

Validiteit van de 4DKL bij mensen met een migratieachtergrond

Berend Terluin, Harm van Marwijk

De Vierdimensionale Klachtenlijst (4DKL) wordt in de huisartsenpraktijk veel gebruikt om psychische klachten in kaart te brengen. Huisartsen vragen zich wel eens af of de 4DKL wel hetzelfde meet bij mensen met een migratieachtergrond als bij mensen met een Nederlandse achtergrond. En mogen we de scores op dezelfde manier interpreteren? Om deze vragen te beantwoorden hebben we een *differential item functioning*-analyse uitgevoerd. Het blijkt dat de 4DKL hetzelfde meet bij personen met een migratieachtergrond – afkomstig uit Turkije, Marokko, Suriname, Antillen of Indonesië – als bij mensen met een Nederlandse achtergrond. Ze moeten daarvoor wel de Nederlandse taal voldoende machtig zijn.

Huisartsen en POH's-ggz gebruiken de Vierdimensionale Klachtenlijst (4DKL) om psychische klachten in kaart te brengen.^{1,2} De 4DKL heeft aparte schalen voor *distress*, depressie, angst en somatisatie. Bij *distress* gaat het om symptomen van psychische spanningen zoals die onvermijdelijk optreden wanneer iemand kampt met te hoge

eisen in de werk- of privésfeer, problemen door verschillende oorzaken of ingrijpende gebeurtenissen. De *distress*-score geeft de huisarts een indruk van hoe moeilijk de patiënt het heeft en helpt het gesprek met de patiënt op gang te brengen om erachter te komen waarmee deze worstelt.³ De depressieschaal meet ernstige depressieve



Bijna een kwart van de Nederlandse bevolking bestaat uit mensen met een migratieachtergrond en ook zij komen met psychische problemen bij de huisarts.

Foto: ANP

WAT IS BEKEND?

- Huisartsen en POH's-ggz gebruiken de Vierdimensionale Klachtenlijst (4DKL) om psychische klachten in kaart te brengen.
- Een kwart van de Nederlandse bevolking heeft een migratieachtergrond.

WAT IS NIEUW?

- Mensen met een migratieachtergrond [afkomstig uit Turkije, Marokko, Suriname, Antillen of Indonesië] vullen de 4DKL niet anders in dan personen met een Nederlandse achtergrond [met uitzondering van 1 item].
- We kunnen de 4DKL-scores van personen met een migratieachtergrond op dezelfde manier interpreteren als die van mensen met een Nederlandse achtergrond.

symptomen en de depressiescore kan de huisarts attenderen op een mogelijke depressieve stoornis.^{4,5} De angstschaal meet ernstige angstsymptomen, waaronder paniek en vermijdingsgedrag. Met de angstschaal kan de huisarts een mogelijke angststoornis achterhalen.^{4,6} Bij somatisatie gaat het om lichamelijke spanningsklachten die tot op zekere hoogte normaal zijn voor mensen die het moeilijk hebben. Relatief hoge somatisatiescores wijzen op het mogelijk bestaan van vicieuze cirkels, waarbij ongerustheid, gedachten en gedrag de klachten onbedoeld versterken, en op mogelijke 'centrale sensitivatie'.^{7,8} Patiënten herkennen zich vaak goed in de dimensies. Een hoge distress-score zonder een verhoogde score op een van de andere dimensies wijst op een 'normale' (vaak voorbijgaande) reactie op overbelasting of tegenslagen.²

Bijna een kwart van de Nederlandse bevolking bestaat uit mensen met een migratieachtergrond.⁹ Tot deze groep behoren zowel mensen die in het buitenland zijn geboren (de eerste generatie), als personen van wie ten minste 1 ouder immigrant was (de tweede generatie). Qua aantallen migranten vormen Turkije, Marokko, Suriname en Indonesië de belangrijkste herkomstlanden. De (voormalige) Nederlandse Antillen komen op de zevende plaats, na Duitsland en Polen. Al die migranten wonen, werken en studeren in Nederland en komen van tijd tot tijd bij hun huisarts, ook met psychische problemen. Huisartsen die gewend zijn om de 4DKL te gebruiken, hebben zich misschien wel eens afgevraagd of zij aan de 4DKL-scores van migranten dezelfde conclusies mogen verbinden als aan die van mensen met een Nederlandse achtergrond. Een terechte vraag, want verschillen in taalbeheersing en culturele achtergrond kunnen ertoe leiden dat vragen van een vragenlijst soms anders worden begrepen en dus gescoord.¹⁰

De vraag of de 4DKL hetzelfde meet bij mensen met een migratieachtergrond als bij mensen met een Nederlandse achtergrond hebben we onderzocht met behulp van *differential item functioning* (DIF)-analyse.¹¹

METHODE

Deelnemers en procedure

Wij gebruikten gegevens die zijn verzameld in het LISS-panel (LISS: Longitudinal Internet Study of the Social Sciences) en het bijbehorende migrantenpanel. Het LISS-panel bestaat uit een representatieve steekproef van de Nederlandse bevolking, waarvan de deelnemers maandelijks online vragenlijsten beantwoorden.¹² Het migrantenpanel is samengesteld uit deelnemers met een migratieachtergrond. In juli 2013 kregen de beschikbare paneeldeelnemers van 16 jaar en ouder een 4DKL voorgelegd. In beide panels identificeerden we personen met een migratieachtergrond, met als herkomstland Turkije, Marokko, Suriname, de (voormalige) Nederlandse Antillen (verder: 'Antillen') of Indonesië. Voor iedere herkomstgroep selecteerden we een referentiegroep van personen met een Nederlandse achtergrond, gematcht op leeftijd (10 jaar-categorieën) en geslacht, in een verhouding van 2:1. Van matching op opleiding hebben we afgezien, omdat daarvoor gegevens ontbraken en eerder bleek dat opleiding op zich nauwelijks met DIF samenhangt.¹⁰

Het meetinstrument

De 4DKL bestaat uit 50 items, verdeeld over 4 schalen: distress (16 items), depressie (6 items), angst (12 items) en somatisatie (16 items). De respondenten scoren de items op een 5-puntschaal van 0 ('nee') tot 4 ('heel vaak of voortdurend'). Voor het berekenen van totaalscores per schaal worden de itemscores 3 en 4 gehercodeerd naar een 2. Zo moeten de totaalscores minder gevoelig worden voor een eventuele neiging om te overdrijven.

Analyse

Vragenlijsten worden vaak gebruikt om constructen te meten die we niet direct kunnen observeren, zoals distress. Met gerichte vragen (bijvoorbeeld over distress-verschijnselen) proberen we een indruk te krijgen van de mate waarin het construct bij een bepaalde persoon aanwezig is. De item-responstheorie (IRT) beschrijft de antwoorden op die vragen (*items*) als functie van het onderliggende construct.¹³ Met DIF-analyse kunnen we onderzoeken of de items van een vragenlijst anders functioneren in verschillende groepen [**bijlage**].

We hebben per 4DKL-schaal een op IRT gebaseerde DIF-analyse gedaan, waarbij we 5 verschillende groepen personen met een migratieachtergrond (de 'focale' groepen) vergeleken met personen met een Nederlandse achtergrond (de referentiegroepen). Voor een uitgebreide technische beschrijving van de methode verwijzen we naar een *open access*-artikel waarin de methode is toegepast op de vraag of een Deense vertaling van de 4DKL hetzelfde meet bij Deense personen als de Nederlandse 4DKL bij Nederlandse personen.¹⁵

RESULTATEN

We hebben de 4DKL voorgelegd aan 6340 beschikbare deelne-

Tabel 1

Kenmerken van de onderzoeksgroepen naar land van herkomst

Kenmerk	Turkije (n = 92)	Marokko (n = 102)	Suriname (n = 102)	Antillen (n = 81)	Indonesië (n = 197)	Nederland (n = 1003)
Leeftijd [gemiddelde/standaarddeviatie]	37 [10]	39 [13]	46 [13]	44 [14]	53 [15]	47 [15]
Geslacht [%]						
vrouw	51	55	51	57	48	52
man	49	45	49	43	52	48
Opleiding [%]						
basisonderwijs	17	20	11	11	4	6
vmbo	18	24	19	23	18	20
havo/vwo	13	7	10	9	15	12
mbo	35	28	25	20	18	26
hbo	14	9	30	23	27	24
wo	2	6	4	12	18	11
onbekend	0	7	2	1	0	0
Burgerlijke staat [%]						
gehuwd	59	63	29	26	46	56
gescheiden	12	10	24	23	13	9
verweduwd	1	1	3	2	7	4
nooit getrouwd	28	26	44	48	34	30
Belangrijkste bezigheid [%]						
betaald werk	50	39	59	46	57	62
werkloos/werkzoekend	9	7	13	15	4	3
arbeidsongeschikt	7	10	8	7	6	3
school of studie	15	14	8	9	5	8
gepensioneerd	13	6	7	10	20	12
huisvrouw/-man	7	16	2	6	6	8
anders	0	9	4	7	3	3
Netto-inkomen per maand [%]						
0-500 euro	30	35	13	7	10	16
501-1500 euro	38	45	41	49	33	34
1501-2500 euro	26	14	28	22	36	33
> 2500 euro	1	3	12	11	17	11
onbekend	4	3	6	10	5	7
4DKL-scores [gemiddelde/standaarddeviatie]						
distress	9,0 [8,7]	9,9 [8,7]	7,9 [8,2]	5,6 [6,6]	6,7 [6,8]	5,1 [5,8]
depressie	2,3 [3,2]	1,9 [3,4]	1,4 [2,8]	0,9 [2,2]	0,9 [2,4]	0,6 [1,8]
angst	3,1 [5,0]	3,5 [5,7]	1,7 [3,6]	1,4 [3,2]	1,3 [3,0]	0,9 [2,4]
somatisatie	8,8 [8,1]	9,9 [7,8]	7,1 [7,3]	5,5 [5,5]	5,5 [4,9]	4,5 [4,5]
Interne consistentie [alfa]						
distress	0,94	0,95	0,95	0,93	0,92	0,91
depressie	0,90	0,94	0,92	0,91	0,94	0,90
angst	0,92	0,95	0,91	0,91	0,89	0,87
somatisatie	0,93	0,92	0,92	0,86	0,82	0,82

mers van het LISS-panel en 1851 deelnemers van het migrantenpanel. De respons bedroeg respectievelijk 78,2% en 74,1%. Van 5213 respondenten was het land van herkomst bekend: Nederland 3871, Turkije 92, Marokko 102, Suriname 102, Antillen 81, Indonesië 197 en andere landen 768. In totaal selecteerden we 1003 deelnemers met een Nederlandse achtergrond verdeeld over de 5 referentiegroepen (waarbij sommige deelnemers in meer dan 1 referentiegroep voorkwamen).

[Tabel 1] toont enkele achtergrondgegevens van de verschillende onderzoeksgroepen. Opvallend, maar wellicht niet verrassend, waren de verschillen tussen personen met een migra-

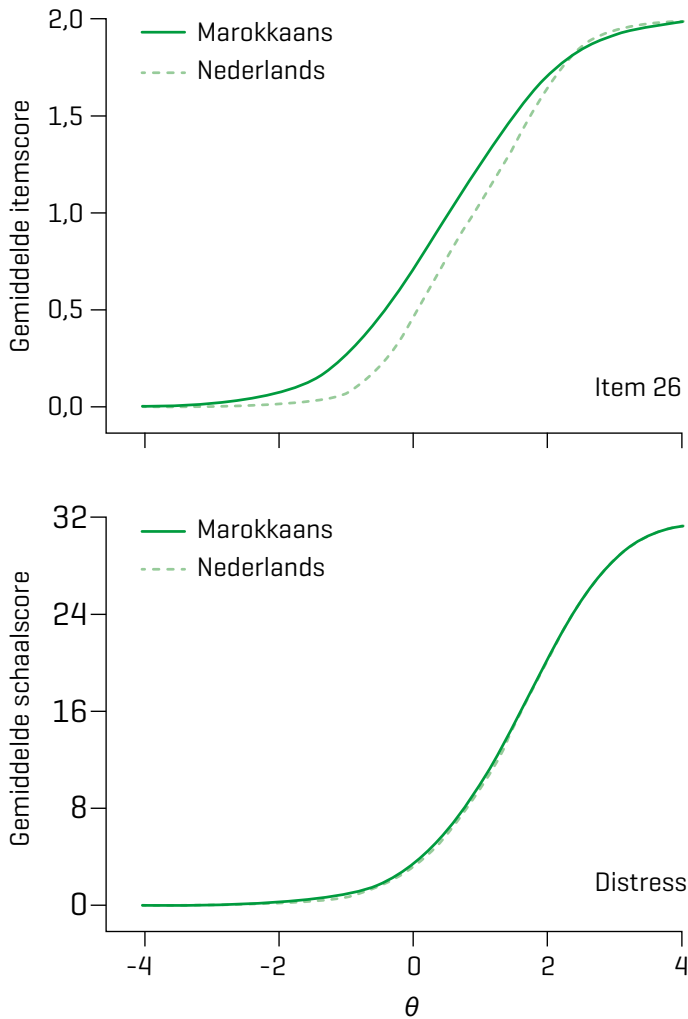
tieachtergrond en personen met een Nederlandse achtergrond voor wat betreft opleiding en inkomen. Ook zien we dat de migrantengroepen (vooral de Turkse en Marokkaanse groepen) duidelijk hogere 4DKL-scores vertoonden dan de groep met een Nederlandse achtergrond. De interne consistentie (Cronbach's alfa) van de 4DKL-schalen was in de groepen met een migratieachtergrond zeker niet lager dan in de groep met een Nederlandse achtergrond, eerder (iets) hoger.

Het onderzoek naar DIF leverde slechts 1 item met DIF op: item 26 ('snel geïrriteerd') van de distress-schaal voor de Marokkaanse groep. [Figuur 1a] toont het effect van deze DIF

Figuur 1a en 1b

Figuur 1a: Gemiddelde score van item 26 in relatie tot het werkelijke niveau van distress (uitgedrukt als Θ) bij mensen met een Marokkaanse dan wel een Nederlandse achtergrond.

Figuur 1b: gemiddelde distress-score in relatie tot het werkelijke niveau van distress (uitgedrukt als Θ) bij mensen met een Marokkaanse dan wel een Nederlandse achtergrond.



Tabel 2

Ter vergelijking: 4DKL-scores van opeenvolgende huisartsenpatiënten [A] en huisartsenpatiënten met psychosociale klachten [B]

4DKL-scores (gemiddelde/standaarddeviatie)	A	B
Distress	9,7 [8,5]	14,8 [9,1]
Depressie	1,2 [2,6]	2,4 [3,5]
Angst	2,5 [4,0]	4,4 [5,2]
Somatisatie	8,4 [6,2]	11,3 [6,7]

Bron: 4DKL-handleiding¹⁹

op de gemiddelde itemscore, afhankelijk van het niveau van distress. Voor respondenten met een Marokkaanse achtergrond was item 26 iets 'makkelijker'. [Figuur 1b] laat zien dat het effect van de DIF in item 26 vrijwel geen effect had op de distress-score. De gemiddelde distress-score, gecorrigeerd voor het verschil in werkelijke distress, was slechts 0,16 punten op een schaal van 32 punten (dus niet meer dan 0,5% van de schaal).

BESCHOUWING

Door te kijken naar het 'functioneren' van de 4DKL-items in de verschillende groepen hebben we onderzocht of de 4DKL bij mensen met een migratieachtergrond hetzelfde meet als bij mensen met een Nederlandse achtergrond. We hebben 5 (groepen) keer 50 items getest, waarbij we slechts 1 item vonden dat anders functioneerde bij mensen met een migratieachtergrond. Dit betrof het item 'snel geïrriteerd' van de distress-schaal voor de Marokkaanse groep. Dit item bleek voor respondenten met een Marokkaanse achtergrond iets 'makkelijker'. Deze personen scoorden dus gemiddeld (iets) sneller op 'snel geïrriteerd' dan personen met een Nederlandse achtergrond. Overigens was het effect klein. Op de distress-score had de gevonden DIF geen noemenswaardige invloed.

We kunnen dus stellen dat de 4DKL dezelfde validiteit heeft voor mensen met een migratieachtergrond als voor personen met een Nederlandse achtergrond. Met dit gegeven kunnen we de verschillen in 4DKL-scores tussen de migrantengroepen en de Nederlandse groep [tabel 1] dan ook met zekerheid interpreteren als een aanwijzing dat personen met een migratieachtergrond (en in het bijzonder personen van Turkse of Marokkaanse komaf) meer distress, depressie, angst en somatisatie ervaren dan mensen met een Nederlandse achtergrond. In het licht van de internationale literatuur is dit geen verrassing.¹⁶⁻¹⁸ Een migratieachtergrond hangt samen met (al of niet traumatische) stress, kansenongelijkheid, armoede en discriminatie. In ons eigen onderzoek zijn het vooral de mensen met een Turkse of Marokkaanse achtergrond die gemiddeld de laagste opleiding en het kleinste inkomen hebben [tabel 1]. Uit ons onderzoek valt niet op te maken wanneer de 4DKL moet worden ingezet. Op basis van onze ervaring adviseren wij de 4DKL bij patiënten met een migratieachtergrond in principe net zo te gebruiken als bij patiënten met een Nederlandse achtergrond. In een nascholingsartikel hebben we praktische adviezen gegeven voor de indicatie van de 4DKL, het interpreteren van de scores en het bespreken van de scores met de patiënt.³

De onderzochte steekproef uit de algemene bevolking verschilt natuurlijk van een typische populatie van huisartsenpatiënten en dat uit zich ook in de gemiddelde 4DKL-scores. [Tabel 2] toont de gemiddelde scores van een groep opeenvolgende huisartsenpatiënten (n = 2127), en van een subset van deze groep, bij wie de huisarts een psychosociale diagnose had gesteld of een sterk vermoeden had dat de klachten een psychosociale achtergrond hadden (n = 665).¹⁹ De onderzochte

mensen met een Nederlandse achtergrond hadden beduidend lagere 4DKL-scores [tabel 1] dan de huisartsenpopulaties. Mensen met een Turkse of Marokkaanse achtergrond scoorden op het niveau van huisartsenpatiënten.

Een mogelijke beperking van het onderzoek ligt in de kleinere aantallen deelnemers in 4 van de 5 migrantengroepen. Het vinden van een DIF-item met een bijna verwaarloosbaar effect op de schaalscore suggereert echter dat er voldoende deelnemers waren om DIF-items met relevante effecten te detecteren.

Een andere beperking is mogelijk de (zelf)selectie van deelnemers met een migratieachtergrond. De betreffende deelnemers waren digitaal vaardig en kennelijk in staat om Nederlandse vragenlijsten in te vullen. Bij een gebrekkiger beheersing van de Nederlandse taal kan de 4DKL eventueel worden afgenomen via een derde persoon die de lijst voorleest, toelicht en/of deels naar de moedertaal vertaalt. Het is mogelijk dat de 4DKL onder dergelijke omstandigheden niet helemaal hetzelfde meet als bij mensen met een betere beheersing van het Nederlands en bij mensen met een Nederlandse achtergrond. In voorkomende gevallen kan een 4DKL in de taal van de patiënt worden gebruikt. Onder meer de volgende vertalingen zijn beschikbaar: Turks, Arabisch, Engels, Frans, Duits en Pools (verkrijgbaar bij de eerste auteur of bij de Mapi Research Trust, mapi-trust.org).

CONCLUSIE

We kunnen concluderen dat we de Nederlandse 4DKL bij de onderzochte migrantengroepen hetzelfde kunnen interpreteren als bij mensen met een Nederlandse achtergrond. Degene die de vragenlijst invult moet de Nederlandse taal dan wel voldoende beheersen. ■

DANKBETUIGING

In dit artikel maakten we gebruik van gegevens van het LISS-panel (LISS: Longitudinal Internet Studies for the Social sciences), beheerd door CentERdata (Tilburg University).

LITERATUUR

1. Terluin B. De Vierdimensionale Klachtenlijst (4DKL). Een vragenlijst voor het meten van distress, depressie, angst en somatisatie. *Huisarts Wet* 1996;39:538-47.
2. Terluin B, Van Marwijk HWJ, Adèr HJ, De Vet HCW, Penninx BWJH, Hermens MLM, et al. The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. *BMC Psychiatry* 2006;6:34.
3. Terluin B, Terluin M, Prince K, Van Marwijk HWJ. De Vierdimensionale Klachtenlijst (4DKL) spoort psychische problemen op. *Huisarts Wet* 2008;51:251-5.
4. Terluin B, Brouwers EPM, Van Marwijk HWJ, Verhaak PFM, Van der Horst HE. Detecting depressive and anxiety disorders in distressed patients in primary care; comparative diagnostic accuracy of the Four-Dimensional Symptom Questionnaire (4DSQ) and the Hospital Anxiety and Depression Scale (HADS). *BMC Fam Pract* 2009;10:58.
5. Terluin B, Spijker J. De depressieve patiënt. In: Van der Horst H, Van Os J, redactie. *De dokter en de patiënt met psychische problemen*. Houten: Bohn Stafleu van Loghum, 2019.
6. Terluin B, Oosterbaan DB, Brouwers EPM, Van Straten A, Van de Ven PM, Langerak W, et al. To what extent does the anxiety scale of the Four-Dimensional Symptom Questionnaire (4DSQ) detect specific types of anxiety disorder in primary care? A psychometric study. *BMC Psychiatry* 2014;14:121.
7. Duddu V, Isaac MK, Chaturvedi SK. Somatization, somatosensory amplification, attribution styles and illness behaviour: a review. *Int Rev Psychiatry* 2006;18:25-33.
8. Den Boer C, Terluin B, Van Wilgen CP. Uitleg aan patiënten met onverklaarde klachten. Sensitiseringsmodel helpt bij acceptatie. *Ned Tijdschr Geneesk* 2020;164:D4642.
9. Centraal Bureau voor de Statistiek. Migratieachtergrond. Hoeveel mensen met een migratieachtergrond wonen in Nederland? 2021. <https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/migratieachtergrond/>. Geraadpleegd op 10 mei 2021.
10. Terluin B, Smits N, Brouwers EPM, De Vet HCW. The Four-Dimensional Symptom Questionnaire (4DSQ) in the general population: scale structure, reliability, measurement invariance and normative data: a cross-sectional survey. *Health Qual Life Outcomes* 2016;14:130.
11. Teresi JA, Fleishman JA. Differential item functioning and health assessment. *Qual Life Res* 2007;16:33-42.
12. Scherpenzeel A. Data collection in a probability-based internet panel: how the LISS panel was built and how it can be used. *Bulletin of Sociological Methodology* 2011;109:56-61.
13. Embretson SE, Reise SP. *Item response theory for psychologists*. Edition 2. Mahwah, NJ: Lawrence Erlbaum, 2009.
14. Woods CM, Cai L, Wang M. The Langer-improved Wald test for DIF testing with multiple groups. *Educ Psychol Meas* 2012;73:532-47.
15. Terluin B, Hoff A, Falgaard Eplow L. Assessing measurement equivalence of the Danish and Dutch Four-Dimensional Symptom Questionnaire using differential item and test functioning analysis. *Scand J Public Health* 2020;49:479-86.
16. Bhugra D. Migration and mental health. *Acta Psychiatr Scand* 2004;109:243-58.
17. Kirmayer LJ, Narasiah L, Munoz M, Rashid M, Ryder AG, Guzder J, et al. Common mental health problems in immigrants and refugees: general approach in primary care. *CMAJ* 2011;183:E959-67.
18. Rousseau C, Frounfelker RL. Mental health needs and services for migrants: an overview for primary care providers. *J Travel Med* 2019;26:tay150.
19. Terluin B, Duijsens IJ. 4DKL. Handleiding van de Vierdimensionale Klachtenlijst. Leiderdorp: Datec, 2002.

Terluin B, Van Marwijk HWJ. Validiteit van de 4DKL bij mensen met een migratieachtergrond. *Huisarts Wet* 2021;64:DOI:10.1007/s12445-021-1241-9. Amsterdam UMC, locatie VUmc, Amsterdam Public Health Onderzoeksinstituut, Amsterdam; dr. B. Terluin, senior onderzoeker, b.terluin@amsterdamumc.nl. University of Brighton, Brighton and Sussex Medical School, Department of Primary Care and Public Health, Brighton; prof. dr. H.W.J. van Marwijk, hoogleraar huisartsgeneeskunde. Mogelijke belangenverstrengeling: B. Terluin bezit de auteursrechten van de 4DKL en ontvangt royalty's voor het commerciële gebruik van de 4DKL. De 4DKL is gratis beschikbaar voor niet-commercieel gebruik.

ABSTRACT

Terluin B, Van Marwijk HWJ. The Four-Dimensional Symptom Questionnaire Validiteit is valid for immigrants too. Huisarts Wet 2021;64 DOI:10.1007/s12445-021-1241-9.

Background The Four-Dimensional Symptom Questionnaire (4DSQ) is a 50-item self-report questionnaire that measures distress, depression, anxiety, and somatization. In general practice, the 4DSQ is frequently used to map mental health problems. It is not known whether the questionnaire is valid for immigrants.

Methods We used 4DSQ data collected from participants of a representative Dutch internet panel, aged 16 and older. We selected first- and second-generation immigrants from Turkey, Morocco, Surinam, Netherlands Antilles, and Indonesia, and age- and sex-matched native Dutch participants for comparison. Item response theory based differential item functioning analysis was used to detect items that functioned differently in the immigrant

groups.

Results The 4DSQ was sent to 8191 participants, of whom 77.3% responded. The country of origin was known for 5213 participants: Netherlands 3871, Turkey 92, Morocco 102, Surinam 102, Netherlands Antilles 81, Indonesia 197, and other countries 768. We selected 1003 native Dutch participants divided into five reference groups. Differential item functioning analysis identified only one item that functioned differently in the Moroccan group: item 26 ['easily irritated'] from the distress scale. This item was slightly more easy for Moroccan responders. The effect on the distress scale score was, however, negligible.

Conclusion The 4DSQ is valid for first- and second-generation immigrants from Turkey, Morocco, Surinam, Netherlands Antilles, and Indonesia, provided that they are sufficiently fluent to complete the 4DSQ in Dutch. The 4DSQ scores of these immigrant groups can be interpreted in the same way as those for native Dutch people.

BIJLAGE

Vragenlijsten worden vaak gebruikt om constructen te meten die we niet direct kunnen observeren, zoals distress. Zo'n construct noemen we daarom *latent*. Met gerichte vragen (bijvoorbeeld over distress-verschijnselen) proberen we een indruk te krijgen van de mate waarin het construct bij een bepaalde persoon aanwezig is. De item-responstheorie (IRT) stelt dat de antwoorden op die vragen (*items*) worden gestuurd door de mate waarin het construct aanwezig is.¹³ In die zin vormen de items indicatoren van het latente construct. Een IRT-meetmodel beschrijft de (mathematische) relatie tussen de kans op een bepaald antwoord op een item enerzijds en eigenschappen van het individu en het item anderzijds. Zo wordt bijvoorbeeld de kans op een bevestigend antwoord op de vraag 'Bent u gespannen?' (met antwoordmogelijkheden 'ja' en 'nee') bepaald door de mate van distress van het individu en door de eigenschappen 'discriminatie' en 'moeilijkheid' van het item.

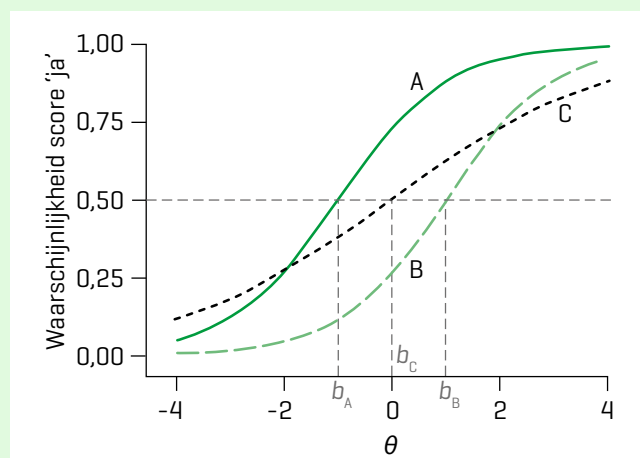
Het meetmodel kunnen we grafisch weergeven met responswaarschijnlijkheidscurves. **[Figuur 2]** toont 3 van zulke curves, die zouden kunnen passen bij 3 verschillende items van een vragenlijst. Op de x-as wordt een maat voor de ernst van het construct weergegeven in een arbitraire maat (aangeduid met de Griekse letter θ). Te zien is dat de kans om op een item te scoren toeneemt met het niveau van het onderliggende construct. De moeilijkheid (of ernst) van een item is de waarde van θ (het niveau van distress) waar de kans 50% is. De discriminatie van het item is de steilte van de responswaarschijnlijkheidscurve. De discriminatie geeft aan hoe goed het item personen met geringe distress kan onderscheiden van personen met ernstige distress. In het distress-voorbeeld zou item A 'Bent u gespannen?' kunnen zijn, en item B 'Kunt u het niet meer aan?' Voor een 'ja' op item B is beduidend meer distress nodig dan voor een 'ja' op item A. Item C zit qua moeilijkheid tussen A en B in, maar is, getuige de minder steile curve, minder goed in staat om een onderscheid tussen geringe en ernstige distress te maken dan items A en B.

Bij een IRT-analyse van een gegevensset van responsen op de items van een vragenlijst worden gelijktijdig de eigenschappen van de items (discriminatie en moeilijkheid) en de eigenschappen van de respondenten (bijvoorbeeld de mate van distress) geschat. Wanneer een item meer dan 2 responsies heeft, bijvoorbeeld 'nee' (code 0), 'soms' (code 1) en 'regelmatig' (code 2), wordt het IRT-model iets ingewikkelder. In dit voorbeeld worden er aparte responswaarschijnlijkheidscurves geschat voor de waarschijnlijkheid van scorecodes ≥ 1 , en de waarschijnlijkheid van code 2.

Idealiter wordt iemands score op een vragenlijst (bijvoorbeeld voor distress) uitsluitend bepaald door het niveau van het onderliggende construct (distress), en niet ook door de specifieke groep (bijvoorbeeld migratieachtergrond) waartoe de persoon behoort. Het hierboven bespro-

Figuur 2

Responswaarschijnlijkheidscurves. De curves tonen de kans op een bevestigend antwoord in relatie tot het niveau van het te meten construct (uitgedrukt als θ). De moeilijkheid wordt aangegeven door de locaties b_A , b_B en b_C op de θ -schaal.



ken meetmodel, waarbij de item(responsie)eigenschappen moeilijkheid en discriminatie de bepalende parameters zijn, vormt de link tussen iemands latente-constructniveau (distress) en diens gerealiseerde schaalscore (voor distress). Wanneer de groep waartoe iemand behoort invloed heeft op de schaalscore, is het meetmodel voor de groepen niet gelijk. Anders gezegd: de eigenschappen moeilijkheid en discriminatie van 1 of meer items van de schaal zijn niet gelijk. Daardoor functioneren ze anders in de ene groep dan in de andere groep. DIF-analyse onderzoekt of (sommige) items van een vragenlijst verschillende moeilijkheden en discriminaties hebben in verschillende groepen. Stel dat de responswaarschijnlijkheidscurves in **[figuur 2]** niet behoren bij verschillende items in 1 groep, maar bij eenzelfde item in verschillende groepen (A, B en C). Dan is te zien dat het item moeilijker is voor personen in groep B dan voor personen in groep A. Wanneer de andere items van de schaal waartoe het item behoort, hetzelfde functioneren, zullen bij gelijke niveaus van het onderliggende construct personen in groep B gemiddeld een lagere schaalscore realiseren dan personen in groep A. Daardoor kunnen de schaalscores van personen van groep B niet op dezelfde manier worden geïnterpreteerd als de schaalscores van de personen van groep A. Het item is voor personen in groep C (iets) makkelijker dan voor personen in groep A wanneer ze laag zitten in het construct, maar moeilijker wanneer ze hoog zitten in het construct. Voor personen in groep C die laag zitten in het construct zal dit leiden tot een gemiddeld hogere schaalscore, terwijl dit voor personen die hoog zitten in het construct leidt tot een gemiddeld lagere schaalscore. Dit betekent ook dat de schaalscores in groep C niet op dezelfde manier kunnen worden geïnterpreteerd als in groep A.