
ONDERZOEK NAAR DE GEVOLGEN VAN HERHAALDELIJKE BLOOTSTELLING AAN GEZICHTEN VAN VERDACHTEN IN POLITIEAPPS VOOR AGENTEN OP PATROUILLE

ACTIVITEITENRAPPORT POLITIE EN WETENSCHAP

DR. ANNA SAGANA

UNIVERSITEIT MAASTRICHT, NEDERLAND

Contactpersoon

Dr. Anna Sagana
Sectie Forensische Psychologie, afdeling Klinische Psychologische Wetenschappen
Faculteit der Psychologie en Neurowetenschappen
Universiteit Maastricht
Postbus 616, 6200 MD Maastricht, Nederland
Tel: +31 43-388 3255
Fax: +31 43-388 4196
mail: anna.sagana@maastrichtuniversity.nl

SAMENVATTING

Het doel van dit project is na te gaan in hoeverre het risico van het aselekt etnisch profileren van groepen burgers als gevolg van herhaalde blootstelling aan gezichten in politiesystemen realistisch is en praktische oplossingen aan te reiken om stereotypering als gevolg van herhaalde blootstelling te verminderen of zelfs te voorkomen. Daartoe hebben we twee experimenten opgezet.

In studie 1 onderzochten we of herhaalde blootstelling aan etnische doelgroepen via politieoperatiesystemen ertoe leidt dat gebruikers criminele gezichten stereotyperen en op hun beurt ongedifferentieerde aanhoudings- en fouilleringstaken uitvoeren. Deelnemers (politieagenten en universiteitsstudenten) voerden een gezichtsherkenningstaak en een staande houden en fouilleren taak uit en beoordeelden de typische kenmerken van een crimineel voor en na blootstelling aan verschillende doelgezichten. De resultaten komen overeen met een typisch "own-race bias" (ORB; eigen-ras bias) effect voor zowel politieagenten als studenten. Deelnemers waren nauwkeuriger in hun herkenningsbeslissingen voor hun eigen gezichten dan voor gezichten van een andere etniciteit. Bovendien was het percentage een staande houden en fouilleringen bij politieagenten hoger voor folies/onschuldigen van een andere dan van de eigen etniciteit. Wat betreft de typische kenmerken van een crimineel, waren er weliswaar geen grote veranderingen voor en na de blootstelling aan de doelgezichten, maar etniciteit stond consequent in de top 5 van kenmerken voor alle delicten.

In studie 2 werd onderzocht hoe de presentatie van gezichten kan worden aangepast (door middel van individualisering, volgorde en instructies) om het risico van stereotypering te verminderen. De taak werd geprogrammeerd en herzien op basis van de feedback van P&W en er werd ethische goedkeuring verkregen. We konden echter niet beginnen met het verzamelen van gegevens, omdat het erg moeilijk was agenten te vinden die wilden deelnemen. Onderbezetting, gecombineerd met de inzet van agenten bij de protesten coronamaatregelen, waren hiervoor de belangrijkste redenen. De studie uitvoeren met een studentensteekproef werd niet als een goed alternatief beschouwd. Bijgevolg werd in onderling overleg besloten dit tweede experiment niet voort te zetten en het project te beëindigen.

Op basis van de resultaten van dit onderzoek kan geconcludeerd worden dat herhaalde blootstelling aan verdachte gezichten, zoals gepresenteerd in deze studie, niet leidt tot willekeurige verdenkingen. Echter, mensen uit andere etnische groepen hebben een groter risico om verkeerd geïdentificeerd te worden of ten onrechte als verdachte te worden aangezien. Belangrijk is dat etniciteit verre van irrelevant is voor de manier waarop agenten denken over criminaliteit.

ONDERZOEK NAAR DE GEVOLGEN VAN HERHAALDELIJKE BLOOTSTELLING AAN GEZICHTEN VAN VERDACHTEN IN POLITIEAPPS VOOR AGENTEN OP PATROUILLE

In het afgelopen decennium steeg het gebruik van nieuwe technologieën en slimme apps in het dagelijks leven. Zo investeerde ook de politie in het opzetten van gespecialiseerde eenheden (bijv. Politie Innovatie Lab, iLab) met als doel om operationele successen te vergroten en nieuwe technologieën te integreren in de dagelijkse praktijk van de politie. Dit heeft geleid tot de ontwikkeling van een breed scala aan systemen en apps, waaronder BlueFocus, Finder, KOPS en de nieuwste app voor de dagelijkse briefing. Deze systemen hebben tot doel om agenten te ondersteunen bij het onthouden van een groot aantal gezochte personen door het tonen van fotoreeksen. BlueFocus is bijvoorbeeld een intranet webapplicatie die gezichten toont voor herkenning en KOPS is een recent ontwikkelde app die politieagenten kan trainen om verdachten op straat te herkennen. De aandacht die wordt besteed aan het beter, sneller en slimmer maken van apps staat echter niet in verhouding tot de cognitieve en gedragsmatige gevolgen van het gebruik ervan. Dit wordt met name duidelijk als blijkt dat agenten wel op technisch niveau leren hoe ze deze apps moeten gebruiken, maar geen richtlijnen ontvangen over welk type inhoud effectief is om het gebruik van deze apps te maximaliseren. Hoewel de bredere ethische en maatschappelijke gevolgen van het gebruik van deze technologieën deel uitmaakten van een recente taskforce binnen de politie, ontbreekt empirisch onderzoek dat gericht is op het toetsen van een aantal van deze zorgen. In dit project willen we de gevolgen onderzoeken van het gebruik van apps die gezichten van verdachten tonen. Ons uiteindelijke doel is om richtlijnen te geven over de beste manier waarop inhoud kan worden getoond met als doel de prestaties in het gebruik van nieuwe technologieën te verbeteren.

ONDERWERP EN PROBLEEMSTELLING VAN HET BEOOGDE ONDERZOEK

Aangezien veel van de nieuwe toepassingen gepaard gaan met herhaalde blootstelling aan gezichten, moeten we nadenken over de manier waarop gezichten vertrouwd raken en hoe herhaald kijken de cognitie beïnvloedt. We weten bijvoorbeeld dat veelvuldige maar passieve blootstelling aan gezichten, zelfs wanneer dit herhaaldelijk plaatsvindt, onvoldoende is om gezichtsherkenning te verbeteren. Onderzoekers zijn het erover eens dat het de kwaliteit en niet de kwantiteit van de blootstelling is die bepalend is voor de herkenbaarheid [1, 2]. Dat komt omdat een succesvolle herkenning van onbekende gezichten de verwerking van kenmerken vereist met een hoge onderscheidingsgevoeligheid. Dit zijn kenmerken die het belangrijkste zijn voor herkenning (bijv. oogkleur, oogvorm, wenkbrauw dikte, lippendikte)[3]. Hoewel de onderzoekers in de loop der jaren verschillende soorten kenmerken hebben overwogen die een rol kunnen spelen bij gezichtsherkenning (configurele informatie vs. afzonderlijke detailkenmerken [4, 5], of externe vs. interne gezichtskenmerken [6, 7]), is de consensus dat ondoeltreffende verwerking van dergelijke kenmerken resulteert in herkenningsfouten.

Een goed gedocumenteerd fenomeen, en misschien wel het duidelijkst om de grenzen van gezichtsherkenning te illustreren, is de vooringenomenheid van het eigen etniciteit (own-race bias, ORB). ORB geeft aan dat we slecht in staat zijn om gezichten van een andere etniciteit te onderscheiden van gezichten van de eigen etniciteit [8-10]. In de loop der jaren is gesuggereerd dat mensen gezichten van de eigen etnische groep nauwkeuriger en efficiënter in het geheugen opslaan. Omgekeerd worden gezichten van andere etnische groepen opgeslagen met minder of ongeschikte herkenningspunten. Daarom heeft het onthouden van gezichten van een andere etniciteit waarschijnlijk een gebrek aan goede diagnostische kenmerken wat resulteert in verwarring en vaagheid [10-12]. Een van de redenen voor dit nadeel is sociale categorisering en gebrek aan individualisering. Sporer [9] stelt dat voor leden van een in-group de gelaatsverwerking automatisch begint met configurele codering van kenmerken, dat betekent dat de focus ligt op de onderlinge relaties tussen de meer opvallende gezichtskenmerken en die kenmerken zelf. Voor de leden van de out-group daarentegen leidt alleen al de perceptie van een out-group tot een indeling in een basisniveau van categorisering voordat men zelfs maar begint met het in detail verwerken van een gezicht. In overeenstemming met dit idee zijn er studies die suggereren dat er zich in de hersenen neuronen in het fusiforme aangezichtshersengebied bevinden die specifiek persoonlijke kenmerken waarnemen van mensen die behoren tot dezelfde etnische groep als de waarnemer zelf, terwijl bij het waarnemen van mensen die behoren tot een andere etnische groep, de neuronen juist meer informatie waarnemen over sociale kenmerken [13]. Daarom ervaren mensen zelfs op neuronaal niveau een intergroepscompetitie die er vaak toe leidt dat individuen gezichten van de in-group tot in detail verwerken, terwijl gezichten van leden van de out-group oppervlakkig worden verwerkt.

Interessant is dat Sporer's model een herkenningsgebrek voorspelt voor gezichten van elke groep van individuen die verschillen op basis van enkele markante karakteristieken; dat betekent dat ze duidelijke herkenningspunten hebben van een out-group (bijv. gezichtstatoeages, zichtbare littekens, opvallend kapsel). In feite zijn de prestaties op het gebied van gezichtsherkenning niet alleen slechter voor een andere etniciteit, maar ook voor andere leeftijdsgroepen (vooringenomenheid van de eigen leeftijd [14, 15]). In een recent onderzoek [2] werd vastgesteld dat neonatologieverpleegkundigen die voortdurend pasgeborenen zien, gezichten van nieuwgeborenen niet beter herkennen dan een controlegroep van gelijke leeftijd. De bevinding geeft aan dat gezichtscategorisering voor elke groep kan gebeuren wanneer er geen noodzaak is om de leden te individualiseren (d.w.z. geen kwalitatieve verwerking). Aangezien de standaardperiode van hospitalisatie voor pasgeborenen 2 tot 4 dagen bedraagt, is het niet nodig dat verpleegkundigen de pasgeborenen op basis van hun gelaatskenmerken leren onderscheiden. Echter, toen deelnemers werden getraind om pasgeboren gezichten te individualiseren door elk gezicht te koppelen aan een naam [zie 16, 17 voor een vergelijkbare aanpak], nam de herkenningsnauwkeurigheid toe voor nieuwe gezichten van pasgeborenen van 59,4% (pre-training) tot 71,7% (post-training).

Een verschuiving in de manier waarop individuen reageren op gezichten van een out-group kan ook te wijten zijn aan de invloed van stereotypering. Mensen vereenvoudigen complexe situaties door vuistregels toe te passen in plaats van moeizaam, normatief denken om dubbelzinnige situaties het hoofd te bieden. Meerdere ontmoetingen

met dezelfde stimulus leiden tot de ontwikkeling van een concept, een norm of een stereotype. In de context van het strafrecht heeft onderzoek criminele stereotypen gedocumenteerd - de opvattingen van mensen over het soort personen dat bepaalde misdrijven pleegt - die verband houden met een grote verscheidenheid aan persoonlijke kenmerken, waaronder o.a. etniciteit, geslacht, leeftijd, opleidingsniveau, sociaaleconomische status en geestelijke gezondheidskwesties [18-21]. Met name Smalarz en collega's [18] toonden aan dat criminele stereotypering een potentiële bron van vooringenomenheid kan zijn, zelfs bij forensische bewijsanalyse. Wanneer de vingerafdruk van een verdachte vergezeld ging van persoonsgegevens (bijv. etniciteit, geslacht van de verdachte) en de verdachte paste in het criminele stereotype, zagen de deelnemers de vingerafdrukken vaker overeenkomen dan wanneer de verdachte niet aan het stereotype voldeed. Met name bleek dat de deelnemers zich niet bewust waren van de mate waarin een crimineel stereotype hun beoordelingen had bevooroordeeld (d.w.z. vooringenomen blinde vlek). Ook de aanwezigheid van een bekentenis verhoogde de schuldbeleving wanneer een verdachte stereotiep bij het misdrijf paste, maar had geen effect op de schuldperceptie wanneer een verdachte contethniciteitereotiep was [22].

Alles bij elkaar genomen suggereren de bevindingen dat wanneer situationele factoren een categorisch denkkader bevorderen, de prestatie van gezichtsherkenning kan worden verslechterd. Daarnaast kan herhaaldelijke blootstelling aan personen die op het eerste gezicht soortgelijke uiterlijke kenmerken hebben (bijv. hetzelfde geslacht, dezelfde leeftijd, gezichtstatoeages of sieraden) stereotiep denken stimuleren of activeren en daardoor negatieve invloed hebben op het vermogen om unieke gezichtskenmerken waar te nemen en leiden tot een voortijdig oordeel over de schuldvraag. Door deze bevindingen te vertalen naar de dagelijkse praktijk, kunnen politieagenten die herhaaldelijk worden blootgesteld aan gezichten van gezochte of verdachte personen, simplistische vuistregels ontwikkelen over groepen mensen, wat kan leiden tot ongedifferentieerd profileren van leden van deze groep. Politieagenten kunnen de identiteit van gezochte verdachten door elkaar gaan halen in plaats van ze te individualiseren, wat kan leiden tot persoonsverwisselingen en valse positieven.

Aan de andere kant is het onduidelijk of herhaaldelijke blootstelling aan bepaalde sociale groepen kan leiden tot een betere of juist slechtere houding tegenover deze groepen in de samenleving [23-25]. Volgens het blootstellingseffect kan het herhaaldelijk waarnemen van een object [26] of persoon [27] de houding tegenover dat object of die groepering verbeteren. Desondanks heeft herhaaldelijke blootstelling aan gezichten met een witte huidskleur geleid tot meer vooroordelen en een negatievere houding ten opzichte van mensen met een donkere huidskleur, in plaats van een betere houding tegenover mensen met een witte huidskleur [24]. Een tweede verklaring voor het ORB is het gebrek aan perceptuele ervaring met de out-group [8, 12, 15, 28]. Volgens deze visie worden de karakteristieken die men gebruikt om gezichten te coderen verfijnd door ervaring en daarom beter geschikt voor gezichten die behoren tot de eigen etniciteit en leeftijdsgroep [15, 29]. Tot slot is een prominente oplossing om de vooroordelen tegen de out-group te verminderen de bevordering van intergroepscontacten [30, 31]. De mate waarin contact de vooroordelen vermindert, lijkt echter te variëren met de aard van de doelgroep en de kwaliteit van het contact [30,

31]. Derhalve heeft herhaalde blootstelling aan gezichten het potentieel om het herkenningsvermogen te verbeteren en zal geen bedreiging vormen, als de presentatiemethode zorgvuldig is uitgewerkt.

Het doel van dit project is na te gaan in hoeverre het gevaar van het aselekt etnisch profileren van groepen burgers als gevolg van herhaalde blootstelling aan gezichten realistisch is en praktische oplossingen aan te reiken om stereotypering als gevolg van herhaalde blootstelling te verzachten of zelfs te voorkomen. In studie 1 wordt met name bestudeerd of herhaalde blootstelling aan doelgroepen ertoe leidt dat app-gebruikers een stereotiep idee van een crimineel gezicht en ongedifferentieerd etnisch profileren. Om te beoordelen of het gevaar voor alle doelgroepen gelijk is, zullen we gebruik maken van andere en eigen etnische gezichten. Studie 2 bestudeert of er mogelijkheden zijn om het gevaar van stereotypering te verminderen door de wijze te veranderen waarop de gezichten worden gepresenteerd, bijv. door individualisering, volgorde en instructies.

STUDIE 1

OPZET

Studie 1 maakt gebruik van een gemengd factoriaal design. Blootstelling aan de doelgroep (voor en na) als interindividuele variabelen. De deelnemersgroepen (politieagenten vs. algemene bevolking) als intergroepsvariabele. De afhankelijke variabelen voor beide onderzoeken zijn: a) de frequentie van staande houden en fouilleren, b) a) de gemeten nauwkeurigheid van staande houden en fouilleren apart voor doelen en folies, c) de gemeten nauwkeurigheid van de herkenning apart voor doelen en folies en d) de mate van overeenstemming tussen de typische kenmerken van een crimineel.

DEELNEMERS

Om de steekproefgrootte te bepalen hebben we a priori vermogensanalyses uitgevoerd voor gemengde ANOVA's met behulp van Gpower [35]. Voor Studie 1 hebben we 132 deelnemers nodig om een gemiddelde effectgrootte $f = 0,20$, met een vermogen van 0,90, op een $p = ,05$ significantieniveau te detecteren.

De politieagenten werden geworven via de politiebureaus in Limburg en Oost-Brabant, de politieacademie en het grote publiek via de Universiteit van Maastricht. De participatiegraad was echter laag. Dit had te maken met onderbezetting bij de politie en politiemensen die de relevantie van het onderzoek niet altijd inzagen of teleurgesteld waren in de resultaten van academisch onderzoek (zie ook [44]). Daarnaast volgden verschillende deelnemers minder dan 3 trainingen en werden daarom uitgesloten van de analyse, omdat blootstelling niet voldoende werd geacht om vertrouwdheid met een gezicht te bereiken. Ten slotte gebruikten sommige deelnemers inconsistente gebruikersnamen, waardoor het onmogelijk was om de verschillende sessies betrouwbaar aan één identiteit te koppelen. Ook deze records werden uit de analyse verwijderd.

Na alle verwijderingen bestond de steekproef uit 67 politieagenten die fase één voltooiden en 34 deelnemers die fase twee voltooiden. Voor het grote publiek waren er 27 deelnemers, bijna allemaal universiteitsstudenten. Het aantal leden van het grote publiek werd laag gehouden om de verschillen tussen de steekproeven klein te houden.

MATERIALEN

KOPS APP. De KOPS-applicatie was een nieuw ontwikkeld systeem dat bedoeld is om agenten op te leiden om verdachten en gezochte personen gemakkelijker te herkennen in een menigte. De KOPS-app diende als proxy, aangezien het in deze app eenvoudiger is om inhoud toe te voegen in vergelijking tot andere gecentraliseerde systemen. Het is opgezet als een 'serious game' en vereist van gebruikers dat zij een aantal vragen beantwoorden door de juiste persoon te selecteren uit vier mogelijke antwoordopties. Deelnemers werd gevraagd de app gedurende 5 dagen minimaal 5 min per dag te gebruiken, gebaseerd op eerder werk [16]. Omdat niet iedereen aan deze instructie voldeed, nemen we deelnemers op die ten minste 3 dagen getraind hadden.

FOTO'S. We hebben gebruik maken van frontale hoofd- en schouderfoto's van Kaukasische en Noord-Afrikaanse personen. Kaukasische personen worden geselecteerd op duidelijke out-group kenmerken (bijv. gezichtstatoeages, opvallend kapsel), gezien de gemiddelde Nederlander en eerder onderzoek [20, 21]. Voor elk persoon worden twee afbeeldingen verzameld: één gepresenteerd tijdens de leerfase (in de KOPS-app) en één gepresenteerd tijdens de testfase (staande houden en fouilleren taak; herkenningstaak). Dit wordt gedaan om fotoherkenning te verminderen in plaats van gezichtsherkenning [36]. We hebben alleen neutrale gezichtsuitdrukkingen opnemen, omdat bekend is dat gelaatsuitdrukkingen de sociale perceptie beïnvloeden [e.g., smile leniency effect 37, 38] en een bemiddelende rol spelen bij gezichtsherkenning [39, 40].

STAANDE HOUDEN EN FOUILLEREN TAAK. Tijdens deze taak werden de deelnemers in een willekeurige volgorde geconfronteerd met doelgezichten (d.w.z. vanuit de app) en afleidende gezichten. Voor elke gepresenteerde persoon moesten de deelnemers aangeven of ze staande houden en fouilleren willen uitvoeren (ja/nee antwoord). Als ze ervoor kiezen om de persoon te fouilleren, moesten ze ook de reden van hun beslissing uitleggen. Hoewel staande houden en fouilleren vaak gebeurt op basis van verdacht gedrag naast uiterlijk, worden gedragskenmerken buiten beschouwing gelaten, om verwarrend gedrag door uiterlijk te voorkomen. Om de frequentie van staande houden en fouilleren te berekenen gebruiken we het aandeel ja-antwoorden per gezichtsgroep. De reacties van de deelnemers werden ook worden geanalyseerd in termen van juiste antwoorden en valse positieven (zie Gegevensanalyseschema).

GEZICHTSHERKENNINGSTAAK. Dit is een standaard oud/nieuw herkenningsparadigma. Het doel is om het werkelijke herkenningsvermogen van deelnemers te meten. Zo kunnen we bepalen of de beslissing om een persoon staande te houden het gevolg is van een verkeerde identificatie (samenvoeging van identiteiten). Deelnemers zien alle gezichten van de staande houden en fouilleren taak in een willekeurige volgorde. Bij elk gezicht dat wordt

getoond moeten deelnemers een ja/nee antwoord geven op de vraag "Herkent u deze persoon?" De reacties van de deelnemers worden geanalyseerd in termen van juiste antwoorden (doelen) en valse positieven (folies).

TYPISCHE KENMERKEN VAN EEN CRIMINEEL. Vergelijkbaar met Smalarz et al. [18] hebben we een aangepaste versie van Katz en Braly's (1933) adjectieve checklistmethode gebruiken om te onderzoeken welke persoonlijke en demografische kenmerken het sterkst geassocieerd worden met 20 verschillende misdrijven. De deelnemers krijgen korte karakterschetsen voor 10 misdrijven (bijv. autodiefstal; het dragen van een verborgen wapen; kindermishandeling; drugshandel). Elke karakterschets van een misdrijf wordt gevolgd door een lijst kenmerken die de deelnemers moeten classificeren. Deze kenmerken omvatten onder andere geslacht; leeftijd; etniciteit; opleiding; en sociale klasse. Tot slot hebben we onderzoeken welke kenmerken het vaakst worden geassocieerd met elk misdrijf (d.w.z. de inhoud van criminele stereotypering), evenals de mate van overeenstemming over die kenmerken (d.w.z. de consensus in criminele stereotypering).

PROCEDURE

Het testen kan worden onderverdeeld in drie afzonderlijke fasen. Deelnemers werd verteld dat het onderzoek gaat over gezichtsherkenning, maar waren blind voor het specifieke doel.

FASE 1 (VOOR BLOOTSTELLING). Na ondertekening van de geïnformeerde toestemming werd de deelnemers gevraagd naar hun demografische informatie, een staande houden en fouilleren taak uit te voeren en de typische kenmerken van een crimineel te beoordelen. De volgorde van de staande houden en fouilleren taak en de beoordeling waren gebalanceerd om volgorde-effecten te voorkomen. Deze maatregelen zijn nodig om de nulmeting van deelnemers vast te stellen, voordat ze via de app worden blootgesteld aan gezichten.

FASE 2 (LEREN). Vervolgens werden de deelnemers geïnstrueerd om de KOPS-app dagelijks te gebruiken, voor minimaal 5 min. per dag, gedurende een week (leerfase).

FASE 3 (NA BLOOTSTELLING). De dag na de laatste KOPS-app sessie werden de deelnemers gevraagd om een gezichtsherkenningstaak uit te voeren, een tweede staande houden en fouilleren taak en de typische kenmerken van een crimineel voor een tweede keer te beoordelen. Deelnemers hebben altijd beginnen met de herkenningstaak om niet in de war te raken en aangeven dat ze een gezicht herkennen van de staande houden en fouilleren taak in plaats van de KOPS-app. Tot slot werden de deelnemers bedankt en debrieft.

RESULTATEN

Er zijn verschillen in de antwoorden van politieagenten en universiteitsstudenten. ($p < .001$), hetgeen kan wijzen op een gebrek aan opleiding voor de laatste groep. Daarom zullen wij de bevindingen voor elke groep afzonderlijk rapporteren, met de nadruk op politieagenten.

TABEL 1 GEMIDDELDE PERCENTAGES EN STANDAARDAFWIJKINGEN VOOR ELKE TAAK IN ELKE TESTFASE

				Politieagenten								Studenten							
				Fase 1				Fase 2				Fase 1				Fase 2			
				Doelen		Folies		Doelen		Folies		Doelen		Folies		Doelen		Folies	
				M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Staande houden en fouilleren (stopfrequentie)																			
			Kaukasisch	.26	.29	.24	.26	.58	.22	.08	.2	.54	.16	.49	.21	.20	.24	.24	.23
			Noord-Afrikaanse	.30	.26	.27	.27	.52	.21	.20	.19	.68	.12	.71	.05	.21	.25	.15	.16
Staande houden en fouilleren (stopnauwkeurigheid)																			
			Kaukasisch					.57	.22	.91	.23					.21	.24	.76	.23
			Noord-Afrikaanse					.52	.21	.79	.19					.22	.25	.84	.17
Herkenning percentage																			
			Kaukasisch					.90	.14	.09	.13					.16	.25	.20	.31
			Noord-Afrikaanse					.79	.15	.17	.10					.24	.33	.17	.25
Herkenningsnauwkeurigheid																			
			Kaukasisch					.90	.14	.90	.13					.24	.33	.80	.31
			Noord-Afrikaanse					.79	.15	.82	.12					.16	.25	.81	.27

Notitie. **Stopfrequentie** = hoe vaak deelnemers aangaven een persoon te willen stoppen; **Stopnauwkeurigheid** = hoe vaak de beslissing om een persoon al dan niet aan te houden juist is; **Herkenning percentage** = hoe vaak deelnemers aangaven dat ze iemand herkenden (positieve beslissing); **Herkenningsnauwkeurigheid** = hoe juist zijn de deelnemers in hun herkenningsbeslissingen

Eerst hebben wij onderzocht of er reeds bestaande verschillen zijn in het stoppercentage tussen de twee etniciteit (d.w.z. voor blootstelling). De kans dat politieagenten Kaukasisch en Noord-Afrikaanse burgers staande hielden en fouilleerden was even groot voordat zij de doelgezichten te zien kregen, al $t(66) < 1,43$, $ps > .126$ (zie Tabel 1, fase 1). Integendeel, de studentenpopulatie staande hielden en fouilleerden meer Noord-Afrikaanse dan Kaukasische burgers, al $t(26) > 5,01$, $ps > .001$.

Vervolgens werd gekeken of het stoppercentage veranderde na blootstelling en hoe nauwkeurig de agenten waren in hun besluit te stoppen. De kans dat agenten Kaukasische en Noord-Afrikaanse doelen/verdachten aanhouden en fouilleren was even groot, $t(33) = -1,43$, $p = .162$. Maar ze houden vaker Noord-Afrikaanse dan Kaukasische folies/onschuldigen aan, $t(1, 34) = 7.41$, $ps < .001$. Evenzo zijn agenten even nauwkeurig in hun beslissingen om Kaukasische en Noord-Afrikaanse doelen/verdachten aan te houden en fouilleren, $t(33) = -1,43$, $p = .162$. Maar ze zijn juist in het aanhouden van Kaukasische dan Noord-Afrikaanse folies/onschuldigen, $t(33) = 6.42$, $p < .001$. Een soortgelijk beeld ontstond voor de studentensteekproef. De kans dat zij Kaukasische en Noord-Afrikaanse doelen/verdachten staande hielden en fouilleerden was even groot, $t(26) = 0.18$, $p = .858$, maar zij hielden vaker Noord-Afrikaanse dan Kaukasische folies/onschuldigen staande, $t(26) = -2,52$, $p = .018$. Bij de folies waren de studenten echter nauwkeuriger in hun beslissing om Noord-Afrikaanse dan Kaukasische aan te houden en te fouilleren, $t(26) = 2.27$, $p = .032$. Dit weerspiegelt een zeer conservatief stopgedrag na blootstelling.

Ten slotte, hebben we gekeken naar hoe vaak politieagenten aangaven dat ze iemand herkenden (positieve beslissing) en naar de nauwkeurigheid van hun beslissingen. Agenten maakten vaker een positieve identificatie (zeiden dat ze herkenden) bij het zien van Kaukasische dan Noord-Afrikaanse doelen/verdachten. En ze hebben meer kans op een positieve (dus fout) identificatie bij het zien van Noord-Afrikaanse dan Kaukasische folies/onschuldigen, al $t(33) > 3.47$, $ps < .001$. Politieagenten maakten meer juiste identificaties ("ja ik ken deze persoon" voor doelen/"nee ik ken deze persoon niet" voor folies) bij het zien van Kaukasische dan Noord-Afrikaanse doelwitten/verdachten en folies/onschuldigen, al $t(33) > -3.58$, $ps < .001$. In het algemeen, waren de studenten beter in het herkennen van Kaukasische dan Noord-Afrikaanse doelen/verdachten, $t(26) = -3.02$, $p = .006$, maar ze waren even goed in het herkennen van Kaukasische doelen als Noord-Afrikaanse folies/onschuldigen, al $t(26) < .823$, $p > .418$.

TYPISCHE KENMERKEN VAN EEN CRIMINEEL

Om te bepalen of herhaalde blootstelling een effect had op het stereotype denken over criminelen, vergeleken we de beoordelingen van de deelnemers van de stereotiepe kenmerken van elk misdrijf voor en na de blootstellingsfase. Tabel 2 toont de top vijf van meest onderschreven kenmerken voor elk misdrijf, gerangschikt van meest naar minst onderschreven voor en na blootstelling. Het geslacht, het ras en de leeftijd van de dader werden op beide tijdstippen met weinig of geen verandering geassocieerd met alle zestien misdrijven. Hieruit blijkt dat stereotiepe associaties

met bepaalde kenmerken en criminaliteit van zeer sterke aard zijn en dus niet gemakkelijk ter discussie worden gesteld.

TABEL 2 LIJST VAN DE 5 BELANGRIJKSTE KENMERKEN DIE VOOR ELK MISDRIJF ZIJN GERAPPORTEERD, VOOR EN NA DE HERHAALDE BLOOTSTELLING AAN DE DOELGEZICHTEN.

Misdaden en kenmerken	Voor (%)	Na (%)
Autodiefstal		
Man	83.6	82.1
Nederlandse	60.0	21.4
18-29 jaar	58.4	60.7
Arme/Lage klasse	29.8	17.8
Werkloos	16.4	14.3
Bankoverval		
Man	77.6	89.3
18-29 jaar	55.3	42.8
Marokkaanse	22.3	-
Arme/Lage klasse	19.4	25.0
30-50 jaar	11.9	25.0
Nederlandse	-	14.3
Dragen van verborgen wapen		
Man	79.2	82.1
Marokkaanse	60.0	21.3
18-29 jaar	52.2	67.8
Nederlandse	18.1	14.3
Moslim	15.0	-
Curaçaose	-	14.3
Kindermisbruik		
Man	85.1	85.7
30-50 jaar	56.7	53.6
Nederlandse	45.1	28.6
50-70 jaar	18.0	7.2
Getrouwd	6.6	-
Alleen leven	-	7.2
Ontvoering onder dwang		
Man	65.7	64.3
Gescheiden	47.8	42.8
30-50 jaar	43.3	53.6
18-29 jaar	28.3	25.0
Nederlandse	19.4	28.6
Seksuele handelingen		
Man	86.6	89.3
18-29 jaar	38.8	28.5
30-50 jaar	32.8	39.3
Nederlandse	23.9	46.4
Krankzinnig/mentaal instabiel	18.0	28.5
Rijden schieten		
Man	79.1	89.3
18-29 jaar	58.2	60.7
Marokkaanse	23.9	14.3
Curaçaoan	16.4	-

30-50 jaar	9.0	7.1
Arme/Lage klasse		14.3
Drugshandel		
Man	77.6	75.0
18-29 jaar	58.2	67.9
Marokkaanse	32.9	42.9
Arme/Lage klasse	20.9	10.6
Werkloos	13.1	21.5
Identiteitsdiefstal		
Man	74.6	82.1
18-29 jaar	38.8	50.0
Nederlandse	25.4	25.0
30-50 jaar	28.4	17.9
Gemiddelde intelligentie	11.9	14.4
Terrorisme		
Man	73.1	75.0
18-29 jaar	55.2	53.6
Moslim	25.3	35.7
Marokkaanse	20.8	28.6
Andere etniciteit	19.6	17.9

Notitie. - geeft aan dat deze functie op dat moment niet tot de top 5 behoorde.

Tabel 3 geeft een overzicht van de consensusscores voor alle in de studie gebruikte misdrijven. In het algemeen was er zowel voor als na de blootstelling sprake van een hoge consensus over de meeste misdrijven. In de meeste gevallen waren er minder dan drie kenmerken nodig om het consensuscriterium te bereiken. Voorts waren er geen grote veranderingen in de consensusscores, behalve voor de misdrijven dragen van een verborgen wapen en ontvoering. Hieruit blijkt dat er een grote mate van overeenstemming tussen de ambtenaren bestond over welke van de hoofdkenmerken het meest waarschijnlijk met de misdrijven in verband werden gebracht.

TABEL 3 CONSENSUSCIJFERS VOOR TYPISCHE DADERKENMERKEN VAN EEN CRIMINEEL VOOR EN NA HERHAALDE BLOOTSTELLING AAN DE DOELGEZICHTEN

Misdaad	Consensus voor	Consensus na	Vershil in consensuspunten voor en na
Kindermisbruik	2,18	2,15	-0,03
Drugshandel	2,43	2,17	-0,26
Rijden schieten	2,53	2,00	-0,53
Terrorisme	2,74	2,60	-0,14
Seksuele handelingen	2,75	2,36	-0,39
Autodiefstal	2,76	2,25	-0,51
Bankoverval	2,88	3,43	+0,55
Dragen van verborgen wapen	3,05	1,83	-1,22
Identiteitsdiefstal	3,05	2,71	-0,34
Ontvoering onder dwang	4,35	2,75	-1,60

STUDIE 2 EN HAALBAARHEID BEOORDELING

Het doel van Studie 2 was te onderzoeken of er manieren zijn om het risico op stereotypering te verminderen door de manier waarop gezichten worden gepresenteerd te veranderen, bijvoorbeeld door individuaties, sequenties en instructies. Aangezien louter blootstelling zonder individuaties de nauwkeurigheid van gezichtsherkenning niet verbetert [16, 17], wilden we onderzoeken of instructies of individuatietraining de prestaties zouden verbeteren.

Terwijl Studie 1 liep, werkten een team van onderzoeksassistenten en ik aan het opzetten van Experiment 2. Dat wil zeggen, na het verkrijgen van ethische goedkeuring hebben we verschillende individuatiebenaderingen getest (bv. met namen, kenmerken, kleuren) en op basis van de resultaten hebben we Experiment 2 geprogrammeerd. Vervolgens hebben wij, na advies en aanbevelingen van P&W, de instructies aangepast om ze nog politievriendelijker te maken.

Hoewel we klaar waren met Studie 2, kwam de moeilijkheid om politieagenten te vinden zoals de deelnemers in het voorstel hadden voorspeld, uit. Na uitputting van de contacten van het iLab in de eenheden Limburg en Oost-Brabant, en in overeenstemming met ons risicobeheersplan (zie aanvraag), wendden wij ons tot de werving van deelnemers van de politieacademie, reeds voor studie 1. De respons was echter gering en de deelnemers werden niet gevonden. De respons was echter laag en verscheidene deelnemers die aan het experiment begonnen, volgden de instructies niet en/of voltooiden het onderzoek niet, ondanks herhaalde herinneringen.

Gezien de moeilijkheden met de werving in Studie 1, en opnieuw in overeenstemming met ons voorstel inzake risicobeheer (zie voorstel), stelden wij toen voor de studie uit te voeren met een studentensteekproef. Dit voorstel kreeg geen bijval wegens bezorgdheid over de ecologische validiteit. Daarom werd, ondanks het feit dat Studie 2 klaar was, gezamenlijk besloten niet door te gaan met Studie 2 en het project te beëindigen.

CONCLUSIE

Al met al suggereren onze bevindingen dat politieagenten, die bijna uitsluitend Kaukasisch waren, minder accuraat waren in hun beslissingen om Noord-Afrikaanse gezichten aan te houden en te herkennen dan Kaukasische gezichten. De bevindingen komen in grote lijnen overeen met de literatuur over raciale vooroordelen en lijken niet het resultaat te zijn van willekeurig richten, als gevolg van herhaalde blootstelling aan doelgezichten.

De resultaten zijn verontrustender als we kijken naar de typische kenmerken van een crimineel. Etniciteit stond in de top 5 van kenmerken voor alle geteste misdrijven, wat erop wijst dat het voor agenten een topkwalificatie blijft en hun denken over misdaad weerspiegelt. Daarom moeten de huidige maatregelen om raciale vooroordelen te verminderen en de aandacht voor etniciteit weg te nemen, worden geïntensiveerd.

Ondanks de beperkingen van dit project geloven wij dat de huidige bevindingen informatief zijn voor beleidsmakers. Ten eerste wijzen de bevindingen op de noodzaak van meer aandacht voor en onderzoek naar de kwestie van ras binnen de politie. Ten tweede, en in het licht van bredere onderzoek ontwikkelingen in de literatuur over gezichtsherkenning, suggereren ze dat het de moeite waard is om interventies te onderzoeken die gericht zijn op het verbeteren van de nauwkeurigheid van herkenning voor patrouilles.

LITERATURE

1. Scott, L.S., et al., *A reevaluation of the electrophysiological correlates of expert object processing*. Journal of Cognitive Neuroscience, 2006. **18**(9): p. 1453-65.
2. Yovel, G., et al., *Can massive but passive exposure to faces contribute to face recognition abilities?*. Journal of Experimental Psychology: Human Perception and Performance, 2012. **38**: p. 285.
3. Abudarham, N. and G. Yovel, *Same critical features are used for identification of familiarized and unfamiliar faces*. Vision Research, 2019(157): p. 105-111.
4. Maurer, D., R. Le Grand, and C.J. Mondloch, *The many faces of configural processing*. Trends in Cognitive Sciences, 2002: p. 255-260.
5. O'Donnell, C. and V. Bruce, *Familiarisation with faces selectively enhances sensitivity to changes made to the eyes*. Perception, 2001. **30**(6): p. 755-764.
6. Ellis, H.D., J.W. Shepherd, and G.M. Davies, *Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition*. Perception 1979. **8**(4): p. 431-9.
7. Young, A.W., et al., *Matching familiar and unfamiliar faces on internal and external features*. Perception, 1985: p. 737-746.
8. Meissner, C.A. and J.C. Brigham, *Thirty years of investigating the own-race bias in memory for faces - A meta-analytic review*. Psychology Public Policy and Law, 2001. **7**(1): p. 3-35.
9. Sporer, S.L., *Recognizing faces of other ethnic groups - An integration of theories*. Psychology Public Policy and Law, 2001. **7**(1): p. 36-97.
10. Wright, D.B., C.E. Boyd, and C.G. Tredoux, *Inter-racial contact and the own-race bias for face recognition in South Africa and England*. Applied Cognitive Psychology, 2003. **17**(3): p. 365-373.
11. Rhodes, G., et al., *Race coding and the other-race effect in face recognition*. Perception, 2009. **38**(2): p. 232-241.
12. Richler, J.J. and I. Gauthier, *A meta-analysis and review of holistic face processing*. Psychological Bulletin, 2014. **140**(5): p. 1281-302.
13. Hughes, B.L., et al., *Neural adaptation to faces reveals racial outgroup homogeneity effects in early perception*. Proceedings of the National Academy of Sciences, 2019. **116**(29): p. 14532-14537.
14. Rhodes, M.G. and J.S. Anastasi, *The own-age bias in face recognition: a meta-analytic and theoretical review*. Psychological Bulletin, 2012. **138**(1): p. 146-74.
15. Proietti, V., et al., *Attending to identity cues reduces the own-age but not the own-race recognition advantage*. Vision Res, 2019. **157**: p. 184-191.
16. McGugin, R.W., et al., *Race-specific perceptual discrimination improvement following short individuation training with faces*. Cognitive Science, 2011. **35**(2): p. 330-347.
17. Tanaka, J.W. and L.J. Pierce, *The neural plasticity of other-race face recognition*. Cognitive, Affective, & Behavioral Neuroscience, 2009. **9**(1): p. 122-131.
18. Smalarz, L., et al., *The perfect match: Do criminal stereotypes bias forensic evidence analysis?* Law and Human Behavior, 2016. **40**(4): p. 420-9.
19. Skorinko, J.L. and B.A. Spellman, *Stereotypic Crimes: How Group-Crime Associations Affect Memory and (Sometimes) Verdicts and Sentencing*. Victims & Offenders, 2013. **8**(3): p. 278-307.
20. MacLin, M.K. and V. Herrera, *The criminal stereotype*. North American Journal of Psychology 2006: p. 197-208.
21. Johnson, B.D. and R.D. King, *Facial Profiling: Race, Physical Appearance, and Punishment*. Criminology, 2017. **55**(3): p. 520-547.
22. Smalarz, L., S. Madon, and A. Turosak, *Defendant stereotypicality moderates the effect of confession evidence on judgments of guilt*. Law and Human Behavior, 2018. **42**(4): p. 355-368.
23. Zebrowitz, L.A., B. White, and K. Wieneke, *Mere exposure and racial prejudice: Exposure to other-race faces increases liking for strangers of that race*. Social Cognition, 2008. **26**(3): p. 259-275.
24. Smith, P.K., A. Dijksterhuis, and S. Chaiken, *Subliminal exposure to faces and racial attitudes: Exposure to Whites makes Whites like Blacks less*. Journal of Experimental Social Psychology, 2008. **44**(1): p. 50-64.
25. Montoya, R.M., et al., *A re-examination of the mere exposure effect: The influence of repeated exposure on recognition, familiarity, and liking*. Psychological Bulletin, 2017. **143**(5): p. 459-498.

26. Kunst-Wilson, W.R. and R.B. Zajonc, *Affective discrimination of stimuli that cannot be recognized*. Science, 1980. **207**(4430): p. 557-558.
27. Bornstein, R.F., D.R. Leone, and D.J. Galley, *The generalizability of subliminal mere exposure effects: Influence of stimuli perceived without awareness on social behavior*. Journal of Personality and Social Psychology, 1987. **53**(6): p. 1070.
28. Chiroro, P. and T. Valentine, *An investigation of the contact hypothesis of the own-race bias in face recognition*. The Quarterly Journal of Experimental Psychology, 1995. **48**(4): p. 879-894.
29. Valentine, T., *A unified account of the effects of distinctiveness, inversion, and race in face recognition*. The Quarterly Journal of Experimental Psychology 1991. **43**(2): p. 161-204.
30. Paluck, E.L., S.A. Green, and D.P. Green, *The contact hypothesis re-evaluated*. Behavioural Public Policy, 2019: p. 129-158.
31. Davies, K., et al., *Cross-group friendships and intergroup attitudes: A meta-analytic review*. Personality and Social Psychology Review, 2011. **15**(4): p. 332-351.
32. *Proactief politieoptreden vormt een risico voor mensenrechten – Etnisch profileren onderkennen en aanpakken*. 2013, Amnesty International
33. Haenen, M., *Politie Rotterdam onderzoekt racisme in appgroep agenten*, in NRC Handelsblad. 2020.
34. Swets, J.A., *Signal detection and recognition in human observers: Contemporary readings*. 1964.
35. Erdfelder, E., F. Faul, and A. Buchner, *GPOWER: A general power analysis program*. Behavior Research Methods Instruments & Computers, 1996. **28**(1): p. 1-11.
36. Bruce, V., *Changing faces: visual and non-visual coding processes in face recognition*. British Journal of Psychology, 1982. **73**(Pt 1): p. 105-16.
37. LaFrance, M. and M.A. Hecht, *Why smiles generate leniency*. Personality and Social Psychology Bulletin, 1995. **21**(3): p. 207-214.
38. Flowe, H.D., *Do characteristics of faces that convey trustworthiness and dominance underlie perceptions of criminality?* PLoS One 2012. **7**(6): p. e37253.
39. Bruce, A.J. and B.G. McDonald, *Face recognition as a function of judgments of likability unlikability*. Journal of General Psychology, 1993. **120**(4): p. 451-462.
40. Lin, T., R. Lendry, and N.C. Ebner, *Face likeability mediates the memory-enhancing effect of face attractiveness in young but not older adults*. Memory, 2016: p. 1396-1406.
41. Bland, J.M. and D.G. Bland, *Statistics Notes: One and two sided tests of significance*. 1994: p. 248.
42. Kätysri, J., *Those virtual people all look the same to me: Computer-rendered faces elicit a higher false alarm rate than real human faces in a recognition memory task* Frontiers in Psychology, 2018. **9**: p. 1362.
43. Perrin, B.M., et al., *Information order and outcome framing: An assessment of judgment bias in a naturalistic decision-making context*. Human Factors, 2001. **43**(2): p. 227-238.
44. Fleming, J. (2010). Learning to work together: police and academics. *Policing: A Journal of Policy and Practice*, 4(2), 139-145.