

Een evaluatie van het NVVK-congres 2015

Exploratieve data-analyse

Guldenmund, FW

Publication date

2015

Document Version

Final published version

Published in

NVVK Info

Citation (APA)

Guldenmund, FW. (2015). Een evaluatie van het NVVK-congres 2015: Exploratieve data-analyse. *NVVK Info*, (4), 13-18.

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



Exploratieve data-analyse

Een evaluatie van het NVVK-congres 2015

Het zal niemand zijn ontgaan: dit jaar vond wederom een NVVK-congres plaats, op 31 maart en 1 april in Papendal. Achteraf is het congres onder de leden van de NVVK geëvalueerd met behulp van een internet-enquête. Hieronder de resultaten en een evaluatie daarvan.

Frank Guldenmund

Van de bijna 3.000 NVVK-leden hebben 1.128 gereageerd op de enquête, waarvoor allereerst dank. Van deze 1.128 respondenten hebben 490 de vraag 'Heeft u het NVVK congres 2015 te Papendal bezocht?' met 'Ja' beantwoord. De enquête telde in totaal 98 vragen: dat zijn dan een heleboel antwoorden. Hoe verwerk en beschrijf je zo iets? Daarover gaat dit artikel.

Eerst zal ik iets vertellen over numerieke gegevens, de resultaten van een enquête, en wat je daarmee kunt doen. Daarna voer ik wat toetsen op schalen uit, die ik heb geconstrueerd op basis van een exploratieve analyse van de gegevens. Het artikel eindigt met een aantal slotopmerkingen.

Numerieke gegevens en meetniveau

Veel gegevens uit vragenlijsten worden omgezet naar nummers. Dat is heel handig, omdat met nummers gerekend kan worden. Met woorden is dat veel lastiger. Maar je moet je wel blijven realiseren waar de nummers voor staan. Het ene nummer is namelijk het andere niet. Dat klinkt misschien vreemd, maar toch is het zo.

Een voorbeeld. In een vragenlijst wordt naar het geslacht van de respondent gevraagd. De antwoordmogelijkheden zijn dan meestal 'Vrouw' en 'Man'. Deze waarden worden dan bijvoorbeeld omgezet in de getallen 1 en 2, waarbij 1 voor 'Vrouw' staat en 2 voor 'Man'. Als wij nu vervolgens naar de getallen kijken, zouden wij, geheel ten onrechte, kunnen beweren: $1 + 1 = 2$, of $2 * 1 = 2$. Wij kunnen deze bewerking niet uitvoeren met deze getallen en dat heeft te maken met hun meetniveau.

Nominale variabelen

Getallen die niets anders uitdrukken dan verschillen, bevinden zich op het nominale meetniveau. Vragenlijsten kunnen er vol mee staan: geslacht, haarkleur, woonplaats, nationaliteit, moedertaal, et cetera. Wij kunnen aan iedere waarde van nominale variabelen (man, blond, Haarlem, Nederlander, Nederlands) een getal toekennen, maar de getallen zeggen niets anders dan dat de respondent verschilt van een vrouw met blond haar die uit Breda komt, Nederlandse is en als moedertaal Noors heeft. Het belangrijkste bij nominale variabelen om te beseffen is dat de waarden willekeurig zijn. Welke waarde wij bijvoorbeeld toekennen aan 'Man' of 'Vrouw' maakt niets uit, dat kan 1, 67, 0.007 of 5687 zijn. De waarden drukken namelijk alleen een verschil uit (man ≠ vrouw, blond ≠ bruin), verder niets.

Vijf meetniveaus

Hoewel er in principe heel veel meetniveaus zijn, onderscheiden wij er in de regel vijf. Dit zijn het nominale, ordinale, interval, ratio en het absolute meetniveau.

Bij het ordinale meetniveau zijn de waarden van de (ordinale) variabele op een bepaalde manier geordend. Dit komt eveneens veel voor in vragenlijsten.

Bijvoorbeeld, bij een variabele als 'Opleiding', zijn de waarden (ongeveer) geordend: lagere school, VMBO, HAVO, MBO, VWO, HBO, WO. Maar ook de waarden van antwoordcategorieën zijn vaak geordend: helemaal mee oneens, oneens, neutraal, mee eens, helemaal mee eens. Deze waarden krijgen doorgaans de getallen 1 tot en met 5 mee. Deze waarden zijn echter niet geheel willekeurig, de waarde 'mee eens' (waarde 4) drukt méér instemming uit dan 'helemaal mee oneens' (waarde 1).

Echter, het is nog steeds de vraag of 4 vier maal 1 is. Dat is precies het kenmerk van ordinale variabelen: de getallen drukken verschillen uit, zoals bij nominale variabelen, en de getallen zijn niet geheel willekeurig, zij drukken namelijk een ordening uit. Maar daar is alles mee gezegd.

Voor variabelen op interval-niveau geldt hetzelfde als voor ordinale variabelen met de bijkomende eigenschap dat de intervallen tussen de waarden nu precies hetzelfde zijn. Hier geldt echter dat het nulpunt willekeurig is gekozen. De temperatuur-schaal is hiervan een goed voorbeeld, of de decibelschaal. Voor het ratio-niveau geldt dat het nulpunt niet willekeurig is, dit geldt bijvoorbeeld voor lengte, gewicht, afstand of hoeveelheid (frequentie).

Waarom is dit nu belangrijk om te vermelden? Welnu, het meetniveau van de getallen bepaalt welke bewerkin-

gen je met de getallen kunt uitvoeren en welke uitspraken je over de getallen kunt doen. Als je gegevens hebt verzameld op ordinaal meetniveau, kun je alleen uitspraken doen op ordinaal niveau: 'A is meer dan B', maar je weet niet hoeveel meer.

Dit geldt ook voor gegevens op interval-niveau: 20°C is niet twee keer zo warm als 10°C. Echter, een stijging van 10°C is wel steeds hetzelfde: van 10°C naar 20°C is dezelfde stijging als van 40°C naar 50°C. Om uitspraken te doen als: 'Y is twee keer zo veel als X' moet je gegevens verzameld hebben op het ratio-meetniveau. In de sociale wetenschappen, en daarmee in enquêtes, komen dergelijke metingen niet zoveel voor.

In de enquête die naar aanleiding van het NVVK-congres is gehouden, komen verschillende soorten vragen voor. Vragen op nominaal niveau ('In welke hoedanigheid bent u werkzaam?', 'Hoe bent u op de hoogte gekomen van ons congres?'), op ordinaal-niveau ('Wat is uw professionele achtergrond?', 'Wat is uw algemene waardering van dit congres?') en op ratio-niveau ('Wat is uw leeftijd?', 'Hoe lang bent u al lid van de NVVK?'). De laatste vragen zijn in principe op ratio-niveau, maar omdat met ongelijke categorieën is gewerkt, zijn de antwoorden teruggebracht naar het ordinale meetniveau.

Datareductie

490 respondenten, die antwoorden op 98 vragen betekent een heleboel gegevens. Om daar beter zicht op te krijgen, kunnen de gegevens of data, 'ingedikt' worden. In de (sociale) wetenschappen wordt daarbij vaak gebruik gemaakt van hetzij Principale Componenten Analyse (PCA) of Factor Analyse (FA). Hoewel deze technieken enigszins op elkaar lijken, zijn zij toch verschillend. De achterliggende gedachte van beide technieken is dat vragen worden beantwoord vanuit een bepaalde invalshoek of perceptie, of met een bepaalde houding of attitude. Het is dan zaak om deze perceptie of attitude te identificeren. Hoe werkt dat?

Als vragen inderdaad vanuit een bepaalde houding worden beantwoord, dan zullen vragen die vanuit dezelfde houding worden beantwoord een vergelijkbaar antwoordpatroon hebben. Dit kun je achterhalen door de samenhang of correlatie tussen vragen te bestuderen.

Vragen die een hoge correlatie hebben (een correlatie ligt altijd tussen -1 en 1), zijn wellicht vanuit dezelfde attitude beantwoord, zo luidt de aanname. Een techniek zoals PCA werkt met correlaties en gebruikt deze om 'Principale Componenten' te bepalen. Eén Principale Component is een gewogen som van alle individuele variabelen; in dit geval de individuele vragen.

Dus: $PC_1 = aX_1 + bX_2 + cX_3$, et cetera, waarbij X_1 de antwoorden op een bepaalde vraag X zijn. De gewichten die elke vraag meekrijgt (a , b , c , enzovoorts), zeggen iets over de bijdrage van een individuele vraag aan de Principale Component. Hoe groter het gewicht, des te belangrijker de bijdrage, des te meer de vraag een rol speelt in de houding die de Principale Component uiteindelijk representeert.

De enquête heeft een aantal vragen die met een waarderingsschaal beantwoord konden worden. Een waarderingsschaal loopt van 'zeer slecht' naar 'zeer goed'. Hierboven is reeds aangegeven dat dit een meting op ordinaal niveau is. Dat betekent dat wij moeten oppassen met onze analyses en interpretaties. Omdat deze schaal heel veel voorkomt in sociaal-wetenschappelijk onderzoek, zijn er speciale technieken ontwikkeld om dergelijke data te analyseren.

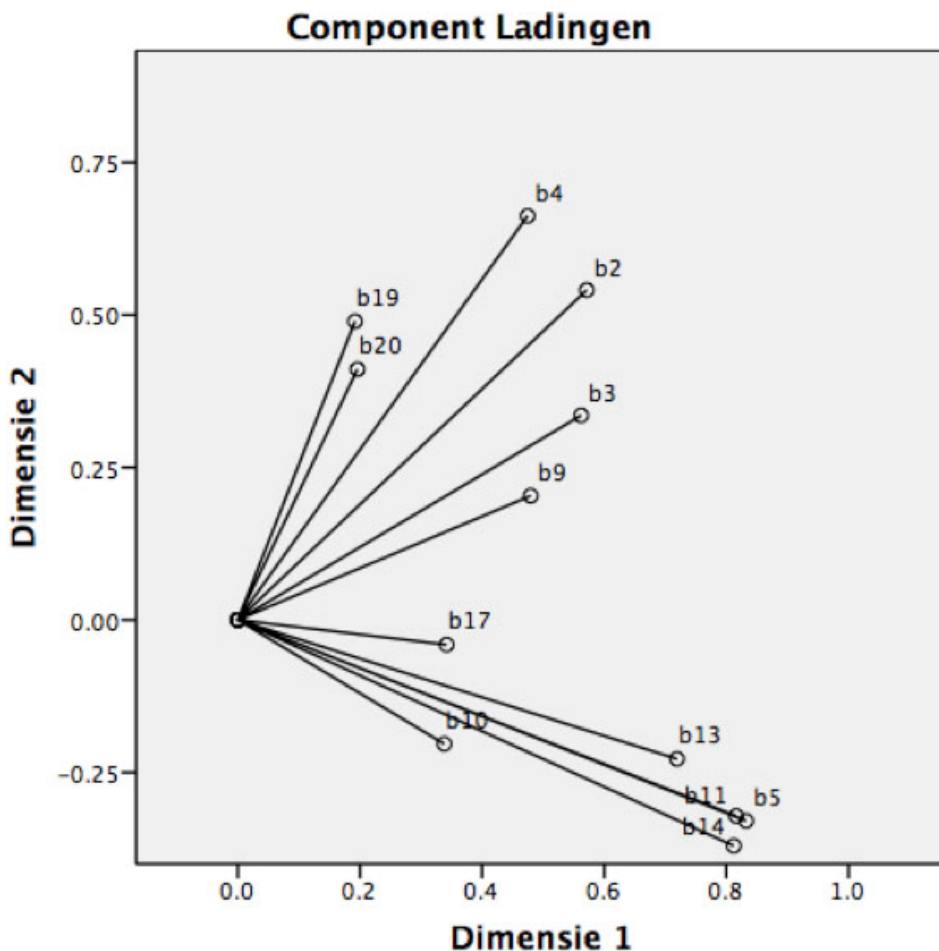
Eén techniek die hiervoor geschikt is, heet PRINCALS en is beschikbaar in het programma SPSS (Statistical Package for the Social Sciences). Het voert te ver om uit te leggen hoe dit precies in zijn werk gaat, maar het programma PRINCALS gebruikt ordinale meetgegevens en produceert daarmee een PCA. De uitkomst van een dergelijke analyse is, onder andere, een plaatje aan de hand waarvan wij groepen vragen kunnen identifi-

ceren. Nadat wij deze groepen hebben geïdentificeerd, laten wij de individuele vragen voor wat zij zijn, en gaan wij verder met de gegroepeerde vragen. Dergelijke groepen vragen worden schalen genoemd.

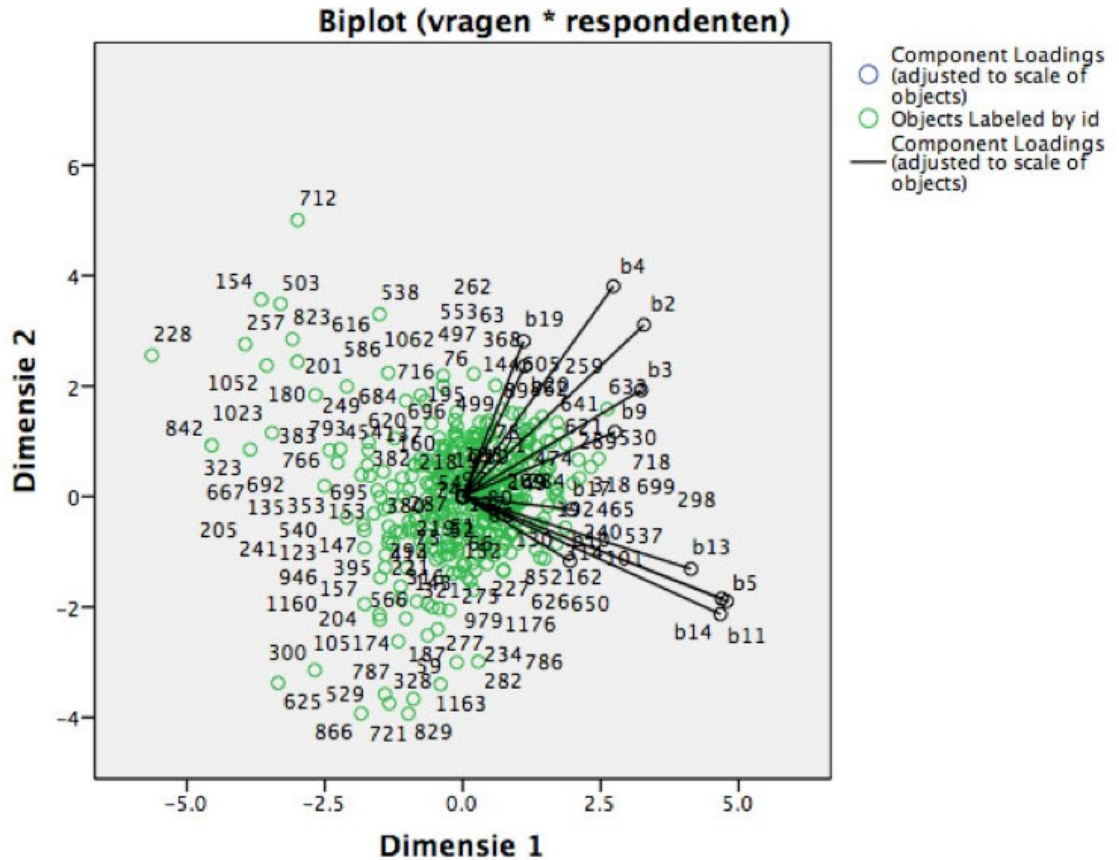
In Figuur 1 staat het resultaat van een PRINCALS-analyse van de waarderingsvragen (zeer slecht t/m zeer goed). Iedere lijn of vector staat voor een vraag. De lengte van een vector geeft de bijdrage van een vraag aan een Component (Dimensie) aan; hoe langer, des te belangrijker.

Er is duidelijk een groepje vragen (b5, b10, b11, b13, b14, b17) dat een bijdrage levert aan Dimensie 1 en een groepje vragen (b2, b3, b4, b9, b19, b20) dat een bijdrage levert aan Dimensie 2. Het is goed te zien in de figuur dat deze twee groepen min of meer loodrecht op elkaar staan (een hoek van 90° maken), wat inhoudt dat ze conceptueel verschillend zijn.

Welnu, het eerste groepje vragen gaat vooral over de inhoud van het congres, het tweede groepje over de vorm van het congres. Dit kun je vooraf misschien al bedenken, maar het is ook mooi als dat uit zo'n analyse komt. Ik heb daarnaast besloten de vragen b10 en b17 eruit te laten, omdat hun bijdrage relatief klein is (korte vector). Vragen b19 en b20 heb ik apart genomen omdat



Figuur 1 - PRINCALS van waarderingsvragen



Figuur 2 - Biplot van waarderingsvragen en respondenten

beide vragen over het avondprogramma gaan. Zo heb ik op basis van deze analyse drie schalen geconstrueerd.

Het is aardig om twee afbeeldingen te combineren. Figuur 1 is een afbeelding van de vragen die zijn meegenomen in de analyse, maar het is ook mogelijk om de respondenten hierin af te beelden. Let wel, het gaat om bijna 500 respondenten, dus het plaatje wordt vol. Zo'n plaatje wordt een biplot genoemd (vragen * respondenten).

Figuur 2 heeft opmerkelijke eigenschappen. Zo is het mogelijk om 'uitbijters' te identificeren. Respondent No. 228 (links in de figuur) ligt bijvoorbeeld aan de overzijde van de vragen b11 en b14. Nu is het zo, dat de vectoren representaties van de onderliggende schaal zijn (zeer slecht t/m zeer goed). Het is bovendien zo, dat de vectoren naar de hoogste waarde 'wijzen' (zeer goed). Dat betekent dat respondent No. 228 'zeer slecht' op deze vragen geantwoord zal moeten hebben, want deze No. 228 ligt aan de andere zijde van vectoren b11 en b14 (en b5 en b13). Dat klopt inderdaad. Respondent No. 866 (linksonder in de figuur) zal niet bijster te spreken zijn over het diner. En inderdaad, dat klopt ook (score 'slecht' op vraag 'Wat is uw oordeel over het diner?').

Ik zal het hierbij laten, maar het zal duidelijk zijn dat dergelijke analyses verder uitgebreid en verdiept kunnen worden, door alleen naar bepaalde groepen te kijken, of naar bepaalde vragen. Ik werk nu verder met de schalen die ik op basis van deze analyse heb geconstrueerd:

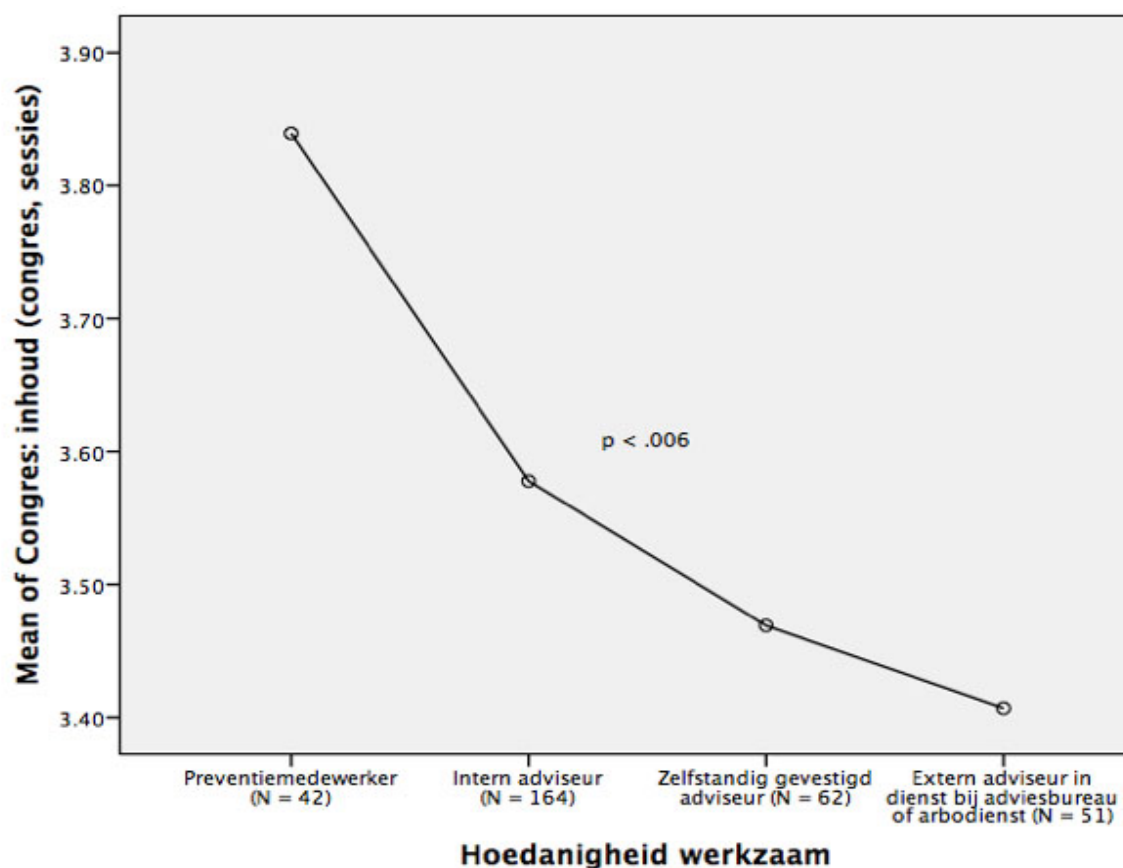
- Schaal 1: Inhoud van het congres;
- Schaal 2: Vorm van het congres;
- Schaal 3: Avondprogramma.

Toetsen van verschillen

Op basis van een aantal nominale variabelen kan vervolgens getoetst worden of er significante verschillen bestaan tussen bepaalde groepen op de schalen; dat blijkt inderdaad het geval. Eén van de vragen in de enquête luidt: 'In welke hoedanigheid bent u werkzaam?'. Ik heb hier vier verschillende groepen van gemaakt:

- 1. Preventiemedewerker;
- 2. Intern adviseur;
- 3. Zelfstandig gevestigd adviseur;
- 4. Extern adviseur in dienst bij adviesbureau of arbodienst.

De scores voor deze vier groepen hebben op Schaal 1 (inhoud van het congres) een verloop zoals aangegeven in Figuur 3.



Figuur 3 - Gemiddelde waardering voor inhoud NVVK-congres tussen preventiemedewerkers, intern, extern en zelfstandig gevestigd adviseurs

Preventiemedewerkers zijn gemiddeld meer te spreken over de inhoud van het congres dan externe adviseurs bij een adviesbureau of arbodienst. Dit verschil is bovendien statistisch significant, wat betekent dat de kans dat dit verschil bij toeval is gevonden kleiner is dan zes promille.

De gemiddelde waardering voor het avondprogramma heeft echter een ander verloop bij deze groepen en is bovendien niet significant (Figuur 4).

Nog steeds hebben preventiemedewerkers een hogere waardering, ditmaal voor het avondprogramma, maar de verschillen tussen de groepen zijn nu niet significant.

De enquête bevatte ook de volgende vraag: 'Heeft het congres inhoudelijk aan uw verwachtingen voldaan?'. Respondenten konden de vraag met 'Ja' of 'Nee' beantwoorden. Als het goed is, scoren de ja-zegggers hoger dan de nee-zegggers op Schaal 1 en dat is inderdaad het geval (3.72 vs. 2.89). Dit verschil is wederom significant ($p < .001$).

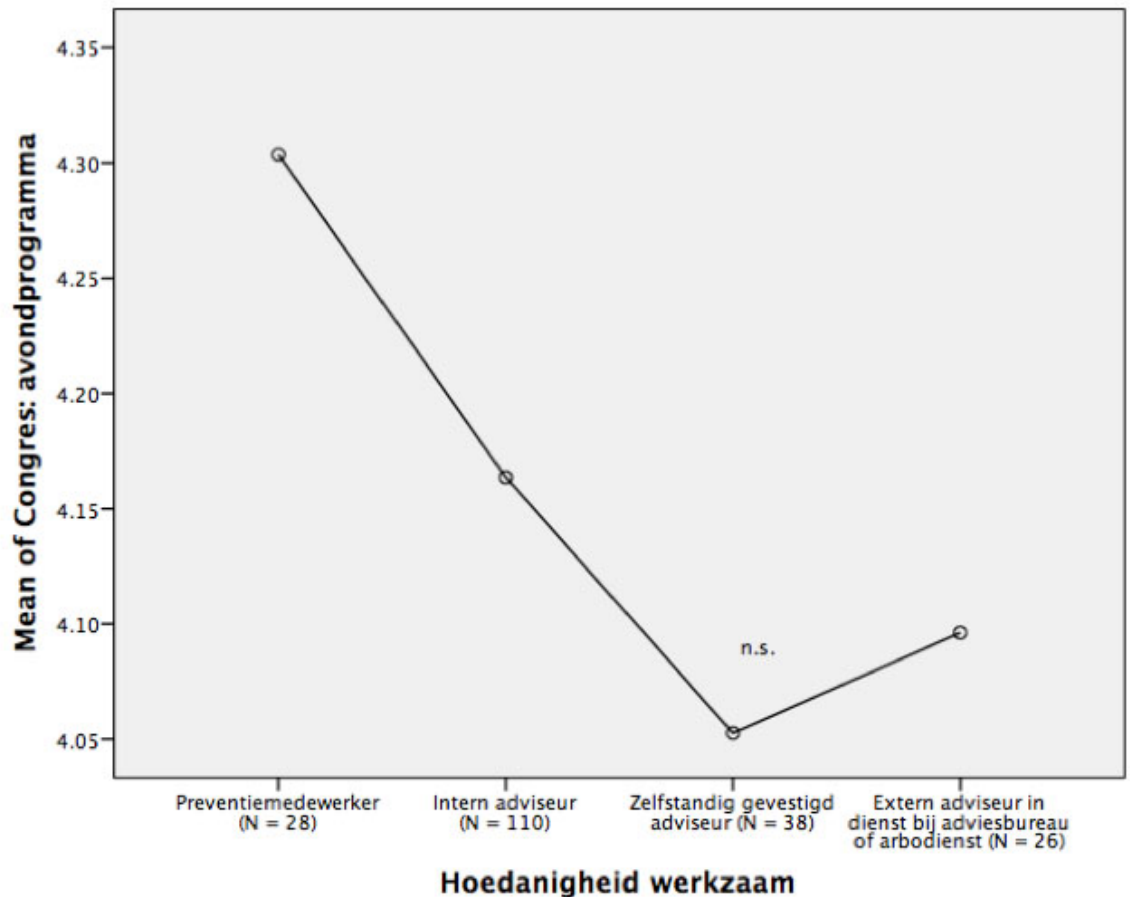
Tot slot

Het is mogelijk om nog meer resultaten te rapporteren, maar ik wil het hierbij laten. De kwalitatieve gegevens (de open vragen) kunnen gebruikt worden om de signi-

ficante verschillen verder te verdiepen of te verklaren. Zo zou het verschil tussen interne en externe adviseurs onderzocht kunnen worden met de veronderstelling dat externe adviseurs naar het congres komen voor 'nieuwigheden' en preventiemedewerkers en interne adviseurs voor meer praktische toepassingen. Een en ander zou kunnen blijken uit de toelichtingen die respondenten op de vragen hebben gegeven.

Ik heb tevens een inkijk willen geven in hoe met gegevens van enquêtes wordt omgesprongen. Zeker als er veel vragen worden gesteld, is het lastig om een goed overzicht te krijgen van de gegevens. Het construeren van schalen via PCA of FA ligt dan voor de hand. Bij Factor Analyse is er sprake van het identificeren van 'latente factoren' en dan is het raadzaam om vanuit een bepaalde theorie te werken die dergelijke latente factoren vooronderstelt. De theorie is dan leidend bij het zoeken en benoemen van factoren.

Bij de constructie van schalen is het gebruikelijk om de interne consistentie van de schalen te rapporteren; dit wordt ook wel de betrouwbaarheid genoemd. Met andere woorden, vormen de items (de vragen) waaruit de schaal bestaat één samenhangend geheel? Om dit te toetsen wordt meestal Cronbach's alpha gebruikt. Cronbach's alpha voor bovenstaande schalen is,



Figuur 4 — Gemiddelde waardering voor avondprogramma NVVK-congres tussen preventiemedewerkers, intern, extern en zelfstandig gevestigd adviseurs

respectievelijk, .814, .749 en .582. De alpha voor Schaal 3 is wat aan de lage kant, maar deze schaal kent ook maar twee vragen.

Het meetniveau van de gegevens legt beperkingen op aan wat er met de data kan worden gedaan, dat wil zeggen: hoe de data gemanipuleerd, geanalyseerd en gerapporteerd kunnen worden. Met bovenstaande gegevens kunnen bijvoorbeeld geen uitspraken worden gedaan in de trant van 'Preventiemedewerkers waarden het congres tweemaal zoveel als adviseurs', wel 'Preventiemedewerkers waarden het congres meer dan adviseurs'. Feitelijk kan er ook geen gemiddelde berekend worden over ordinale variabelen, de mediaan is dan een betere maat. Voor het gemak heb ik voor bovenstaande toetsen wel het gemiddelde gebruikt.

Tenslotte: het congres van 2015 is door bijna 77% van de respondenten als 'goed' tot 'zeer goed' beoordeeld. Ruim 85% vond dat het congres bij de dagelijkse praktijk aansloot en 89% van de respondenten had 'waar voor zijn geld' gekregen. Mooie cijfers.

De congrescommissie gaat met deze en andere resultaten aan de slag om in 2017 weer een interessant en gevarieerd congres voor de leden te presenteren.

Frank Guldenmund is universitair docent bij de Sectie Veiligheidskunde van de Technische Universiteit Delft. Frank maakt tevens deel uit van het hoofdbestuur van de NVVK, het Wetenschappelijk Bureau, de redactie van NVVKinfo en de Congrescommissie.