

# Wanneer beïnvloedt statistisch leren sociale perceptie van gezichten?

## De rol van reeds bestaande informatieve cues

Anouk M. van de Klundert

*Studente Master Sociale en Organisationspsychologie, Universiteit Utrecht, studentnummer: 4116127*

Scriptiebegeleider: Ron Dotsch

Tweede beoordelaar: Reine van der Wal

Datum: 17 juli 2017

### Abstract

Uit voorgaand onderzoek is gebleken dat gezichtsbeoordelingen afhankelijk zijn van statistisch leren, waarbij mensen de verdeling van gezichten in hun omgeving aanleren. Huidige studie heeft onderzocht of statistisch leren ook sociale perceptie beïnvloedt in de aanwezigheid van informatieve cues voor dominantie en betrouwbaarheid. In het huidige experiment werden proefpersonen in een leerfase blootgesteld aan computer-gegenereerde gezichten afkomstig uit een dominantiedimensie of een betrouwbaarheidsdimensie. De gezichten volgden op deze dimensies een normaalverdeling, waarvan het gemiddelde tussen proefpersoon verschilde, met een gemiddelde z-score van -5 of +5 op de aangeboden dimensie. In een daaropvolgende testfase beoordeelden proefpersonen nieuwe gezichten op dominantie, betrouwbaarheid en gemiddeldheid. Verwacht werd dat de beoordelingen in de testfase zouden verschillen op basis van de aangeleerde verdelingen en dimensies in de leerfase. Hier is geen bewijs voor gevonden. De proefpersonen leken de verdelingen in de leerfase niet te hebben aangeleerd, waardoor er geen sprake was van statistisch leren.

*Sleutelwoorden:* gezichtsbeoordelingen, statistisch leren, dominantie, betrouwbaarheid, gemiddeldheid, gezichtsdimensies

Abstract

Previous research has found that face evaluations depend on statistical learning, in which people learn what average faces are based on their environments. The current study has investigated whether statistical learning also influences social perception when informative cues for dominance and trustworthiness are present. In the learning phase of the current experiment, participants were exposed to faces originating from a dominance dimension or a trustworthiness dimension of computer generated faces. The faces followed a normal distribution on these dimensions, of which the average differed between subjects. The average could be either an average z-score of -5 or +5 on the shown dimension. In the subsequent test phase, participants evaluated novel faces on dominance, trustworthiness, and typicality. The evaluations in the test phase were expected to differ based on the learned dimensions and the divisions in the learning phase. These effects were not found in the results of the experiment. The participants seemingly have not learned the divisions in the learning phase, which means there was no indication of statistical learning.

*Keywords:* face evaluation, statistical learning, dominance, trustworthiness, typicality, face dimensions

### Wanneer Beïnvloedt Statistisch Lerens Sociale Perceptie van Gezichten?

Mensen kunnen al een impressie vormen van een persoon op basis van een gezicht na een blootstelling van 34 milliseconden (Todorov, Pakrashi, & Oosterhof, 2009). In een dergelijke korte tijd vormen mensen al een mening over diverse karaktereigenschappen, zoals betrouwbaarheid of agressiviteit (Willis & Todorov, 2006). Deze mening kan van grote invloed zijn in praktijksituaties: hoe betrouwbaar iemands gezicht gevonden wordt, kan bijvoorbeeld de zwaarte van een straf in een rechtszaak beïnvloeden (Porter, ten Brinke, & Gustaw, 2010). Ook hebben individuen met onbetrouwbare gezichten een grotere kans om als misdadiger geselecteerd te worden uit een politie line-up en worden zij daarnaast ook eerder berecht (Flowe & Humphries, 2011). Desondanks zijn evaluaties van gezichten een ontoereikende manier om stabiele persoonlijkheidskenmerken vast te stellen (Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Mensen zijn vaak niet accuraat in hun oordelen over persoonlijkheid op basis van gezichten. De vaak ontorechte maar stevige gevolgen van sociale evaluaties van gezichten geven aan hoe belangrijk het is om te achterhalen hoe deze evaluaties tot stand komen.

Er is al veel onderzoek verricht naar welke specifieke gezichtskenmerken informatief zijn voor welke specifieke beoordeling, zoals bijvoorbeeld mannelijkheid/vrouwelijkheid voor de mate van dominantie, of glimlachen voor de mate van betrouwbaarheid (Boothroyd et al., 2007; Buckingham et al., 2006; Krumhuber et al., 2007; Oosterhof & Todorov, 2008). Hoewel er onder mensen een grote mate van overeenstemming is in de sociale evaluatie van gezichten (Cogsdill et al., 2014; Secord et al., 1954), is deze overeenstemming niet perfect. Mensen verschillen soms van mening bij het beoordelen van gezichten en de huidige modellen verklaren niet waarom dat zo is.

Zo is in nieuw onderzoek naar de rol van culturele invloed op gezichtsevaluaties gebleken dat mensen van verschillende culturen gezichten uit hun eigen cultuur betrouwbaarder vinden dan gezichten uit een andere cultuur (Sofer et al., 2017). Deze bevindingen geven dus aan dat evaluaties van gezichten systematisch kunnen verschillen, waarvoor een verklaring nodig is. Recent onderzoek geeft een mogelijke verklaring: het is mogelijk dat de gezichtskenmerken die mensen gebruiken voor hun oordeel deels universeel gedeeld zijn, maar ook deels afhankelijk zijn van iemands omgeving en cultuur (Dotsch, Hassin, & Todorov, 2016). De achterliggende gedachte hierbij is dat mensen door middel van statistisch leren tot zich nemen wat gemiddelde gezichten zijn in hun omgeving. Vervolgens vertrouwen ze gezichten die dichter bij dat gemiddelde liggen meer dan gezichten die daar verder vanaf liggen (Sofer et al., 2015).

## WANNEER BEÏNVLOEDT STATISTISCH LEREN SOCIALE PERCEPTIE VAN GEZICHTEN?

Deze mogelijke verklaring is empirisch getoetst door Dotsch, Hassin, & Todorov (2016). Proefpersonen kregen 500 gezichten te zien die systematisch verschilden op specifieke gezichtskenmerken. Ze volgden op die gezichtskenmerken een normaalverdeling, waarbij de positie van de piek van de verdeling (i.e., het gemiddelde) tussen proefpersonen gemanipuleerd werd. Gezichten werden inderdaad positiever beoordeeld naarmate ze dichter bij het aangeleerde midden van de aangeboden verdeling kwamen (Dotsch, Hassin & Todorov, 2016). Deze resultaten ondersteunen het idee dat variatie in de structuur van iemands omgeving kan leiden tot variatie in de evaluaties van gezichten en verklaart mogelijk waarom er individuele en culturele verschillen in sociale evaluaties bestaan.

In het onderzoek van Dotsch en collega's (2016) zijn de gebruikte gezichten gegenereerd door middel van een statistisch 'face space' model, wat de variantie van een grote hoeveelheid echte gezichten met 130 orthogonale (statistisch onafhankelijke) dimensies samenvat, waaronder 50 symmetrische vormdimensies, 30 asymmetrische vormdimensies en 50 textuurdimensies. Een dimensie beschrijft de variantie van een bepaalde set covariërende gezichtskenmerken, waarbij de score op die dimensie aangeeft hoe sterk de eigenschap aanwezig is in het gezicht dat op dat moment door het model gerepresenteerd wordt. De score is een z-score die aangeeft hoeveel standaardafwijkingen het gezicht voor die gezichtskenmerken afwijkt van het gemiddelde, het midden van de dimensie en de ruimte. In een dergelijke gezichtenruimte is elk mogelijk gezicht te representeren als een coördinaat in die ruimte en elk coördinaat is door de computer weer te visualiseren als gezicht. Het model maakt het mogelijk om met behulp van wiskunde uiterst precieze en gecontroleerde manipulaties van gezichten te genereren.

Om het statistisch leren effect op sociale evaluatie van gezichten te isoleren zonder de invloed van kenmerken waarvan bekend is dat ze evaluatie beïnvloeden, creëerden Dotsch en collega's (2016) nieuwe random dimensies. Dit werden gewogen combinaties van bestaande dimensies in de gezichtenruimte, die zijn gecreëerd door per random dimensie 50 willekeurige waarden te genereren voor de vormdimensies. Deze werden georthogonaliseerd (i.e., statistisch onafhankelijk gemaakt) aan tien eerder geïdentificeerde sociale dimensies, zoals betrouwbaarheid, dominantie, competentie, aantrekkelijkheid, dreiging en extraversie (Oosterhof & Todorov, 2008), zodat verandering in de nieuwe random dimensie niet resulteerde in verandering in de sociale dimensies. Er werden negen kandidaatdimensies geselecteerd die zwak met elkaar correleerden, om de variantie in de uiteindelijke set van dimensies te maximaliseren. Gezichten variërend op deze negen kandidaatdimensies werden beoordeeld door 30 studenten op betrouwbaarheid, sociale betekenis (social meaning),

## WANNEER BEÏNVLOEDT STATISTISCH LEREN SOCIALE PERCEPTIE VAN GEZICHTEN?

voorspelbaarheid en gemiddeldheid (typicality) op een zevenpuntsschaal. Uiteindelijk werd de kandidaatdimensie met de laagste betrouwbaarheidsvariantie geselecteerd als experimentele dimensie voor het verdere onderzoek.

Gezichten in de leerfase van het onderzoek naar statistisch leren van Dotsch en collega's (2016) varieerden dus systematisch volgens een normaalverdeling op deze ene experimentele dimensie. Hoewel de experimentele dimensie georthogonaliseerd was ten opzichte van andere sociale dimensies, inclusief dominantie, volgden oordelen van dominantie een opvallend patroon. Waar andere oordelen zoals betrouwbaarheid piekten rond het aangeleerde gemiddelde gezicht op de experimentele dimensie en afwijkingen negatiever beoordeeld werden, werden gezichten naarmate ze hoger scoorden op de experimentele dimensie steeds dominanter beoordeeld. Het is mogelijk dat er ondanks de orthogonalisering toch dominantie-informatie aanwezig was in de doeldimensie.

Als het inderdaad klopt dat in de aanwezigheid van informatieve cues over een oordeel het statistisch leren effect wegvalt voor dat specifieke oordeel, dan is de generaliseerbaarheid van de resultaten beschreven door Dotsch et al. (2016) ernstig in het geding. Het zou betekenen dat het statistisch leren effect alleen optreedt in zeer uitzonderlijke en gecontroleerde situaties waarin er geen andere informatie beschikbaar is in het gezicht om de evaluatie te informeren. Het huidige onderzoek richt zich op de vraag of statistisch leren ook sociale perceptie beïnvloedt in de aanwezigheid van informatieve cues, zoals cues die samenhangen met betrouwbaarheid of dominantie. Betrouwbaarheid en dominantie zijn twee van de belangrijkste inferenties die mensen maken, en verklaren een groot deel van de variantie in gezichtsbeoordelingen (Oosterhof & Todorov, 2008; Sutherland et al., 2013). Wanneer er informatie aanwezig is die betrekking heeft op de specifieke gezichtsbeoordeling van mensen, dan is het mogelijk dat zij hun oordeel vormen op basis van deze informatie in plaats van de aangeleerde statistische informatie. Dit onderzoek zal de invloed van overgebleven informatie testen bij dimensies van dominantie en betrouwbaarheid. Het doel hierbij is dat proefpersonen aanleren wat een gemiddeld gezicht is door middel van blootstelling aan gezichten die een normaalverdeling volgen op een dominantiedimensie of een betrouwbaarheidsdimensie. Tussen proefpersonen wordt gemanipuleerd welke dimensie wordt aangeboden, de dominantie dimensie of de betrouwbaarheidsdimensie. Daarnaast wordt ook gemanipuleerd waar het gemiddelde van die distributie is. Voor de ene conditie ligt dit gemiddelde bij een z-score van -5, wat inhoudt dat er gezichten worden getoond die vijf standaarddeviaties afliggen van het midden van de gezichtenruimte. Voor de andere conditie ligt dit gemiddelde bij een z-score van +5. Wanneer op een later moment getest wordt hoe de

## WANNEER BEÏNVLOEDT STATISTISCH LEREN SOCIALE PERCEPTIE VAN GEZICHTEN?

proefpersonen nieuwe gezichten die variëren op dominantie en betrouwbaarheid beoordelen op dominantie en betrouwbaarheid, wordt verwacht dat ze gebruik maken van de eerder aangeleerde statistische informatie voor beoordelingen die niet geïnformeerd kunnen worden door de variërende dimensie. Dit houdt in dat er verwacht wordt dat de dominantie-oordelen van gezichten variërend op dominantie en de betrouwbaarheidsoordelen van gezichten variërend op betrouwbaarheid niet beïnvloed worden door statistisch leren, omdat er diagnostische kenmerken voor het oordeel in de gezichten aanwezig zijn. In deze gevallen wordt een lineaire samenhang tussen het oordeel en de positie van het gezicht op de dimensie verwacht. Hoe hoger de z-score is, des te extremer zal het oordeel op die dimensie zijn. Verwacht wordt dat de dominantie-oordelen van gezichten variërend op betrouwbaarheid, en de betrouwbaarheidsoordelen van gezichten variërend op dominantie, daarentegen wel beïnvloed worden door het statistisch leren. Wanneer proefpersonen in de leerfase blootgesteld zijn aan variatie die zichtbaar is in de gezichten in de testfase, zal het oordeel pieken rond het gemiddelde uit de leerfase. Dit wordt zichtbaar als een curvilineaire relatie tussen de positie van het gezicht op de geleerde dimensie en het oordeel op die dimensie. Wanneer proefpersonen in de leerfase blootgesteld zijn aan variatie die niet zichtbaar is in de gezichten in de testfase, wordt op basis van Todorov et al. (2015) verwacht dat er nog steeds een curvilineaire relatie tussen oordeel en positie op de dimensie bestaat, maar met een piek rond een z-score van 0. Omdat betrouwbaarheid en dominantie licht negatief correleren is het tevens mogelijk dat een deel van de lineaire relatie tussen oordeel en positie van gezicht op een dimensie ook in de laatste twee gevallen zichtbaar wordt.

Bij de toegevoegde manipulatiecheck, waarbij gemiddeldheid wordt gemeten, wordt hetzelfde curvilineaire effect verwacht voor alle condities. De piek zal rond het gemanipuleerde gemiddelde liggen bij beoordelingen van de gezichten die variëren op de aangeleerde dimensie. De piek zal dan rond 0 liggen bij beoordelingen van de gezichten die variëren op de andere dimensie.

## **Methode**

### **Design**

De studie omvatte een mixed design. Er werd gebruik gemaakt van twee verschillende dimensies; dominantie en betrouwbaarheid. Per dimensie werd gemanipuleerd wat het theoretische gemiddelde was van de normaalverdeling die de gezichten in de leerfase op een van de twee dimensies volgden: een gemiddelde z-score van -5 of +5. Daarnaast werden in de testfase gezichten beoordeeld die binnen proefpersonen varieerden op dominantie of

betrouwbaarheid. Dit maakte deze studie tot een  $2 \times 2 \times 2$  design (Doeldimensie [dominantie, betrouwbaarheid]  $\times$  Z-scores [-5, +5]  $\times$  Testdimensie [dominantie, betrouwbaarheid]). Hierbij varieerden de eerste twee variabelen tussen proefpersonen en de laatste variabele binnen proefpersonen. De afhankelijke variabelen waren de gerapporteerde dominantie, betrouwbaarheid en gemiddeldheid in de testfase.

### **Proefpersonen**

Om vast te stellen hoeveel proefpersonen nodig waren voor het onderzoek is een poweranalyse uitgevoerd. Bij het uitvoeren is een ANOVA aangehouden met herhaalde metingen, en within en between subjects interactie. Uit de analyse bleek dat een steekproefgrootte van 128 proefpersonen nodig was om een effect van tenminste  $d = 0.15$  te kunnen detecteren met 80% power.

De proefpersonen zijn geworven via de website Prolific ([www.prolific.ac](http://www.prolific.ac)), waar ze een vergoeding van £2 ontvingen na het voltooien van het experiment. Mensen met een goede Engelse taalvaardigheid en die niet kleurenblind zijn konden zich aanmelden om deel te nemen.

### **Stimuli**

Voor deze studie zijn 116 gezichten gegenereerd per experimentele conditie. De gezichten zijn gegenereerd zoals in Dotsch et al. (2016). Hiervoor werd FaceGen 3.1 gebruikt, een multidimensionaal computergegenereerde 'face space'. Deze face space bevat 50 symmetrische vormdimensies, 30 asymmetrische vormdimensies en 50 symmetrische textuurdimensies.

De gegenereerde gezichten waren driedimensionaal en verschilden op zowel de dimensies als op willekeurige manieren, orthogonaal ten opzichte van de doeldimensie. De scores van de dimensie volgden een normaal verdeelde distributie met  $\mu = 0$  s.d. en  $\sigma = 3$  s.d. Deze zijn omgezet naar gemiddelde z-scores van -5 of +5, variërend tussen proefpersonen.

Ook werden silhouetafbeeldingen gecreëerd voor de leerfase. Elk gezicht werd 90 graden geroteerd rondom zijn verticale as om een zijaanzicht te creëren. Alle kleurinformatie werd vervangen door grijswaarden. Voor de testfase is gebruik gemaakt van drie blokken van 21 gezichten die zijn gegenereerd met variërende z-scores van -5 tot +5 op de doeldimensie, met gelijke intervallen van een z-score van 0.5. De stimuli in de testfase varieerden allen als functie van de positie op de doeldimensie, wat alleen invloed had op de vorm van het gezicht.

## **Procedure**

Proefpersonen werden toegewezen aan een van de experimentele condities, zonder de proefpersoon te informeren over het bestaan van de verschillende condities. Iedere proefpersoon kreeg 116 gezichten te zien in de leerfase van het experiment. Om de proefpersonen te motiveren de aandacht bij de taak te blijven houden, werd ze bij één op de tien willekeurige trials gevraagd om aan te geven of een getoond zijaanzicht van een gezicht overeen kwam met het gezicht dat voorafgaand werd getoond. Bij de helft van deze zijaanzichten kwam dit silhouet overeen met het voorafgaande gezicht. Aan de proefpersonen werd verteld dat deze taak test hoe goed ze een gezicht van de zijkant kunnen herkennen wanneer ze het gezicht eerst van voor hebben gezien. Deze gegevens zijn vervolgens niet geanalyseerd.

In de daaropvolgende testfase beoordeelden de proefpersonen gezichten op dominantie, betrouwbaarheid en gemiddeldheid op een zevenpuntsschaal. Deze aspecten werden aangeboden in drie blokken van 21 gezichten per blok. Proefpersonen beoordeelden de gezichten in de testfase op gemiddeldheid als manipulatiecheck. Dit blok was altijd het laatste blok.

## **Dataverwerking en analyses**

Om de verzamelde data te kunnen gebruiken voor analyses, zijn eerst de pieken per proefpersoon geanalyseerd. Dit houdt in dat er geschat werd waar de piek ligt van de beoordelingen van een proefpersoon op gezichten uit een bepaalde dimensie. Voor het schatten van de pieken is gebruik gemaakt van de ‘local polynomial regression fitting approach’ (LOESS), door middel van de loess functie in R. Bij het uitvoeren van verdere analyses zijn de geschatte pieken per proefpersoon gebruikt als data voor gezichtsbeoordelingen. Voor 4 proefpersonen werd er meer dan één enkele piek gedetecteerd. Doordat de data van deze proefpersonen geen variantie in respons liet zien, kon er geen piek geschat worden. Deze proefpersonen zijn niet meegenomen in de verdere analyses, omdat er voor hen geen werkbare data beschikbaar is. Dit brengt het proefpersonen aantal terug naar 124 personen.

Voor de manipulatiecheck, waarbij gemiddeldheid werd gemeten, is een  $2 \times 2$  between subjects ANOVA uitgevoerd, waarbij de twee gemanipuleerde dimensies zijn meegenomen en de twee condities die verschillen in gemanipuleerde gemiddelde z-score. De gemanipuleerde dimensies waren de doeldimensies van dominantie en betrouwbaarheid waar proefpersonen in de leerfase van het experiment aan werden blootgesteld. Het andere



## WANNEER BEÏNVLOEDT STATISTISCH LEREN SOCIALE PERCEPTIE VAN GEZICHTEN?

onderscheid dat werd gemaakt is tussen de gemiddelde z-scores van -5 en +5 in de condities van de leerfase. Deze analyse is alleen uitgevoerd over beoordelingen van gemiddeldheid waarbij de beoordeelde gezichten afkomstig waren uit dezelfde dimensie als de aangeboden dimensie in de leerfase. Dit houdt in dat de beoordelingen zijn meegenomen van gezichten uit de dominantiedimensie door proefpersonen die de dominantiedimensie aangeboden hebben gekregen in de leerfase, aangezien hier een effect wordt verwacht. Hetzelfde geldt voor de betrouwbaarheidsdimensie.

Vervolgens is er voor de overige factoren een  $2 \times 2 \times 2 \times 2$  repeated measures ANOVA uitgevoerd. Hierin zijn meegenomen:

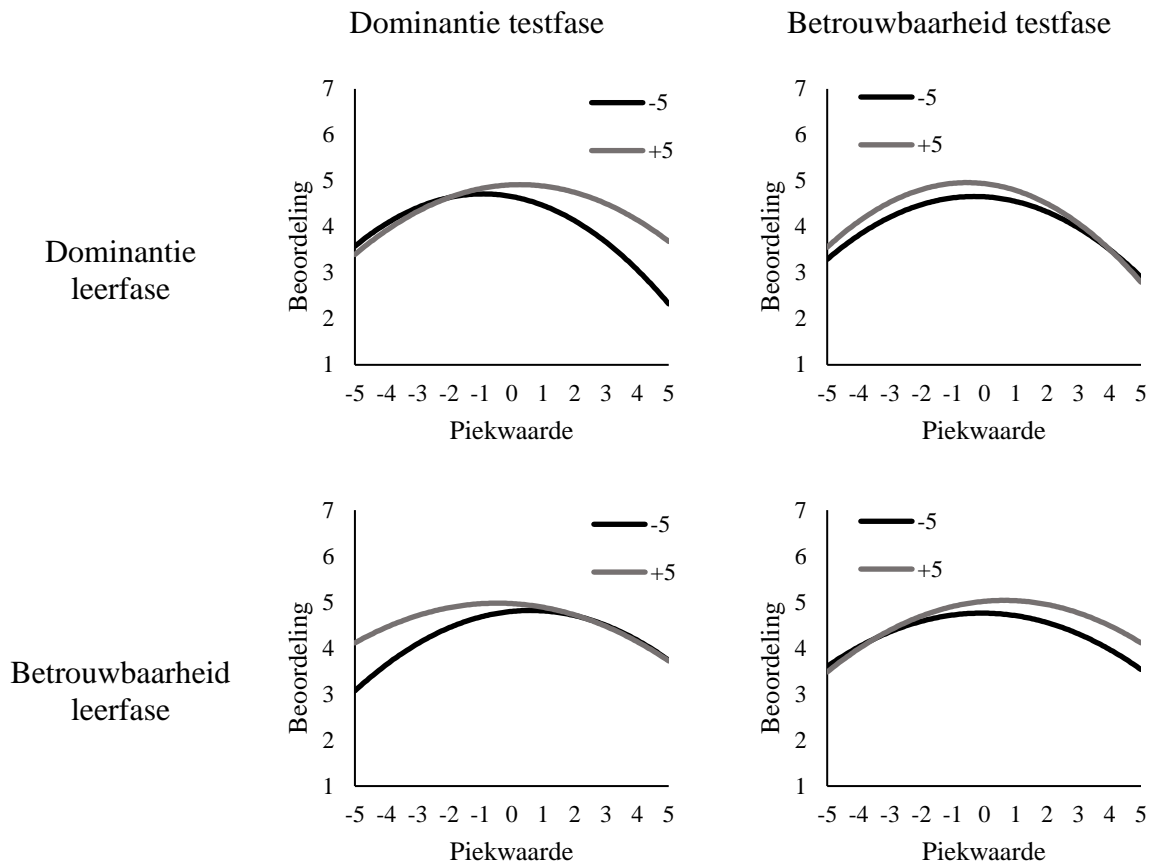
- De doeldimensie: de conditie waarin de proefpersoon in de leerfase wordt ingedeeld. Hierbij waren er twee mogelijkheden, namelijk gezichten uit de dimensie met dominantie-informatie, of gezichten uit de dimensie met betrouwbaarheidsinformatie. Omdat deze conditie verschilde tussen proefpersonen is er hier sprake van een between subjects factor.
- De z-scores: de twee soorten van de doeldimensie (dominantie en betrouwbaarheid) verschilden in de leerfase ook wat betreft de gemiddelde z-scores. Dit houdt in dat sommige proefpersonen gezichten te zien kregen die rond een z-score van -5 liggen op de doeldimensie, terwijl andere proefpersonen gezichten zagen van rond een z-score van +5 op de doeldimensie. Ook dit betrof een between subjects factor aangezien de conditie verschilt tussen proefpersonen.
- De taakdimensie: in de testfase van het experiment beoordeelden de proefpersonen de gezichten die ze zagen op dominantie, betrouwbaarheid en gemiddeldheid. Deze drie eigenschappen vormen de taakdimensie. Alle proefpersonen doorliepen de beoordelingsblokken binnen de testfase, wat maakt dat de taakdimensie een within subjects factor is. Gemiddeldheid wordt in deze analyse niet meegenomen omdat hier alleen naar werd gevraagd als manipulatiecheck. Dit werd getest door middel van de eerdergenoemde  $2 \times 2$  between subjects ANOVA.
- De testdimensie: de gezichten die de proefpersonen in de testfase beoordeelden, waren afkomstig uit ofwel een dimensie met dominantie-informatie, ofwel met betrouwbaarheidsinformatie. Deze gezichten waren hetzelfde voor alle proefpersonen, wat de testdimensie een within subjects factor maakte.

De between subjects ANOVA en de repeated measures ANOVA zijn uitgevoerd door middel van het programma IBM SPSS 22.

## Resultaten

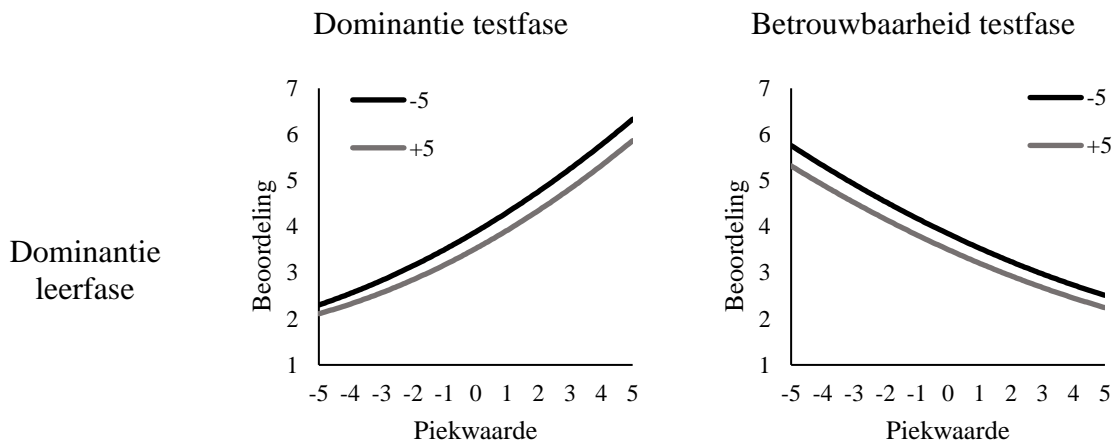
### Beschrijvende statistieken

#### GEMIDDELDHEID

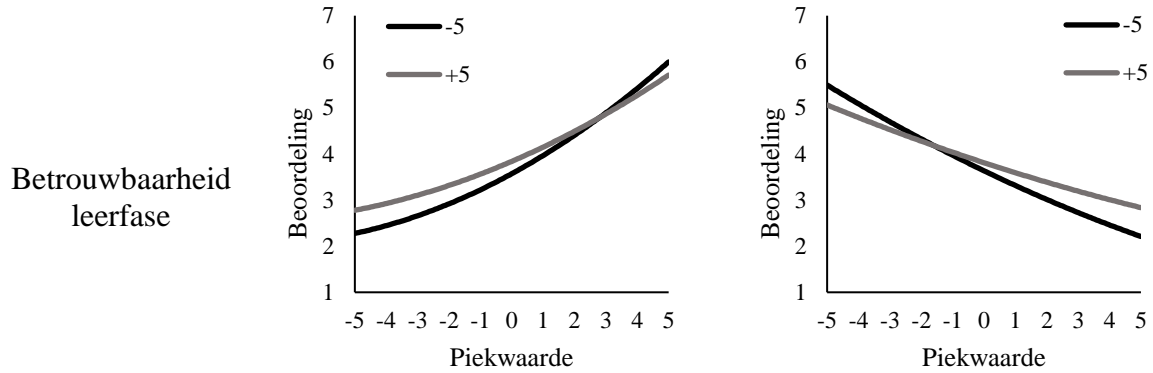


*Figuur 1.* Weergave van de geschatte pieken voor beoordelingen van gemiddeldheid, waarbij onderscheid wordt gemaakt tussen de gemanipuleerde gemiddelde z-scores van -5 en +5.

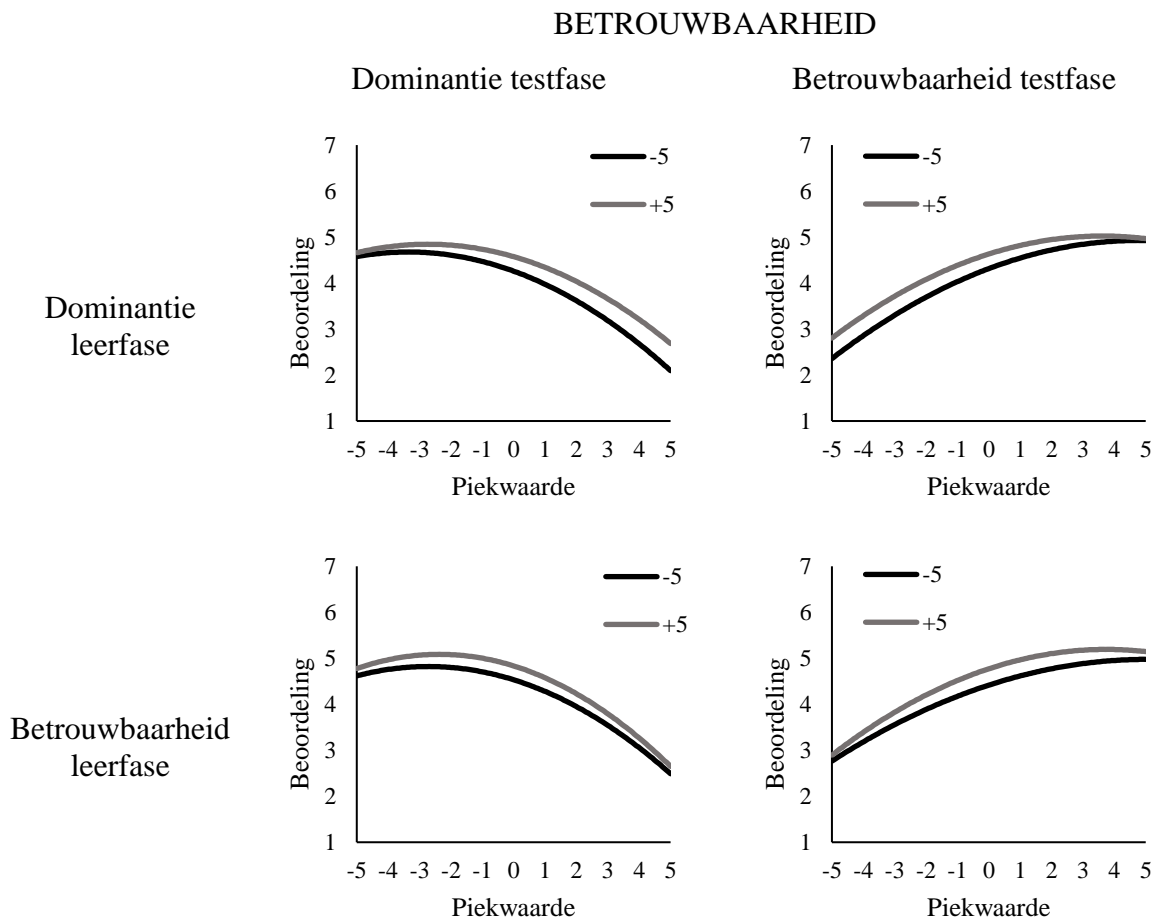
#### DOMINANTIE



# WANNEER BEÏNVLOEDT STATISTISCH LEREN SOCIALE PERCEPTIE VAN GEZICHTEN?



Figuur 2. Weergave van de geschatte pieken voor beoordelingen van dominantie, waarbij onderscheid wordt gemaakt tussen de gemanipuleerde gemiddelde z-scores van -5 en +5.



Figuur 3. Weergave van de geschatte pieken voor beoordelingen van betrouwbaarheid, waarbij onderscheid wordt gemaakt tussen de gemanipuleerde gemiddelde z-scores van -5 en +5.

Tabel 1

*De gemiddelde geschatte pieken per factor in het onderzoeksdesign*

Taakdimensie	Doeldimensie		Testdimensie		Z-score	
	Dominantie	Betrouwbaarheid	Dominantie	Betrouwbaarheid	-5	+5
Dominantie	4.70	-4.54	0.05	0.18	0.03	0.16
Betrouwbaarheid	-2.55	3.38	0.41	0.43	0.40	0.44
Gemiddeldheid	-0.18	0.44	0.08	0.26	0.02	0.29

**Statistische analyses**

Uit de  $2 \times 2$  (Doeldimensie [dominantie, betrouwbaarheid]  $\times$  Z-scores [-5, +5]) between subjects ANOVA die is uitgevoerd om de manipulatiecheck te meten, is gebleken dat er geen significante effecten optreden voor gemiddeldheid in de situaties waarin de aangeboden dimensie in de leerfase gelijk was aan de beoordeelde dimensie in de testfase. Er is geen hoofdeffect gevonden voor de doeldimensie op beoordelingen van gemiddeldheid,  $F(1,120) = 1.484$ ,  $p = .226$ . Ook bleek er geen hoofdeffect te zijn van de verschillende gemanipuleerde z-scores,  $F(1, 120) = 0.177$ ,  $p = .675$ . Tenslotte is er ook geen interactie-effect gevonden tussen de doeldimensie en de gemanipuleerde z-scores,  $F(1, 120) = 1.407$ ,  $p = .238$ .

In de  $2 \times 2 \times 2 \times 2$  (Doeldimensie [dominantie, betrouwbaarheid]  $\times$  Z-scores [-5, +5]  $\times$  Taakdimensie [dominantie, betrouwbaarheid]  $\times$  Testdimensie [dominantie, betrouwbaarheid]) repeated measures ANOVA zijn alle factoren, met uitzondering van gemiddeldheid, geanalyseerd. Er bleek geen hoofdeffect te zijn voor de taakdimensie,  $F(1, 120) = 3.210$ ,  $p = 0.76$ . Ook zijn er geen interactie-effecten gevonden voor de taakdimensie in combinatie met de verschillende gemanipuleerde z-scores uit de leerfase ( $F(1, 120) = 0.220$ ,  $p = .640$ ), de doeldimensie uit de leerfase ( $F(1, 120) = 0.356$ ,  $p = .552$ ), of de drie factoren gecombineerd ( $F(1, 120) = 0.795$ ,  $p = .374$ ).

Daarentegen is er wel een hoofdeffect gevonden voor de testdimensie,  $F(1, 120) = 49.814$ ,  $p < .001$ ,  $\eta_p^2 = .293$ . Dit betekent dat er een significant verschil bestaat tussen de beoordelingen die zijn gedaan voor gezichten uit de dominantiedimensie en gezichten uit de betrouwbaarheidsdimensie in de testfase van het experiment. Er traden echter geen interactie-effecten op van de testdimensie in combinatie met de gemanipuleerde z-scores ( $F(1, 120) = 0.139$ ,  $p = .710$ ), de doeldimensie ( $F(1, 120) = 0.004$ ,  $p = .952$ ), of een combinatie van deze drie factoren ( $F(1, 120) = 0.005$ ,  $p = .944$ ).

De taakdimensie en de testdimensie interacteren wel significant met elkaar,  $F(1, 120) = 736.447$ ,  $p < .001$ ,  $\eta_p^2 = .860$ . Dit houdt in dat de beoordelingen op gezichten uit de

dominantie- en betrouwbaarheidsdimensie in de testfase, verschillend waren voor de gevraagde dominantiebeoordeling dan voor de gevraagde betrouwbaarheidsbeoordeling van de gezichten in de testfase. Er treedt vervolgens geen interactie-effect op van de taakdimensie en testdimensie met de verschillende gemanipuleerde z-scores,  $F(1, 120) = 0.670, p = .415$ . Ook is er geen interactie-effect van de taakdimensie en testdimensie met de doeldimensie,  $F(1, 120) = 0.359, p = .550$ . Tenslotte laat de combinatie van de vier factoren (taakdimensie, testdimensie, z-scores en doeldimensie) geen interactie-effect zien,  $F(1, 120) = 0.267, p = .606$ .

Voor alle uitgevoerde analyses is er als significantieniveau een alfa-waarde van .05 gehanteerd.

### Discussie

Het doel van de huidige studie was om te onderzoeken of statistisch leren ook sociale perceptie beïnvloedt in de aanwezigheid van informatieve cues, in dit geval cues die samenhangen met betrouwbaarheid of dominantie. In een experiment werd gemanipuleerd tussen proefpersonen of ze gezichten uit een dominantie- of betrouwbaarheidsdimensie aangeboden kregen, evenals de locatie van de piek van de normaalverdeling die de gezichten op die dimensies volgden. Verwacht werd dat bij het beoordelen van nieuwe gezichten, dominantie-oordelen van gezichten uit de dominantiedimensie niet gevormd zouden worden op basis van statistisch leren, aangezien er al diagnostische kenmerken voor het oordeel in het gezicht aanwezig waren. Ditzelfde effect werd verwacht voor betrouwbaarheidsoordelen van gezichten uit de betrouwbaarheidsdimensie. Daarnaast werd verwacht dat er wel sprake zou zijn van statistisch leren bij dominantie-oordelen van gezichten variërend op betrouwbaarheid en betrouwbaarheidsoordelen van gezichten variërend op dominantie. In het geval dat de aangeboden dimensie uit de leerfase overeenkwam met de beoordeelde dimensie in de testfase, was de verwachting dat het oordeel van de proefpersonen zou pieken rond het gemiddelde uit de leerfase en een curvilineaire relatie zou laten zien. In het geval dat de dimensie uit de leerfase en de dimensie uit de testfase niet met elkaar overeenkwamen, werd verwacht dat er een curvilineaire relatie uit zou volgen met een piek rond een z-score van 0. Uit de resultaten van de uitgevoerde analyses is gebleken dat de verwachte resultaten van het onderzoek niet zijn uitgekomen.

Aangezien er geen effecten zijn gevonden van de aangeboden dimensies in de leerfase en de locatie van de piek van de normaalverdeling die de gezichten op die dimensies volgden

## WANNEER BEÏNVLOEDT STATISTISCH LEREN SOCIALE PERCEPTIE VAN GEZICHTEN?

op de beoordelingen van gemiddeldheid, lijkt het erop dat de proefpersonen de dimensies in de leerfase van het experiment niet hebben aangeleerd.

Uit de resultaten van de uitgebreide analyse, waarin alle factoren zijn meegenomen, is gebleken dat de verschillende condities van de taakdimensie niet leidden tot onderscheidende beoordelingen. De verschillen tussen de antwoorden op de vragen over dominantie en de antwoorden op de vragen over betrouwbaarheid in de testfase van het experiment verschilden niet genoeg van elkaar om te kunnen spreken van een effect. De verschillende dimensies die de proefpersonen aangeboden kregen in de leerfase van het experiment hadden ook geen invloed op de beoordelingen van dominantie en betrouwbaarheid in de testfase. Daarbij komt ook dat de verschillende gemanipuleerde z-scores in de leerfase niet bijdroegen aan verschillende beoordelingen.

Wat wel uit de analyse is gebleken, is dat de beoordelingen van gezichten uit de dominantiedimensie verschilden van de beoordelingen van gezichten uit de betrouwbaarheidsdimensie in de testfase van het experiment. Deze verschillende beoordelingen zijn echter niet het gevolg van de twee soorten aangeboden dimensies of de twee soorten gemanipuleerde z-scores uit de leerfase. Het gevonden verschil laat zien dat proefpersonen simpelweg de dominantie- en betrouwbaarheidsinformatie van de gezichten gebruiken voor hun beoordelingen ervan. Dominantie en betrouwbaarheid correleren negatief met elkaar (Todorov, Dotsch, Porter, Oosterhof & Falvello, 2013), wat een verklaring biedt voor het gevonden patroon.

Vervolgens is er ook gebleken dat er een effect bestaat tussen de taakdimensie en de testdimensie. De vragen die gesteld werden om te achterhalen hoe dominant en betrouwbaar de proefpersonen de gezichten in de testfase beoordeelden, lijken te hebben geleid tot verschillende beoordelingen van gezichten uit de dominantie- en betrouwbaarheidsdimensie. Dit effect hangt echter niet samen met de verschillende gemanipuleerde z-scores en aangeboden dimensies uit de leerfase.

De visuele weergaven van de data in Figuren 1, 2 en 3 bevestigen de gevonden resultaten uit de analyses. De grafieken van de beoordelingen op gemiddeldheid (Figuur 1) laten curvilineaire lijnen zien, terwijl de lijnen voor de beoordelingen van dominantie (Figuur 2) en betrouwbaarheid (Figuur 3) een meer lineaire trend lijken te volgen. De twee grafieken van de beoordelingen op dominantie bij gezichten uit de betrouwbaarheidsdimensie laten een gelijksoortig effect zien als de grafieken uit het onderzoek van Dotsch en collega's (2016). De grafieken laten echter niet de curvilineaire relaties zien die aanvankelijk werden verwacht.

Opvallend is dat de grafieken van dominantiebeoordelingen bij gezichten uit de dominantiedimensie dezelfde stijgende lijn laten zien als de grafieken van betrouwbaarheidsbeoordelingen bij gezichten uit de betrouwbaarheidsdimensie. De grafieken van dominantiebeoordelingen bij gezichten uit de betrouwbaarheidsdimensie laten daarentegen een dalende lijn zien, wat ook optreedt bij de grafieken van betrouwbaarheidsbeoordelingen bij gezichten uit de dominantiedimensie. Deze uitkomsten zijn het gevolg van het interactie-effect tussen de taakdimensie en de testdimensie. De grafieken verschillen niet op enkel de aangeboden dimensies in de leerfase, wat ook uit de uitgevoerde analyse is gebleken. Daarnaast is aan de grafieken te zien dat de lijnen in de meeste gevallen vrijwel parallel aan elkaar zijn, wat het gebrek aan een effect voor de gemanipuleerde z-scores bevestigt.

### **Beperkingen**

Het ontbrekende effect van de aangeleerde dimensies op gemiddeldheid is in tegenstrijd met de bevindingen van Dotsch en collega's (2016). Een reden voor dit onverwachte resultaat kan zijn dat er minder gezichten werden getoond in de leerfase van het experiment dan in het onderzoek van Dotsch en collega's (2016). Er is gekozen om in huidig experiment minder gezichten te tonen aangezien dit in een eerder onderzoek nog steeds significante resultaten opleverde (R. Dotsch, persoonlijke communicatie, 13 februari 2017). Wellicht was dit een eenmalige bevinding, of verschilde het experiment zodanig dat het heeft kunnen leiden tot een ander resultaat.

De verwachte effecten die niet zijn opgetreden als resultaat van de testfase van het experiment kunnen een gevolg zijn van het feit dat er minder gezichten zijn beoordeeld in deze testfase. In het experiment van Dotsch en collega's (2016) werden er 29 gezichten beoordeeld op de onderzochte kenmerken, waar er in huidig experiment 21 gezichten werden beoordeeld. Er zijn minder gezichten gebruikt in de testfase omdat er ook minder gezichten zijn gebruikt in de leerfase ten opzichte van voorgaand onderzoek.

Wat ook een bijdrage heeft kunnen leveren aan de onverwachte resultaten is het feit dat de proefpersonen niet gelijk verdeeld waren over de verschillende condities in de leerfase van het experiment. Doordat een aantal groepen maar weinig proefpersonen bevatte, is dit statistisch nadelig voor het vinden van een verschil tussen deze groepen.

Bij het uitvoeren van een vervolgonderzoek is het aan te raden om voorafgaand van het daadwerkelijke experiment een pilot test te doen waarmee getest kan worden of de gekozen dimensies daadwerkelijk aangeleerd kunnen worden met minder gezichten in de leerfase.

## WANNEER BEÏNVLOEDT STATISTISCH LEREN SOCIALE PERCEPTIE VAN GEZICHTEN?

Tevens is het uitvoeren van het experiment van belang dat de proefpersonen evenredig verdeeld worden over de verschillende condities van het onderzoek.

### **Conclusie**

Voorgaand onderzoek heeft aangetoond dat een gezicht verschillende beoordelingen kan uitlokken op basis van zijn statistische eigenschappen. Het huidige onderzoek is dieper ingegaan op statistisch leren, waarbij is onderzocht of statistisch leren invloed heeft op sociale perceptie wanneer er informatieve cues voor dominantie en betrouwbaarheid aanwezig zijn in de getoonde gezichten. De verwachting was dat het statistisch leren effect zou optreden wanneer de informatieve cues in de gezichten niet informatief waren voor de beoordeling ervan, maar de proefpersonen er wel statistische informatie over hadden aangeleerd. Dit effect is echter niet terug te zien in de resultaten van het uitgevoerde experiment. Dit onderzoek wijst uit dat het statistisch leren effect niet generaliseert naar situaties waarin gezichten duidelijk variëren op gezichtskenmerken die diagnostisch zijn voor bepaalde sociale kenmerken. Het zou cruciaal geweest kunnen zijn geweest voor het onderzoek van Dotsch et al. (2016) om een dimensie te gebruiken die georthogonaliseerd was aan deze sociale dimensies.

Het statistisch leren effect blijkt minder generaliseerbaar te zijn dan verwacht op basis van voorgaand onderzoek. De onverwachte resultaten kunnen echter het gevolg zijn van beperkingen zoals de onevenredige verdeling van proefpersonen over de condities. Ook is het niet zeker of de ontbrekende effecten wellicht wel gevonden waren bij blootstelling aan meer gezichten in de leerfase van het experiment. Tenslotte biedt ook de negatieve correlatie tussen dominantie en betrouwbaarheid een mogelijke verklaring. Vervolgonderzoek en replicaties zullen moeten uitwijzen of de verwachte effecten wel uit een gelijksoortig onderzoek zouden kunnen volgen.



Referenties

- Boothroyd, L. G., Jones, B. C., Burt, D. M., & Perrett, D. I. (2007). Partner characteristics associated with masculinity, health and maturity in male faces. *Personality and Individual Differences, 43*(5), 1161-1173. <https://doi.org/10.1016/j.paid.2007.03.008>
- Buckingham, G., DeBruine, L. M., Little, A. C., Welling, L. L., Conway, C. A., Tiddeman, B. P., & Jones, B. C. (2006). Visual adaptation to masculine and feminine faces influences generalized preferences and perceptions of trustworthiness. *Evolution and Human Behavior, 27*(5), 381-389. <https://doi.org/10.1016/j.evolhumbehav.2006.03.001>
- Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces a developmental study. *Psychological science, 25*(5), 1132-1139. <https://doi.org/10.1177/0956797614523297>
- Dotsch, R., Hassin, R. R., & Todorov, A. (2016). Statistical learning shapes face evaluation. *Nature Human Behaviour, 1*, 0001. <https://doi.org/10.1038/s41562-016-0001>
- Flowe, H. D., & Humphries, J. E. (2011). An examination of criminal face bias in a random sample of police lineups. *Applied Cognitive Psychology, 25*(2), 265-273. <https://doi.org/10.1002/acp.1673>
- Krumhuber, E., Manstead, A. S., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion, 7*(4), 730. <https://doi.org/10.1037/1528-3542.7.4.730>
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*(32), 11087-11092. <https://doi.org/10.1073/pnas.0805664105>
- Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: The impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime & Law, 16*(6), 477-491. <https://doi.org/10.1080/10683160902926141>
- Secord, P. F., Dukes, W. F., & Bevan, W. (1954). Personalities in faces: I. An experiment in social perceiving. *Genetic psychology monographs.*
- Sofer, C., Dotsch, R., Oikawa, M., Oikawa, H., Wigboldus, D. H., & Todorov, A. (2017). For Your Local Eyes Only: Culture-Specific Face Typicality Influences Perceptions of Trustworthiness. *Perception, 0301006617691786*. <https://doi.org/10.1177/0301006617691786>
- Sofer, C., Dotsch, R., Wigboldus, D. H., & Todorov, A. (2014). What is typical is good the

- influence of face typicality on perceived trustworthiness. *Psychological science*, 0956797614554955. <https://doi.org/10.1177/0956797614554955>
- Sutherland, C. A., Oldmeadow, J. A., Santos, I. M., Towler, J., Burt, D. M., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, 127(1), 105-118. <https://doi.org/10.1016/j.cognition.2012.12.001>
- Todorov, A., Dotsch, R., Porter, J. M., Oosterhof, N. N., & Falvello, V. B. (2013). Validation of data-driven computational models of social perception of faces. *Emotion*, 13(4), 724. <https://doi.org/10.1037/a0032335>
- Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, 66, 519-545. <https://doi.org/10.1146/annurev-psych-113011-143831>
- Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, 27(6), 813-833. <https://doi.org/10.1521/soco.2009.27.6.813>
- Willis, J., & Todorov, A. (2006). First impressions making up your mind after a 100-ms exposure to a face. *Psychological science*, 17(7), 592-598. <https://doi.org/10.1111/j.1467-9280.2006.01750.x>

Bijlage A  
Syntax SPSS

**Tabel 1**

```
DATASET ACTIVATE DataSet1.
* Custom Tables.
CTABLES
  /VLABELS VARIABLES=task testdim peakloc targetdim mu DISPLAY=LABEL
  /TABLE task BY testdim > peakloc [MEAN] + targetdim > peakloc
[MEAN] + mu > peakloc [MEAN]
  /CATEGORIES VARIABLES=task testdim targetdim mu ORDER=A KEY=VALUE
EMPTY=EXCLUDE.
```

**2 × 2 ANOVA gemiddeldheid**

```
USE ALL.
COMPUTE filter_$=(task = 'Typ' & testdim='Dom' & targetdim = 'Dom' |
task='Typ' & testdim = 'Trust'
& targetdim='Trust').
VARIABLE LABELS filter_$ "task = 'Typ' & testdim='Dom' & targetdim =
'Dom' | task='Typ' & "+
"testdim = 'Trust' & targetdim='Trust' (FILTER)".
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

```
DATASET ACTIVATE DataSet1.
UNIANOVA peakloc BY mu targetdim
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /POSTHOC=mu targetdim(TUKEY BONFERRONI)
  /EMMEANS=TABLES(mu) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(targetdim) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(mu*targetdim)
  /PRINT=ETASQ HOMOGENEITY DESCRIPTIVE
  /CRITERIA=ALPHA(.05)
  /DESIGN=mu targetdim mu*targetdim.
```

**2 × 2 × 2 × 2 repeated measures ANOVA**

```
USE ALL.
COMPUTE filter_$=(task = 'Dom' | task = 'Trust').
VARIABLE LABELS filter_$ "task = 'Dom' | task = 'Trust' (FILTER)".
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

```
IF (task = 'Dom' & testdim='Dom')
taskdimensie_dom_testdim_dom=peakloc.
EXECUTE.
```

## WANNEER BEÏNVLOEDT STATISTISCH LEREN SOCIALE PERCEPTIE VAN GEZICHTEN?

```
IF (task = 'Dom' & testdim='Trust')
taskdimensie_dom_testdim_trust=peakloc.
EXECUTE.
```

```
IF (task = 'Trust' & testdim='Dom')
taskdimensie_trust_testdim_dom=peakloc.
EXECUTE.
```

```
IF (task = 'Trust' & testdim='Trust')
taskdimensie_trust_testdim_trust=peakloc.
EXECUTE.
```

```
GLM taskdimensie_dom_testdim_dom taskdimensie_dom_testdim_trust
taskdimensie_trust_testdim_dom
  taskdimensie_trust_testdim_trust BY plusmin targetdimensie
  /WSFACTOR=taskdimensie 2 Polynomial testdimensie 2 Polynomial
  /METHOD=SSTYPE(3)
  /SAVE=SRESID
  /POSTHOC=plusmin targetdimensie(TUKEY BONFERRONI)
  /PLOT=PROFILE(taskdimensie*plusmin*targetdimensie)
  /EMMEANS=TABLES(plusmin) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(targetdimensie) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(taskdimensie) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(testdimensie) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(plusmin*targetdimensie)
  /EMMEANS=TABLES(plusmin*taskdimensie)
  /EMMEANS=TABLES(plusmin*testdimensie)
  /EMMEANS=TABLES(targetdimensie*taskdimensie)
  /EMMEANS=TABLES(targetdimensie*testdimensie)
  /EMMEANS=TABLES(taskdimensie*testdimensie)
  /EMMEANS=TABLES(plusmin*targetdimensie*taskdimensie)
  /EMMEANS=TABLES(plusmin*targetdimensie*testdimensie)
  /EMMEANS=TABLES(plusmin*taskdimensie*testdimensie)
  /EMMEANS=TABLES(targetdimensie*taskdimensie*testdimensie)
  /EMMEANS=TABLES(plusmin*targetdimensie*taskdimensie*testdimensie)
  /PRINT=DESCRIPTIVE ETASQ HOMOGENEITY
  /CRITERIA=ALPHA(.05)
  /WSDESIGN=taskdimensie testdimensie taskdimensie*testdimensie
  /DESIGN=plusmin targetdimensie plusmin*targetdimensie.
```