat het in heel specifieke gevallen, bijvoorbeeld bij de samenstelling van een relevante ‘referentiecategorie’, nuttig kan zijn om een tweedeling te maken. In dit paper doen we daartoe een voorstel.

Voordat we deze culturele indeling uitwerken, zetten we eerst nog eens op een rij wat de relevante criteria zijn bij clustering en labeling.

Relevante eisen bij de classificatie van herkomstgroepen

In eerdere publicaties hebben we uiteengezet dat aan het classificeren naar herkomst verschillende eisen moeten worden gesteld.⁵ Classificatie bestaat uit verschillende componenten, waaronder het clusteren en het labelen. Clustering of categorisering is de

⁴ Zie uitgebreid Bovens et al. 2021.

⁵ Bovens et al. 2016 en 2021.

indeling van gegevens in groepen om zo de complexiteit te reduceren. Labelen is het toewijzen van termen aan clusters. In tekstkader 1 geven we aan welke eisen van belang zijn voor het ontwikkelen van classificaties naar herkomst.⁶

Tekstkader 1: Hoe classificeren naar herkomst?

Bij de beoordeling van de geschiktheid van clusters en termen daarvoor zijn twee soorten eisen van belang:

1. *Informatieve eisen*: clusters en termen moeten
 - a. voldoende empirische verfijning bieden;
 - b. continuïteit en validiteit van onderliggende indelingen waarborgen;
 - c. zo min mogelijk begripsmatige verwarring veroorzaken.
2. *Performatieve eisen*: de termen moeten
 - a. zo min mogelijk uitsluitende werking hebben;
 - b. zo min mogelijk negatieve associaties oproepen.
 - c. zoveel mogelijk nevenschikken en niet onderschikken.

Daarnaast hebben we een aantal criteria opgesteld om na te gaan of het überhaupt verantwoord is om in beleid en onderzoek onderscheid naar herkomst te maken.⁷ Deze staan in tekstkader 2.

Tekstkader 2: Wanneer classificeren naar herkomst?

1. Dient classificatie een legitiem doel (*legitimiteit*)?
2. Draagt classificatie bij aan het bereiken van het doel (*functionaliteit*)?
3. Kan het doel niet op een andere manier worden bereikt (*subsidiariteit*)?
4. Wegen de voordelen op tegen de nadelen (*proportionaliteit*)?

Daarbij is vooral het derde criterium, *subsidiariteit*, hier van belang. Het is in veel gevallen niet nodig om herkomstlanden te clusteren en te labelen, omdat de onderzoeksvraag in analyses ook te beantwoorden is door de relevante factoren mee te nemen als continue variabelen.

⁶ Bovens et al. 2016: 60.

⁷ Bovens et al. 2016: 59.

Dat speelt in het bijzonder bij sociaal-economische factoren. Als onderzoekers of beleidsmakers onderscheid willen maken naar het welvaartsniveau van de landen waar migranten vandaan komen, dan kunnen ze gewoon het bruto binnenlands product (bbp) per hoofd van de verschillende landen meenemen als een continue variabele in de analyse. Het subsidiariteitscriterium komt erop neer dat het niet nodig is om landen in te delen op basis van kenmerken die uit te drukken zijn in cijfers, zoals het gemiddelde aantal jaren scholing en het bbp per hoofd van de bevolking. Ons advies is in dat geval: neem gewoon de betreffende variabele zelf mee in de analyse en noem het land niet. Het is dan bijvoorbeeld niet nodig om een indeling te maken in lagere, midden- en hogere welvaartslanden, of een vierdeling in zeer ontwikkelde landen, hoog ontwikkelde landen, matig ontwikkelde landen en laag ontwikkelde landen, zoals in de Human Development Index van de VN.⁸ Dit vermindert ook de kans op onwenselijke associaties. In geval van de Human Development Index kan de kwalificatie van een herkomstland als 'laag ontwikkeld' ertoe leiden dat individuele migranten uit de betreffende landen als laag ontwikkeld worden gezien.

Classificeren op basis van culturele achtergrond bij integratievraagstukken

Voor het integratieonderzoek zijn twee dimensies van betekenis: enerzijds een sociaal-economische of structurele dimensie en anderzijds een sociaal-culturele dimensie.⁹ Het sociaal-economische deel van het integratievraagstuk is, zoals we hebben beargumenteerd, vaak goed in cijfers uit te drukken. Het Stelsel van Sociaal-statistische Bestanden (SSB) van het CBS bevat inmiddels een uitgebreide set aan sociaal-economische achtergrondvariabelen, die onderzoekers kunnen hanteren.

Maar bij een reeks van integratie- en samenlevingsvraagstukken, zoals schoolsucces, sociale contacten, deelname aan vrijwilligerswerk, religiositeit, huwelijkspatronen of deviant gedrag, zijn sociaal-economische variabelen niet voldoende om verschillen tussen groepen migranten, of tussen migranten en gevestigde groepen, te verklaren. We moeten dan op zoek naar achtergrondvariabelen die niet goed in registerdata vallen uit te drukken, maar wel van invloed kunnen zijn op het gedrag van personen. Dit kan zowel het gedrag zijn van de betrokken persoon die zelf tot een bepaalde herkomstgroep behoort, als de manier waarop die persoon gepercipieerd en gepositioneerd is in de Nederlandse samenleving. Het tweede

⁸ Zie bijvoorbeeld UNDP 2019.

⁹ Dagevos 2001.

kan van invloed zijn op het gedrag van andere personen, bijvoorbeeld in het geval van discriminatie. Velen komen dan uit bij het begrip ‘cultuur’.¹⁰

Cultuur is echter een moeilijk te definiëren fenomeen, zelfs voor de culturele antropologie, een vakgebied dat haar bestaansrecht ontleent aan de culturele factor. De Amerikaanse antropologen Alfred Kroeber en Clyde Kluckhohn¹¹ inventariseerden aan het begin van de jaren vijftig van de vorige eeuw meer dan 160 verschillende definities van cultuur. Ook in eenentwintigste-eeuwse publicaties zoals *Culture and public action* van Vijayendra Rao en Michael Walton, met bijdragen van Arjun Appadurai, Mary Douglas en Amartya Sen, zien we dat cultuur op zeer uiteenlopende manieren wordt gedefinieerd.¹² Daarnaast wijzen auteurs op mogelijke gevaren van culturele analyses, waaronder een essentialisering van cultuur – cultuur is immers aan verandering onderhevig¹³ – en aan een mogelijk stigmatiserende werking ervan. Tegelijkertijd onderkennen de auteurs dat de culturele factor van invloed is op het gedrag van mensen en de ontwikkeling van samenlevingen.

Ondanks het feit dat cultuur lastig is te definiëren, zijn er aspecten van cultuur die objectief zijn vast te stellen, zoals taal en religie. Onderzoekers die landen proberen te clusteren op basis van cultuur gebruiken deze aspecten dan ook vaak, zij het soms impliciet.¹⁴ De indeling in herkomstlanden die wij in dit paper presenteren, is eveneens gebaseerd op taal en religie.

De culturele clustering van herkomstlanden die we verderop uitwerken, is in de eerste plaats gebaseerd op taalverwantschap. Sommige taalfamilies zijn meer verwant aan de Noord-Europese taalfamilie waarvan de Nederlandse taal een onderdeel is. Onderzoek laat zien dat migranten gemakkelijker een vreemde taal leren als hun moedertaal verwant is aan die taal.¹⁵ Migranten in Nederland die een verwante taal spreken, zoals Deens, Duits of Engels, leren over het algemeen dan ook sneller Nederlands dan migranten die een taal spreken die daar niet aan verwant is, zoals Arabisch, Mandarijn of Japans.¹⁶ Taalbeheersing – je verstaanbaar

¹⁰ Een bekend voorbeeld is de culturele clustering op basis van de World Value Survey. Zie: www.worldvaluessurvey.org/WVSContents.jsp, en ook Inglehart en Welzel 2010. Een ander bekend voorbeeld is de culturele clustering van Hofstede 1980.

¹¹ Kroeber en Kluckhohn 1952.

¹² Rao en Walton 2004.

¹³ Burke (2009: 102) formuleert het kernachtig: “In our world, no culture is an island”.

¹⁴ Zie bijvoorbeeld Cattell 1950; Bakacsi et al. 2002; Inglehart en Welzel 2010.

¹⁵ Zie bijvoorbeeld Van Tubergen en Kalmijn 2005; Isphording en Otten 2014.

¹⁶ Zie bijvoorbeeld Schepens et al. 2016.

kunnen maken – is een belangrijk onderdeel van participatie en integratie. Zij is cruciaal voor veel aspecten van integratie, zoals werk¹⁷, sociale contacten¹⁸ en politieke participatie¹⁹.

De indeling naar taalgroepen is wetenschappelijk uitgebreid onderbouwd. Zij biedt de nodige verfijning, omdat in de literatuur verschillende taalgroepen, op verschillende niveaus van aggregatie worden onderscheiden. Ook is zij door de tijd heen redelijk stabiel, omdat de *lingua franca* in een herkomstland niet van de ene dag op de andere zal veranderen. Daarbij is het mogelijk om taalsubgroepen te onderscheiden als dat relevant is voor de specifieke migratie naar ons land. Dat maakt het bijvoorbeeld ook mogelijk om onderscheid te maken tussen de taal van de migranten in ons land en de standaardtaal in het herkomstland. Taalgroepen zijn ook weinig waardengeladen labels, waardoor de negatieve performatieve effecten ervan redelijk beperkt zullen zijn.

In de tweede plaats is de indeling gebaseerd op religie. Religie is namelijk een belangrijke bron van normen en waarden en van culturele symbolen.²⁰ Ook een seculiere samenleving als de Nederlandse is doordrenkt van christelijke symbolen – in taal en cultuur, in omgangsvormen, feestdagen, rites en ethische beginselen. Bovendien valt religie redelijk goed objectief vast te stellen. Onderzoekers die landen proberen te clusteren op basis van cultuur, gebruiken dit aspect (impliciet) dan ook vaak.²¹

In het vervolg van dit paper presenteren wij een classificatie op basis van herkomstlanden die we oorspronkelijk hebben ontwikkeld voor de WRR-verkenning *De nieuwe verscheidenheid*.²² Deze classificatie, die uit twaalf geaggregeerde groepen bestaat, dikken we vervolgens verder in tot een aantal clusters dat zich beter leent om gegevens naar herkomstgroep te presenteren, bijvoorbeeld in tabellen op CBS-Statline. We zijn ons ervan bewust dat elke

¹⁷ Zie bijvoorbeeld Dustmann en Fabbri 2003; Auer 2017; Adserà en Ferrer 2021.

¹⁸ Zie bijvoorbeeld Thuesen 2016.

¹⁹ Zie bijvoorbeeld Wong et al. 2011; Ortensi en Riniolo 2020.

²⁰ Zie bijvoorbeeld Berger en Luckmann 1966; Bellah 1967.

²¹ Zie bijvoorbeeld Cattell 1950; Bakacsi et al. 2002.

²² Jennissen et al. 2018. Door voortschrijdend inzicht wijkt de hier gepresenteerde classificatie iets af van die in de WRR-verkenning. Deze indeling is voor een groot deel gebaseerd op de door Blom en Jennissen (2014) gehanteerde indeling, die is gebruikt in een onderzoek naar de betrokkenheid van verschillende etnische groepen bij verschillende typen van criminaliteit. Blom en Jennissen splitsen de groep met een westerse migratieachtergrond echter niet verder op dan in de categorieën ‘westers’ en ‘Oost-Europees’. We willen deze groep in deze studie opdelen in meer dan twee categorieën. De landenindeling van het GLOBE-project (zie bijvoorbeeld House et al. 2004) hebben we hierbij als inspiratiebron gebruikt.

indeling een reductie van de complexe werkelijkheid is en dat ook andere indelingen van landen mogelijk zijn.

Kanttekeningen bij onze culturele clustering

Voor we deze indeling nader uitwerken, is het goed om uit te leggen hoe we daarbij te werk zijn gegaan. Een eerste kanttekening is dat we bij de aggregatie uitgaan van de kenmerken van de in Nederland aanwezige migrantenpopulatie en niet van de overheersende taal en religie in het herkomstland zelf. Deze twee komen meestal overeen, maar dat is niet altijd het geval. Het belangrijkste voorbeeld hiervan in de Nederlandse context is de bevolking met als herkomstland Indonesië/voormalig Nederlands-Indië. Deze heeft in Nederland heel andere kenmerken dan in het huidige Indonesië.

In onze clustering is de aggregatie van herkomstlanden die dicht bij Nederland liggen gemiddeld genomen verfijnder dan bij verder gelegen landen. Daarbij speelt een praktische overweging een rol, namelijk de omvang van de herkomstgroepen. Op dit moment komt een zeer groot deel van de migranten in Nederland uit Europa, het Middellandse zeegebied of het Midden-Oosten. Een risico bij het creëren van een groot aantal of zeer specifieke clusters van verder weg gelegen herkomstlanden is dat de spoeling van de cijfers dunner wordt. In bepaalde clusters kunnen op een gegeven moment zo weinig personen zitten dat de cijfers niet meer gepresenteerd mogen worden, omdat anders de identiteit van individuen kan worden achterhaald (onthullingsrisico's).

De indelingen die we voorstellen, zijn niet in beton gegoten. Migratie heeft een dynamisch karakter en de kenmerken van de groepen die naar ons land komen, kunnen door de tijd heen veranderen. Het beste voorbeeld is wederom de migratie uit Indonesië. In de toekomst zal het merendeel van de migranten uit Indonesië vermoedelijk van huis uit geen Nederlands meer spreken en voor een groot deel islamitisch zijn. Dat kan reden zijn om Indonesië in de toekomst onder te brengen bij Zuidoost-Azië.

Een etno-linguïstische indeling van herkomstlanden in twaalf clusters

Zoals gezegd, verblijven in Nederland migranten uit meer dan 200 verschillende herkomstlanden. Op basis van taal en religie hebben we deze geaggregeerd tot twaalf verschillende grotere groepen: de Angelsaksische landen, de Continentaal-Germaanstalige landen, Mediterraan Europa, Midden- en Oost-Europa, de Arabische landen, Sub-Sahara-Afrika, Latijns-Amerika, de Cariben, Zuid-Azië, Centraal-Azië, Oost-Azië, en Zuidoost-Azië en de Pacific. Deze clustering komt grotendeels overeen met die van Vipin Gupta en zijn

coauteurs.²³ Wel kent onze indeling meer verfijning in Azië, een cultureel zeer heterogeen continent, waar ongeveer 60 procent van de wereldbevolking woont en waar het herkomstland ligt van ongeveer een derde van de inwoners van Nederland met een migratieachtergrond. Verder hebben wij de Cariben als een aparte groep onderscheiden. De precieze indeling van alle individuele herkomstlanden is te vinden in bijlage 1. Hieronder lichten we de indeling nader toe.

Angelsaksische landen

De voertaal van de meeste inwoners van deze landen is Engels. Ook Zuid-Afrika wordt tot deze landen gerekend.²⁴ Deze cluster is de enige waarvan de landen niet enigszins bij elkaar in de buurt liggen. Dit komt door de sterke invloed die het *British empire*, ‘waar de zon nooit onderging’, op deze landen heeft gehad.²⁵

Continental-Germaanstalige landen

Tot deze groep behoren de Europese landen die het Duits, Nederlands of één van de Noord-Germaanse talen als voertaal hebben. Indonesië maakt ook deel uit van deze groep landen omdat de meeste inwoners van Nederland met wortels in Indonesië/voormalig Nederlands-Indië (kinderen van) *totoks* of *indo's* zijn.²⁶ Ook België, Zwitserland en Luxemburg rekenen

²³ Gupta et al. 2002.

²⁴ Gedurende het apartheidsregime in Zuid-Afrika vond er al een uittocht plaats van hoogopgeleide Engels- en Afrikaanstalige Zuid-Afrikanen naar West-Europa (voornamelijk het Verenigd Koninkrijk), Noord-Amerika, Australië en Nieuw-Zeeland. In de jaren na de omverwerping van dit regime nam deze zogenoemde ‘*white brain drain*’ sterk toe (Kalule-Sabiti et al. 2003). De Zuid-Afrikaanse gemeenschap in Nederland bestaat voor het grootste deel uit Zuid-Afrikanen die cultureel het beste aansluiten bij de in Nederland aanwezige Angelsaksische gemeenschap.

²⁵ Ashkanasy et al. 2002.

²⁶ Tijdens de Japanse bezetting van Nederlands-Indië verbleven ongeveer 300.000 Europeanen in de archipel. Deze groep bestond uit circa 100.000 *totoks* (Nederlandse immigranten en hun nakomelingen) en circa 175.000 Indo-Europeanen, ook wel *indo's* genoemd. De laatstgenoemde groep bestaat uit nakomelingen uit relaties tussen (Indo-)Europeanen en inlanders die door de (Indo-)Europese ouder wettelijk waren erkend. Daarnaast woonden er 20.000 à 30.000 niet-Nederlandse Europeanen in Nederlands-Indië (Oostindie 2010). In de periode van 1946 (de *bersiap*-periode) tot en met 1962 (de Nieuw-Guineacrisis) kwamen er veel repatrianten naar Nederland. Omdat dit voornamelijk totoks en indo's waren, rekenen we mensen met een Indonesische/voormalig Nederlands-Indische achtergrond tot de individuen met een Continental-Germaanstalige achtergrond. Dit geldt ook voor de niet-Europese migranten die in de decennia na de Tweede Wereldoorlog vanuit (voormalig) Nederlands-Indië naar Nederland kwamen, zoals Zuid(oost)-Molukkers, Peranakan-Chinezen en Papoea's. Op basis van de gegevens uit het

we tot dit cluster ondanks dat een groot deel van de bevolking van deze landen een niet-Germaanse taal als moedertaal heeft. Dat geldt in het bijzonder voor Finland, dat ook tot deze groep behoort ondanks dat het merendeel van de bevolking een Finoegrise taal spreekt. Zweeds is echter, naast het Fins en Samisch, een officiële landstaal in Finland. De meeste Finnen beschikken over enige kennis van de Zweedse taal, omdat het voor alle scholieren in Finland verplicht is om Zweeds te leren.²⁷

Midden- en Oost-Europa

In veel Midden- en Oost-Europese landen spreekt het grootste deel van de bevolking een (Balto-)Slavische taal. Dit geldt niet voor Albanië, Armenië, Estland, Georgië, Hongarije, Kosovo, Moldavië en Roemenië. Al deze landen hebben echter in meer of mindere mate een Slavisch sprekende minderheidsgroep binnen hun landsgrenzen.²⁸ Daarnaast heeft ook de niet-Slavische bevolking vaak wel enige kennis van een Slavische taal, bijvoorbeeld Russisch of Servisch (in Kosovo).

Alle personen die voor 1991 geboren zijn in de voormalige Sovjet-Unie – dus ook in de voormalige Aziatische Sovjetrepublieken – worden ook gezien als een persoon met een Oost-Europese achtergrond. Dit is onvermijdelijk omdat we van het overgrote deel van deze mensen de republiek waar ze geboren zijn, niet meer kunnen achterhalen in de bevolkingsstatistiek. Deze problematiek is waarschijnlijk niet erg groot, aangezien we kunnen aannemen dat het grootste deel van de in Nederland verblijvende mensen met de voormalige Sovjet-Unie als herkomstland, behoort tot de zogenoemde Oost-Slaven (Russen, Oekraïners en Wit-Russen)²⁹ of tot de volkeren van de Baltische staten. In Nederland verblijvende mensen met een migratieachtergrond die na 1991 zijn geboren in een voormalige Aziatische

bevolkingsregister kunnen we deze groepen niet onderscheiden van de (Indo-)Europeanen met een Indonesische/voormalig Nederlands-Indische achtergrond.

²⁷ Latomaa en Nuolijärvi 2002.

²⁸ Hongarije is het enige Midden- en Oost-Europese land waarvoor dit nauwelijks van toepassing is. En ook in Albanië en Roemenië is het aandeel van de bevolking dat van huis uit Slavisch spreekt erg beperkt.

²⁹ Zelfs indien deze migranten geboren zijn in een niet-Slavische Sovjetrepubliek, is de kans vrij groot dat zij etnisch Slavisch zijn. Ingegeven door de ontwikkeling van nieuwe landbouwgebieden in Centraal-Azië en een politiek van sovjetisering, die gepaard ging met russificatie, werden in de periode vanaf de jaren dertig tot en met de jaren zeventig veel Slaven aangezet om naar de niet-Slavische delen van de Sovjet-Unie te verhuizen (Öberg en Boubnova 1993). Na het uitvallen van de Sovjet-Unie keerden veel Slaven terug naar (Wit-)Rusland en Oekraïne of kozen zij voor emigratie naar landen buiten de voormalige Sovjet-Unie (Jennissen et al. 2006).

Sovjetrepubliek³⁰, behoren niet tot de groep Midden- en Oost-Europeanen. Hierbij hebben we weer een uitzondering gemaakt voor Armenië en Georgië. Deze overwegend christelijke landen in Transkaukasië worden, in tegenstelling tot het overwegend islamitische Turkstalige buurland Azerbeidzjan, ook tot de Midden- en Oost-Europese landen gerekend.³¹

Mediterraan Europa

In deze landen spreekt de bevolking overwegend een Romaanse taal of Grieks. Dit geldt overigens niet voor Malta³² en Israël, waar de bevolking voor het grootste deel Semitische talen spreekt. Het laatstgenoemde land heeft sowieso een wat bijzondere positie binnen deze groep.³³

Latijns-Amerika

Onder de Latijns-Amerikaanse landen verstaan we hier alle landen in de Amerika's waar Spaans, Portugees of Frans één van de officiële talen is, met uitzondering van Canada en Puerto Rico. Door het koloniale verleden en de grootschalige immigratie vanuit Zuid-Europa vinden we in deze landen culturele kenmerken uit mediterrane landen.

³⁰ We hanteren hier de waterscheiding van de Oeral en de Kaukasus als grens tussen Europa en Azië en bekijken vervolgens in welk continent het merendeel van de bevolking van de betreffende staat woont. Zo rekenen we Russen wel tot de Midden- en Oost-Europeanen, maar Kazachen niet.

³¹ De in Nederland verblijvende Bulgaren zijn ook een groep waarover een discussie mogelijk is of deze tot de Midden- en Oost-Europese landen zou moeten behoren. Bulgarije is namelijk een multi-etnisch land waarin naast Bulgaren ook veel minderheden wonen. De grootste minderheidsgroep zijn de Turken, die volgens de officiële bevolkingsstatistieken ongeveer 10 procent van de totale bevolking van Bulgarije uitmaken. Gijsberts en Lubbers (2013) interviewden voor hun onderzoek circa 400 in Nederland geregistreerde Bulgaren. Hiervan was volgens de onderzoekers ongeveer de helft van Turks-Bulgaarse afkomst. Onduidelijk hierbij is of zij de Roma uit Bulgarije tot de etnische Bulgaren of etnische Turken rekenden. Ongeveer de helft van de Bulgaarse Roma behoort tot de Horahane (Crowe 2008). Dit zijn Turkssprekende Roma. Aangezien het overgrote deel van de Turkse Bulgaren ook Bulgaars (een Zuid-Slavische taal) spreekt, hebben we er toch voor gekozen om hen in te delen bij de groep Midden- en Oost-Europese landen.

³² Een groot deel van de Maltese bevolking spreekt echter – naast het Maltees, wat een Semitische taal is – ook Engels, de tweede officiële taal, en/of Italiaans (zie bijvoorbeeld Caruana 2013).

³³ Zo ligt Israël überhaupt niet in Europa en heeft het land allerlei eigenschappen die behoorlijk afwijken van die van de overige landen (taal, religie, sterke transnationale banden met de diaspora). Dat Israël is ingedeeld bij de mediterrane landen, is dan ook vooral omdat een beter alternatief ontbreekt.

Cariben

De landen in het Caribisch gebied waar Engels of Nederlands een officiële taal is, vormen samen de groep Cariben. Ook Belize, Guyana en Suriname behoren tot deze cluster, ofschoon ze in feite op het Amerikaanse vasteland liggen. Binnen deze groep hebben personen met een Surinaamse³⁴ of Antilliaanse achtergrond een speciale positie, aangezien zij een relatief grote kennis van de Nederlandse taal en cultuur hebben.

Arabische Landen

Dit zijn de landen waarin het merendeel van de bevolking Arabisch als moedertaal heeft. Deze landen zijn gesitueerd in Noord-Afrika en in het zuidwesten van Azië. De islam is in deze landen de godsdienst met verreweg de meeste aanhangers. Tsjaad en Djibouti maken geen deel uit van deze groep landen, ook al is Arabisch, naast Frans, de officiële taal van deze twee landen. We hebben hiervoor gekozen omdat meer dan de helft van de inwoners (en waarschijnlijk ook van de in Nederland woonachtige mensen met een migratieachtergrond die afkomstig zijn uit deze landen) tot etnisch niet-Arabische bevolkingsgroepen behoort. Ondanks dat veel in Nederland verblijvende personen met een Marokkaanse achtergrond tot de niet-Arabische Berber-bevolking behoren, wordt ook Marokko tot deze groep gerekend. De Berber-bevolking in Nederland spreekt overigens veelal Tamazight/Tarifit, wat, net zoals Arabisch, onderdeel is van de Afro-Aziatische taalfamilie.³⁵

Centraal-Azië

Dit zijn de landen in het geografische hart van Azië waar een Turkse of Iraanse taal wordt gesproken. De islam is de meest voorkomende godsdienst in deze landen. Pakistan zou ook tot deze landen gerekend kunnen worden. Echter, hoewel er ook in Pakistan Iraanse taalvarianten (bijvoorbeeld Pasjtje of Beloetsji) worden gesproken, spreekt de meerderheid van de inwoners van dit land een Indo-Arische taal (bijvoorbeeld Punjabi, Sindhi of Urdu).³⁶

³⁴ Hierbij zijn we ons ervan bewust dat er binnen de categorie personen met een Surinaamse achtergrond een grote diversiteit bestaat. De Surinaamse gemeenschap in Nederland bestaat voor ongeveer 50 procent uit Creolen en Marrons (Choenni en Harmsen 2007). De overige Surinamers zijn Hindostanen (34 procent van het totaal), Javanen (13 procent) en een aantal kleinere groepen (Chinezen, Boeroes, Libanezen, Cariben en Arowakken) die samen slechts 3 procent van de in Nederland verblijvende Surinamers uitmaken.

³⁵ Zie bijvoorbeeld Achab 2012.

³⁶ Zie bijvoorbeeld Manan en David 2014.

Zuid-Azië

In deze landen leven vele verschillende etno-linguïstische en religieuze groepen. De meeste mensen spreken er een Indo-Arische taal.³⁷ Dit zijn de landen die voortgekomen zijn uit het voormalige Britse rijk op het Indisch subcontinent. Ook personen met als herkomstland Nepal en Bhutan worden in deze studie als personen met een Zuid-Aziatische achtergrond gezien. Door het koloniale verleden zijn ook Britse gewoonten en tradities in de cultuur doorgesijpeld.

Oost-Azië

De meeste inwoners van de Oost-Aziatische landen spreken een Sino-Tibetaanse of een (Macro-)Altaïsche³⁸ taal. Confucianistische denkbeelden hebben de samenleving in deze landen sterk beïnvloed. Het neoconfucianisme, waarbij klassiek confucianisme is verweven met boeddhistische en taoïstische aspecten, streeft naar een stabiele samenleving waarin respect voor elkaar hoog in het vaandel staat.³⁹

Zuidoost-Azië en de Pacific

Dit zijn landen in tropisch Azië die niet op het Indisch subcontinent liggen, met uitzondering van Indonesië, en de landen in Oceanië, met uitzondering van Australië en Nieuw-Zeeland. In deze landen spreken de meeste mensen een taal die behoort tot de Austroaziatische of Austronesische taalfamilie of tot de Tai-Kadaitalen^{40, 41}. Zoals reeds vermeld, maakt Indonesië geen deel uit van deze groep landen.

Sub-Sahara-Afrika

Dit zijn alle Afrikaanse landen met uitzondering van de landen die behoren tot de Arabische landen en Zuid-Afrika. Op de landen in de Hoorn van Afrika⁴² en

³⁷ De grote uitzondering hierop is het zuiden van India waar een Dravidische taal wordt gesproken (Venkateswarlu 2020).

³⁸ Zie bijvoorbeeld Miller 1971.

³⁹ Zie bijvoorbeeld Yuan 2007; Chen 2017.

⁴⁰ In de praktijk gaat het in dit cluster alleen om de Zuidwestelijke Thai-talen die in Thailand en Laos worden gesproken. In de overige landen waar Tai-Kadaitalen oorspronkelijk worden gebezigd, is een andere taal die niet tot deze taalfamilie behoort dominant (zie bijvoorbeeld Sindwell en Jenny 2021).

⁴¹ Zie bijvoorbeeld Enfield en Comrie 2015.

⁴² De meeste inwoners van de landen in de Hoorn van Afrika spreken een Koesjitische taal. Koesjitische talen behoren tot de Afro-Aziatische taalgroep, die ook wel de Hamito-Semitische taalgroep wordt genoemd (zie bijvoorbeeld Porkhomovshy 2020). Koesjitische talen zijn dan ook meer verwant aan het Arabisch dan aan de overige Sub-Sahara-Afrikaanse talen. Er was dan ook

Madagaskar⁴³ na, spreekt de bevolking van deze landen voornamelijk Niger-Congo- en in mindere mate Nilo-Saharaanse talen.⁴⁴ Deze landen hebben veelal kunstmatig getrokken grenzen en kennen veel etno-linguïstische groepen.

Naar meer behapbare clusters

Een belangrijke informatieve eis bij de clustering van herkomstlanden is dat er voldoende empirische verfijning is. Vandaar dat we een meervoudige indeling in twaalf clusters gebruiken. Die geeft veel meer specifieke informatie over de positie van migrantengroepen dan een grove tweedeling. Niettemin zullen gebruikers van data waarbij een onderscheid naar herkomstgroep wordt gemaakt, soms ook behoefte hebben aan verdere vereenvoudiging in een geringer aantal categorieën. Vaak spelen daarbij praktische overwegingen een rol, zoals het geringe aantal migranten uit een bepaalde herkomstgroep of de beperkte omvang van een steekproef. De twaalf hierboven beschreven etno-linguïstische clusters kunnen in dat geval worden geaggregeerd tot een kleiner aantal clusters.

Zes of zeven etno-linguïstische clusters

Een eerste stap om dit te bewerkstelligen is door de Angelsaksische en Continentaal-Germaanstalige landen samen te voegen tot de groep ‘Anglo-Germaanstalige landen’. Het ligt vanwege de taalkundige overeenkomst tussen Spaans, Portugees, Frans en Italiaans ook voor de hand om de categorieën Middellandse Europa en Latijns-Amerika te verenigen. Daarnaast is er vanwege de religieuze overeenkomst veel voor te zeggen om de groepen Arabische landen en Centraal Azië onder te brengen in één cluster. De islam is in de landen van deze twee clusters namelijk de dominerende godsdienst. Vervolgens kunnen de landen in Azië die niet in het geografisch centrale hart van het continent liggen, samen met de kleine eilandstaten in de Stille Oceaan, worden gecategoriseerd in de overkoepelende cluster ‘Overig Azië en de Pacific’.

Ten slotte kunnen onderzoekers en beleidsmakers een zevende cluster ‘Cariben’ onderscheiden als ze geen apart onderscheid maken naar de landen die tot nu toe worden

zeker wat voor te zeggen geweest om de landen in de Hoorn van Afrika en Niger in dezelfde cluster onder te brengen die in onze indeling alleen de Arabische Landen bevat. In Niger wordt overwegend Hausa een Afro-Aziatische taal, behorend tot de Tsjadische tak van deze taalfamilie, gesproken (zie bijvoorbeeld Beidou 2014).

⁴³ Ofschoon Madagaskar relatief dicht bij het Afrikaanse vasteland ligt, werd het pas vrij laat bevolkt. Pas rond 650 arriveerden er zeevaarders uit de Indonesische archipel. Zij waren de eerste mensen die zich definitief op het eiland vestigden. De talen die op Madagaskar worden gesproken, behoren dan ook tot de Austronesische taalfamilie (zie bijvoorbeeld Serva et al. 2013).

⁴⁴ Zie bijvoorbeeld Greenberg 1963.

aangeduid als de ‘grote vier klassieke niet-westerse herkomstlanden’, te weten Turkije, Marokko, Suriname en de voormalige Nederlandse Antillen⁴⁵. Deze landen worden in veel publicaties afzonderlijk weergegeven. De bevolking van de cluster Cariben bestond begin 2020 voor 98,4 procent uit personen met een Surinaamse of Nederlands-Antilliaanse achtergrond. De overige personen in deze cluster vormen dan ook een relatief erg kleine groep. Aangezien de landen die dan overblijven allemaal officieel Engelstalig zijn, kunnen zij het beste worden ondergebracht bij de Anglo-Germaanstalige landen.⁴⁶

Uiteindelijk komen we dan tot deze zes of zeven clusters:

	<i>Mogelijke benaming in tabellen</i>
1. Anglo-Germaanstalige landen	Anglo-Germaans
2. Midden- en Oost-Europese landen	MOE-landen
3. Latijnstalige en andere mediterrane landen	Latijnstalig/M
4. Arabische landen en Centraal-Azië	Arabisch/CA
5. Overig Azië en de Pacific	Overig Azië/P
6. Sub-Sahara-Afrika	SS-Afrika
(7. Cariben	Cariben)

⁴⁵ Het land de Nederlandse Antillen omvatte sinds de totstandkoming van het Statuut voor het Koninkrijk der Nederlanden in 1954 zes eilanden in de Caribische Zee (Aruba, Bonaire, Curaçao, Sint Eustatius, Sint Maarten en Saba) die politieke autonomie binnen het Koninkrijk der Nederlanden verkregen. Vanaf 1 januari 1986 maakt Aruba geen deel meer uit van de Nederlandse Antillen. Het verkreeg op deze datum de gewenste ‘status aparte’ binnen het koninkrijk, waardoor het werd verlost van de als dominant ervaren positie van Curaçao. Op 10 oktober 2010 hield het land de Nederlandse Antillen op te bestaan. Naast Aruba werden nu ook Curaçao en Sint Maarten autonome landen binnen het Koninkrijk der Nederlanden. Bonaire, Sint Eustatius en Saba zijn weer deel gaan uitmaken van het Nederlandse staatsbestel. Tegenwoordig bestaat het Koninkrijk der Nederlanden dan ook uit vier landen: Nederland, Aruba, Curaçao en Sint Maarten. De Caribische delen van het koninkrijk noemen we in deze publicatie ‘de voormalige Nederlandse Antillen’. Daarbij verwijzen we dus naar de situatie van vóór 1 januari 1986 (de status aparte van Aruba).

⁴⁶ Als cijfers voor personen met een Surinaamse en Nederlands-Antilliaanse achtergrond apart worden weergegeven, behoren de overige landen uit het cluster Cariben tot de Angelsaksische landen indien de indeling met twaalf clusters wordt gehanteerd.

Een dichotomie voor specifieke analyses

In de inleiding van dit paper en in andere publicaties hebben we beargumenteerd dat het af te raden is om gebruik te maken van dichotomieën. Omdat dichotomieën de sociale werkelijkheid onnodig simplificeren, zijn uitkomsten weinig informatief zijn en is het risico op stereotypering groot. Er zijn echter enkele specifieke gevallen denkbaar waarin het maken van een dichotomie toch waardevol kan zijn. Zo kunnen daarmee relevante referentiecategorieën worden gecreëerd waarmee migranten uit specifieke herkomstlanden vergeleken kunnen worden. We hebben daarom de zeven onderscheiden clusters verder gereduceerd tot twee clusters. Daarvoor maken we opnieuw gebruik van taalkundige en religieuze scheidslijnen.

- In Anglo-Germaanstalige, Midden- en Oost-Europese en Latijntalige en andere mediterrane en Caribische landen spreekt het overgrote deel van de bevolking een Indo-Europese taal. Bovendien bevat de heersende cultuur onmiskenbaar christelijke elementen. Het is te verwachten dat migranten afkomstig uit deze landen, na correctie voor sociaal-economische achtergrondkenmerken, minder moeite hebben om te integreren in de Nederlandse samenleving.⁴⁷ Het Nederlands is ook een Indo-Europese taal en in Nederland zijn de sporen van het christendom duidelijk aanwezig. Een vergelijkbare religieuze achtergrond van migranten en de reeds aanwezige bevolking zorgt er doorgaans voor dat de integratie van migranten voorspoediger verloopt.⁴⁸
- Ook de landen in de overige drie clusters, Arabische landen en Centraal Azië, Overig Azië en de Pacific en Sub-Sahara-Afrika, kunnen we samennemen. In bijna alle landen in deze groep spreekt het grootste deel van de bevolking een niet-Indo-Europese taal en/of is het christendom niet een duidelijk aanwezige godsdienst.⁴⁹ De hypothese is dat migranten uit deze landen meer moeite kunnen hebben om te integreren in de Nederlandse samenleving, ook al zouden zij een sociaal-economische positie hebben die vergelijkbaar is met die van de doorsnee-inwoner van ons land.

⁴⁷ Daarmee blijft overigens wel de mogelijkheid buiten beeld dat het niet (alleen) om de daadwerkelijke culturele afstand gaat, maar ook om percepties en beeldvorming door anderen, waaronder racisme en discriminatie.

⁴⁸ Zie bijvoorbeeld Owen et al. 1981; Foner en Alba 2008; Kogan et al. 2020.

⁴⁹ De uitzonderingen hierop zijn allemaal relatief kleine eilandstaten. Personen met een achtergrond uit één van deze eilandstaten zijn nauwelijks woonachtig in Nederland. Kaapverdië is het enige herkomstland van deze staten dat, met bijna 23.000 personen in 2020, in kwantitatief opzicht van enige betekenis is. Het grootste deel van de bevolking op de Kaapverdische Eilanden is namelijk christen en spreekt een op het Portugees gebaseerde creolentaal (*Kriolu*) (zie bijvoorbeeld Carter en Aulette 2009).

Deze tweedeling is geen vervanging van de tweedeling westers/niet-westers. We adviseren met klem om deze tweedeling uitsluitend te gebruiken als referentiecategorie. Zo is deze tweedeling te gebruiken als referentiegroep bij onderzoek naar één specifiek cluster uit de hierboven genoemde etno-linguïstische clusters, bijvoorbeeld naar Latijns-Amerika, Midden- en Oost-Europese landen, of Arabische landen en Centraal Azië. Zo is het beter om alle mensen uit Sub-Sahara-Afrika te vergelijken met de tweede overkoepelende cluster dan met de totale populatie. Hetzelfde speelt wanneer we willen onderzoeken welke problemen migranten uit specifieke herkomstlanden, zoals Polen of Eritrea, ondervinden bij het vinden van betaalde arbeid. Het heeft weinig zin om de participatie van deze groepen te vergelijken met de arbeidsparticipatie van de totale bevolking of met die van de gehele bevolking met een migratieachtergrond. Om inzicht te krijgen in de specifieke problemen van Polen of Eritreeërs is het zinvoller om Polen te vergelijken met personen die qua taalfamilie en religie meer met hen verwant zijn. Zijn er over grote aantallen data beschikbaar, dan kunnen Polen vergeleken worden met de groep Midden- en Oost-Europese talen. Zijn die data niet beschikbaar, dan ligt een vergelijking voor de hand met de hele cluster van Angelsaksische, Continentaal-Germaanse, Midden- en Oost-Europese en Latijnstalige- en andere mediterrane en Caribische landen. En hetzelfde geldt voor Eritreeërs. Het ligt voor de hand hen te vergelijken met de cluster Arabische landen en Centraal Azië, Overig Azië en de Pacific en Sub-Sahara-Afrika. Als hun arbeidsparticipatie na correctie voor sociaal-economische factoren nauwelijks afwijkt van die van andere personen in dezelfde cluster van herkomstlanden, is er weinig aanleiding om specifiek beleid voor deze groep op te zetten.

Voor beschrijvende analyses is deze tweedeling niet geschikt. Het heeft geen nut om bijvoorbeeld schoolprestaties of werkloosheid van deze twee grote clusters te vergelijken in beschrijvende tabellen of grafieken. Daarvoor zijn ze intern veel te verschillend. Onderzoekers en beleidsmakers kunnen dan beter gebruikmaken van de bovengenoemde indelingen in twaalf of zeven clusters.

Een open vraag is nog welke *labels* we voor deze twee clusters zouden moeten hanteren. In tekstkader 3 bespreken we enkele opties.

Tekstkader 3: Mogelijke labels voor de tweedeling op basis van culturele verwantschap

De basis voor het hier gemaakte onderscheid is culturele verwantschap. Dat zou pleiten voor het label ‘verwant’ versus ‘niet-verwant’. Nadeel hiervan is dat dit opnieuw een onderschikking of een negatieve tegenstelling suggereert. Beter is het om hier een *glijdende schaal* te gebruiken, bijvoorbeeld door te kiezen voor labels als ‘cultureel dichterbij’ en ‘cultureel verder weg’.

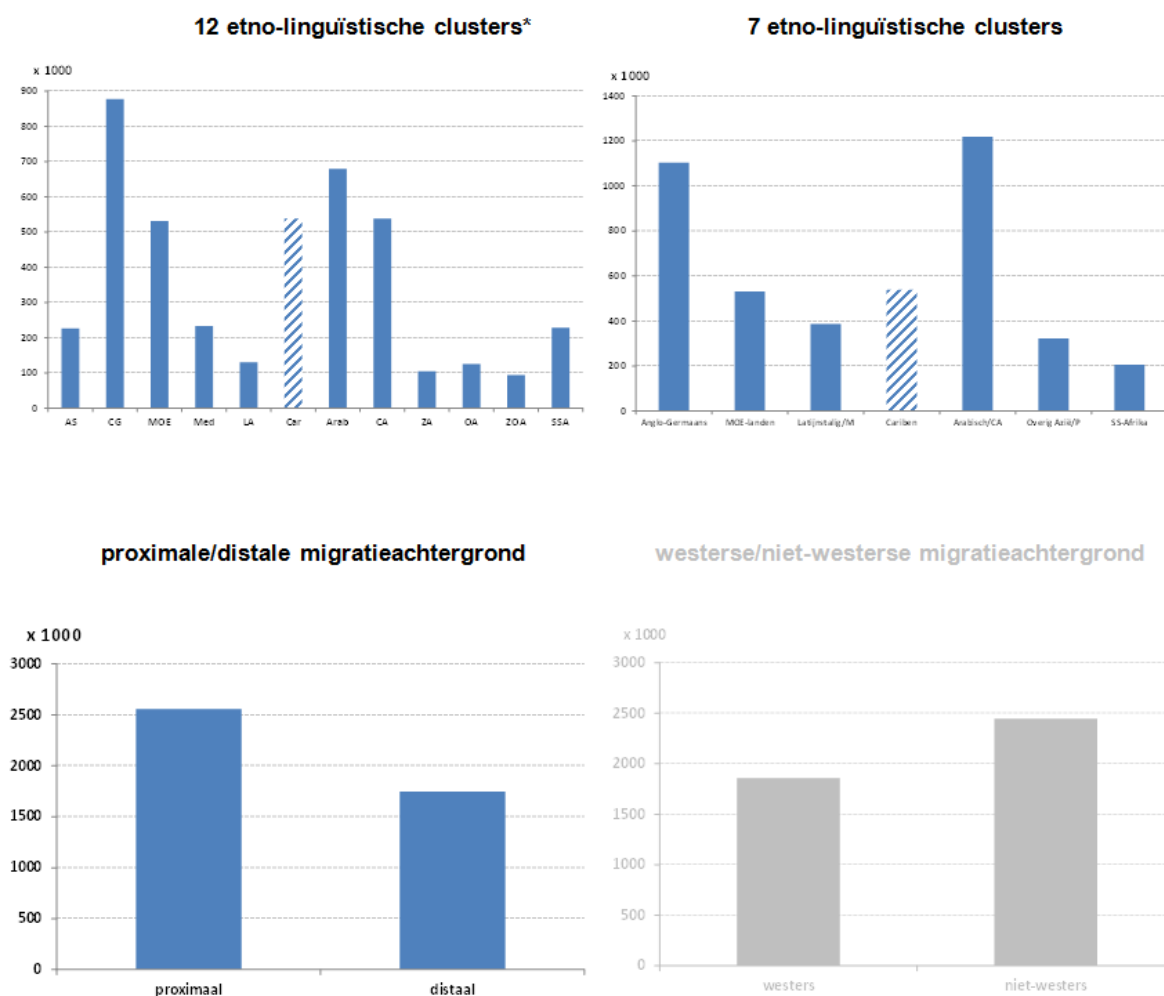
Een andere optie is om hiervoor *technische* labels te gebruiken die nog geen sterke performatieve effecten hebben. Om vergelijkbare reden hebben Hilde Verweij-Jonker en de WRR ooit het begrippenpaar autochtoon/allochtoon geleend uit de biologie en de geologie.⁵⁰ Dat is ook hier mogelijk. In de taalkunde worden bijvoorbeeld de termen ‘proximaal’ en ‘distaal’ gebruikt om dichterbij en verder weg gelegen taalgroepen aan te duiden. Hetzelfde geldt voor de aanduiding van sedimenten in de geologie en voor de anatomie, waar bijvoorbeeld onderscheid wordt gemaakt tussen proximale en distale lichaamsdelen. Dan zou de technische aanduiding van de referentiegroepen zijn: groepen met een ‘proximale migratieachtergrond’ en groepen met een ‘distale migratieachtergrond’.

Cijfermatige consequenties

Een belangrijke vraag is wat de cijfermatige consequenties zijn van alternatieve clusterings. Daarvoor kijken we naar de omvang van de verschillende groepen die uit de clustering voortkomen. Elk cluster dient in ieder geval te bestaan uit een substantieel aantal personen om geschikt te zijn voor publicatie op CBS-Statline. Figuur 1 toont de omvang van de verschillende clusters in het begin van het jaar 2020. Behalve de alternatieve clusters hebben we hier ter vergelijking ook de huidige verdeling in westers en niet-westers weergegeven.

⁵⁰ Zie bijvoorbeeld Groenendijk 2007

Figuur 1. De omvang van de verschillende clusters met herkomstgroepen, exclusief Nederlandse achtergrond, 1-1-2021



* AS: Angelsaksische landen; CG: Continentaal-Germaanstalige landen; MOE: Midden- en Oost-Europese landen; Med: Mediterrane landen; LA: Latijns-Amerika; Car: Cariben; CA: Centraal-Azië; ZA: Zuid-Azië; OA: Oost-Azië; ZOA: Zuidoost-Azië en de Pacific; SSA: Sub-Sahara-Afrika

Bron: CBS

Conclusie

Onderzoekers en beleidsmakers moeten clusters en labels gebruiken die zo informatief en neutraal mogelijk zijn. Daarom hebben we in de WRR-Verkenning *Migratie en Classificatie* en in de WRR-Policy Brief *Afscheid van westers en niet-westers* voorgesteld om in onderzoek en beleid afscheid te nemen van de standaardtweedeling ‘westers/niet-westers’. In de eerdere publicaties hebben we echter geen standaardalternatief voorgesteld. We menen namelijk dat de clustering altijd moet afhangen van de onderzoeksvraag. Onderzoekers en beleidsmakers

worden op deze manier aangespoord om zelf een meest passende clustering toe te passen, in plaats van de ingesleten, maar ook versleten, tweedeling ‘westers/niet-westers’ te gebruiken. In dit paper hebben we specifiek gekeken naar vraagstukken op het gebied van integratie en sociale samenhang waarbij culturele factoren relevant zijn. We hebben voor het integratieonderzoek en voor beleidsstatistieken alternatieve indelingen ontwikkeld om de veelheid aan herkomstlanden op een theoretisch onderbouwde, werkbare en informatieve manier te clusteren. De belangrijkste elementen hiervan zijn:

- De indeling is gebaseerd op de notie van culturele verwantschap. De gedachte hierachter is dat dit een relevante factor is bij de mate en snelheid van integratie van migranten en dat cultuur niet of nauwelijks valt te herleiden uit de overige registerdata van het CBS.
- Indicatoren voor culturele verwantschap zijn taal en religie. Deze indicatoren zijn objectiveerbaar en meetbaar.
- We gaan bij de aggregatie uit van de kenmerken van de in Nederland aanwezige migrantenpopulatie en niet van de overheersende taal en religie in het herkomstland zelf. Deze twee komen meestal overeen, maar dat hoeft niet altijd het geval te zijn.
- In onze clustering is de aggregatie van herkomstlanden die dicht bij Nederland liggen, gemiddeld genomen verfijnder dan bij verder weg gelegen landen.
- De bovenstaande uitgangspunten leiden tot een indeling in twaalf groepen van herkomstlanden, die kan worden teruggebracht tot zes of zeven clusters. Elke individuele groep bestaat uit een substantieel aantal personen en is geschikt voor publicatie op CBS-Statline.
- Mocht er behoefte zijn aan referentiecategorieën voor verdiepende analyses, dan is het eventueel mogelijk om uit deze zeven clusters een tweedeling te construeren. Deze tweedeling is niet geschikt voor beschrijvende statistieken.

Literatuur

- Achab, K. (2012) *Internal structure of verb meaning: A study of verbs in Tamazight (Berber)*, Cambridge: Cambridge Scholars Publishing.
- Adserà, A. en A.M. Ferrer (2021) 'Linguistic proximity and the labour market performance of immigrant men in Canada', *Labour*, 35(1): 1-23.
- Ashkanasy, N.M., E. Trevor-Roberts en L. Earnshaw (2002) 'The Anglo cluster: Legacy of the British empire', *Journal of World Business*, 37(1), 28-39.
- Auer, D. (2018) 'Language roulette – the effect of random placement on refugees' labour market integration', *Journal of Ethnic and Migration Studies*, 44(3): 341-362.
- Bakacsi, G., S. Takács, A. Karácsonyi en V. Imre (2002) 'Eastern European cluster: Tradition and transition', *Journal of World Business*, 37(1): 69-80.
- Beidou, H. (2014) *Dynamiques linguistiques au Niger: Cohabitation et/ou concurrence entre le Français et les langues locales*, Montréal: Université du Québec.
- Bellah, R.N. (1967) 'Biblical religion and civil religion in America', *Dædalus: Journal of the American Academy of Arts and Sciences*, 96(1): 1-21.
- Berger, P.L. en T. Luckmann (1966) *The social construction of reality: A treatise in the sociology of knowledge*, New York: Doubleday.
- Blom, M. en R. Jennissen (2014) 'The involvement of different ethnic groups in various types of crime in the Netherlands', *European Journal on Criminal Policy and Research*, 20(1): 51-72.
- Bovens, M., M. Bokhorst, R. Jennissen en G. Engbersen (2016) *Migratie en classificatie: Naar een meervoudig migratie-idioom*, WRR Verkenning nr. 34, Den Haag: Wetenschappelijke Raad voor het Regeringsbeleid.
- Bovens, M., R. Jennissen, M. Bokhorst en G. Engbersen (2021) *Afscheid van westers en niet-westers: Alternatieven voor de indeling van herkomstgroepen*, WRR Policy Brief nr. 8, Den Haag: Wetenschappelijke Raad voor het Regeringsbeleid.
- Burke, P. (2009) *Cultural hybridity*, Cambridge: Polity Press.
- Carter, K. en J. Aulette (2009) 'Creole in Cape Verde: Language, identity and power', *Ethnography*, 10(2): 213-236.
- Caruana, S. (2013) 'Italian in Malta: A socio-educational perspective', *International Journal of Bilingual Education and Bilingualism*, 16(5): 602-614.
- Cattell, R.B. (1950) 'The principal culture patterns discoverable in the syntal dimensions of existing nations', *The Journal of Social Psychology*, 32(2): 215-253.

- Chen, L. (2017) 'Historical and cultural features of confucianism in East Asia', blz. 102-111 in R.T. Ames and P.D. Herschok (red.) *Confucianisms for a changing world cultural order*, Honolulu: University of Hawaii Press.
- Choenni, C. en C. Harmsen (2007) 'Geboorteplaats en etnische samenstelling van Surinamers in Nederland', *Bevolkingstrends*, 1e kwartaal 2007: 74-78, Den Haag: Centraal Bureau voor de Statistiek.
- Crowe, D.M. (2008) 'The Roma in post-communist Eastern Europe: Questions of ethnic conflict and ethnic peace', *Nationalities Papers*, 36(3): 521-552.
- Dagevos, J. (2001) *Perspectief op integratie; over de sociaal-culturele en structurele integratie van etnische minderheden in Nederland*, WRR-werkdocument nr. 121, Den Haag: Wetenschappelijke Raad voor het Regeringsbeleid.
- Dustmann, C. en F. Fabbri (2003) 'Language proficiency and labour market performance of immigrants in the UK', *The Economic Journal*, 113: 695-717.
- Enfield, N.J. en B. Comrie (2015) 'Mainland Southeast Asian languages', blz. 1-28 in N.J. Enfield en B. Comrie (red.), *Languages of mainland Southeast Asia: The state of the art*, Berlijn: De Gruyter Mouton.
- Foner, N. en R. Alba (2008) 'Immigrant religion in the US and Western Europe: Bridge or barrier to inclusion?', *International Migration Review*, 42(2): 360-392.
- Gijsberts, M. en M. Lubbers (2013) *Nieuw in Nederland: Het leven van recent gemigreerde Bulgaren en Polen*, Den Haag: Sociaal en Cultureel Planbureau.
- Greenberg, J.H. (1963) *The languages of Africa*, Bloomington: Indiana University Press.
- Groenendijk, K. (2007) 'Allochtonen of burgers: definitiemacht in debat en wetgeving over immigranten', blz. 101-112 in J.W. Duyvendak, G. Engbersen, M. Teeuwen en I. Verhoeven (red.) *Macht en verantwoordelijkheid. Essays voor Kees Schuyt*, Amsterdam: Amsterdam University Press.
- Gupta, V., P.J. Hanges en P. Dorfman (2002) 'Cultural clusters: Methodology and findings', *Journal of World Business*, 37(1): 11-15.
- Hofstede, G. (1980) *Culture's consequences: International differences in work related values*, Newbury Park (CA): SAGE Publication.
- House, R.J., P.J. Hanges, M. Javidan, P. Dorfman en V. Gupta (2004) *Culture, leadership and organizations. The GLOBE study of 62 societies*, Thousand Oaks (CA); SAGE Publication.
- Inglehart, R. en C. Welzel (2010) 'Changing mass priorities: The link between modernization and democracy', *Perspectives on Politics*, 8(2): 551-567.
- Isphording, I.E. en S. Otten (2014) 'Linguistic barriers in the destination language acquisition of immigrants', *Journal of Economic Behavior & Organization*, 105: 30-50.

- Jennissen, R.P.W., N.L. van der Gaag en L.J.G. van Wissen (2006) 'Searching for similar international migration trends across countries in Europe', *Genus: Journal of Population Sciences*, 62(2): 37-64.
- Jennissen, R., G. Engbersen, M. Bokhorst en M. Bovens (2018) *De nieuwe verscheidenheid: Toenemende diversiteit naar herkomst in Nederland*, WRR Verkenning nr. 38, Den Haag: Wetenschappelijke Raad voor het Regeringsbeleid.
- Kalule-Sabiti, I., D. Lucas en Y. Amoateng (2003) 'Stripping South Africa of its human capital', *Development Bulletin*, 62: 126-128.
- Kogan, I., E. Fong en J.G. Reitz (2020) 'Religion and integration among immigrant and minority youth', *Journal of Ethnic and Migration Studies*, 42(2): 197-216.
- Kroeber, A.L. en C. Kluckhohn (1952) *Culture: a critical review of concepts and definitions*, Papers of the Peabody Museum of Archaeology & Ethnology, Harvard: Harvard University.
- Latomaa, S. en P. Nuolijärvi (2002) 'The language situation in Finland', *Current Issues in Language Planning*, 3(2): 95-202.
- Manan, S.A. en M.K. David (2014) 'Mapping ecology of literacies in educational setting: The case of local mother tongues vis-à-vis Urdu and English languages in Pakistan', *Language and Education*, 28(3): 203-222.
- Miller, R.A. (1971) *Japanese and the other Altaic languages*. Chicago: University of Chicago Press.
- Mooij, M. de, D. Dieleman en S. de Regt (red.) (2020) *Jaarrapport Integratie 2020*, Den Haag: Centraal Bureau voor de Statistiek.
- Öberg, S. en H. Boubnova (1993) 'Ethnicity, nationality and migration potentials in Eastern Europe', blz. 234-256 in R. King (red.) *Mass migration in Europe: The legacy and the future*, Londen: Belhaven.
- Oostindie, G. (2010) *Postkoloniaal Nederland: Vijfenzestig jaar vergeten, herdenken, verdringen*, Amsterdam: Bert Bakker.
- Ortensi, L.E. en V. Riniolo (2020) 'Do migrants get involved in politics? Levels, forms and drivers of migrant political participation in Italy', *Journal of International Migration and Integration*, 21(1): 133-153.
- Owen, C.A., H.C. Eisner en T.R. McFaul (1981) 'A half-century of social distance research: National replication of the Bogardus' studies', *Sociology and Social Research*, 66(1): 80-98.
- Porkhomovsky, V. (2020) 'Afro-Asiatic overview', blz. 269-274 in R. Vossen en G.J. Dimmendaal (red.) *The Oxford handbook of African languages*, Oxford: Oxford University Press.

- Rao, V. en M. Walton (eds.) (2004) *Culture and public action*, Stanford: Stanford University Press.
- Schepens, J.J., F. van der Slik en R. van Hout (2016) 'L1 and L2 distance effects in learning L3 Dutch', *Language Learning*, 66(1): 224-256.
- Serva, M., F. Petroni, D. Volchenkov en S. Wichmann (2013) 'From Indonesia to Madagascar: In search of the origins of the Malagasy language', *Indian Ocean Review of Science and Technology*, artikel 1-2013.
- Sindwell, P. en M. Jenny (2021) 'Introduction', blz. 1-20 in P. Sindwell en M. Jenny (red.), *The languages and linguistics of mainland Southeast Asia: A comprehensive guide*. Berlijn: De Gruyter Mouton.
- Thuesen, F. (2017) 'Linguistic barriers and bridges: Constructing social capital in ethnically diverse low-skill workplaces', *Work, Employment and Society*, 31(6): 937-953.
- Tubergen, F. van en M. Kalmijn (2005) 'Destination-language proficiency in cross-national perspective: A study of immigrant groups in nine western countries', *American Journal of Sociology*, 110(5): 1412-1457.
- UNDP (2019) *Human development report 2019*, New York: UNDP.
- Venkateswarlu, K. (2020) *Colonialism, orientalism and the Dravidian languages*, New Delhi: Routledge.
- Wong, J.S., S.K. Ramakrishnan, T. Lee, J. Junn en J. Wong (2011) *Asian American political participation: Emerging constituents and their political identities*, New York: Russell Sage Foundation.
- WRR (2020) *Samenleven in verscheidenheid: Beleid voor de migratiesamenleving*, WRR Rapport nr. 103, Den Haag: Wetenschappelijke Raad voor het Regeringsbeleid.
- Yuan, Z. (2007) 'Dressing for power: Rite, costume, and state authority in Ming Dynasty China', *Frontiers of History in China*, 2(2): 181-212.

Bijlage 1. De indeling van herkomstlanden in geaggregeerde groepen op basis van culturele verwantschap (taal en religie)

Anglo-Germaanstalige landen

Angelsaksische landen – Australië, Canada, Falklandeilanden, Gibraltar, Ierland, Isle of Man, Kanaaleilanden, Nieuw-Zeeland, Puerto Rico, Sint Helena, Verenigde Staten, Verenigd Koninkrijk, Zuid-Afrika

Continentaal-Germaanstalige landen – België, Denemarken, Duitsland, Faeröer Eilanden, Finland, Groenland, Indonesië/voormalig Nederlands-Indië, IJsland, Liechtenstein, Luxemburg, (Nederland), Noorwegen, Oostenrijk, Zweden, Zwitserland

Midden- en Oost-Europese landen

Midden- en Oost-Europa – Albanië, Armenië, Bosnië-Herzegovina, Bulgarije, Estland, Georgië, Hongarije, Kosovo, Kroatië, Letland, Litouwen, Moldavië, Montenegro, Noord-Macedonië, Oekraïne, Polen, Roemenië, Rusland, Servië, Slovenië, Slowakije, Tsjechië, voormalig Joegoslavië, voormalige Sovjet-Unie, voormalig Tsjechoslowakije, Wit-Rusland

Latijnstalige en andere mediterrane landen

Mediterrane landen – Andorra, Cyprus, Frankrijk, Griekenland, Israël, Italië, Malta, Monaco, Portugal, San Marino, Spanje

Latijns-Amerika – Argentinië, Bolivia, Brazilië, Chili, Colombia, Costa Rica, Cuba, Dominicaanse Republiek, Ecuador, El Salvador, Frans Guyana, Guadeloupe, Guatemala, Haïti, Honduras, Martinique, Mexico, Nicaragua, Panama, Paraguay, Peru, Saint Pierre en Miquelon, Uruguay, Venezuela

Cariben

Cariben – Amerikaanse Maagdeneilanden, Anguilla, Antigua en Barbuda, Aruba, Bahama's, Barbados, Belize, Bermuda, Britse Maagdeneilanden, Dominica, Caribisch Nederland, Curaçao, Grenada, Guyana, Jamaica, Kaaiman eilanden, Montserrat, (voormalige) Nederlandse Antillen, Sint Kitts en Nevis, Sint Lucia, Sint Maarten, Sint Vincent en Grenadines, Suriname, Trinidad en Tobago, Turks en Caicoseilanden

Arabische landen en Centraal-Azië

Arabische landen – Algerije, Bahrein, Egypte, Irak, Jemen, Jordanië, Koeweit, Libanon, Libië, Marokko, Mauritanië, Oman, Gazastrook en Westelijke Jordanoever, Qatar, Saudi-Arabië, Soedan, Syrië, Tunesië, Verenigde Arabische Emiraten

Centraal-Azië – Afghanistan, Azerbeidzjan, Iran, Kazachstan, Kirgizië, Oezbekistan, Tadzjikistan, Turkije, Turkmenistan

Overig Azië en de Pacific

Zuid-Azië – Bangladesh, Bhutan, Brits territorium in de Indische Oceaan, India, Maldiven, Nepal, Pakistan, Sri Lanka

Oost-Azië – China, Hongkong, Japan, Macau, Mongolië, Myanmar, Noord-Korea, Taiwan, Zuid-Korea

Zuidoost-Azië en de Pacific – Amerikaans Samoa, Brunei, Cambodja, Cookeilanden, Fiji, Filipijnen, Frans-Polynesië, Guam, Kiribati, Laos, Maleisië, Marshall eilanden, Micronesië, Nauru, Nieuw-Caledonië, Niue, Noordelijke Marianen, Norfolk, Palau, Papoea Nieuw-Guinea, Pitcairneilanden, Singapore, Solomoneilanden, West-Samoa, Thailand, Timor-Leste, Tokelau, Tonga, Tuvalu, Vanuatu, Vietnam, Wallis en Futuna

Sub-Sahara-Afrika

Sub-Sahara-Afrika – Angola, Benin, Botswana, Burkina Faso, Burundi, Centraal Afrikaans Republiek, Comoren, Congo-Brazzaville, Congo-Kinshasa, Djibouti, Equatoriaal-Guinea, Eritrea, Ethiopië, Gabon, Gambia, Ghana, Guinee, Guinee-Bissau, Ivoorkust, Kaapverdië (→ Proximale migratieachtergrond/Latijnstalig, zie voetnoot 49), Kameroen, Kenia, Lesotho, Liberia, Madagaskar, Malawi, Mali, Mauritius, Mayotte, Mozambique, Namibië, Niger, Nigeria, Réunion, Rwanda, Sao Tomé en Príncipe, Senegal, Seychellen, Sierra Leone, Somalië, Swaziland, Tanzania, Togo, Tsjaad, Uganda, Zambia, Zimbabwe, Zuid-Soedan