

MA-scriptie Niek Gorris

Themaonderwerp: Overtuigen met de mening van anderen.

Nederlandse titel: Een Bayesiaans perspectief op argumentkwaliteit: een vergelijking tussen het autoriteitsargument en het *ad populum*-argument en de invloed daarop van numerieke voorkeur en wiskundige vaardigheden.

English title: A Bayesian perspective on argument quality: a comparison between the authority argument and the *ad populum*-argument and the influence of numerical preference and mathematical skills.

Radboud Universiteit



Student: Niek Gorris (s4613082)
Begeleider: Dr. J.M.A. Hornikx
Tweede beoordelaar: Dr. A.P.J.V. van Hooft
Universiteit: Radboud Universiteit Nijmegen
Opleiding: Communicatie- en informatiewetenschappen
Aantal woorden: 12.954
Datum: 21-6-2018

Samenvatting

Dagelijks worden wij geconfronteerd met informatie en meningen van anderen, bijvoorbeeld van een individuele expert of van kleine of grote groepen leken, die onze eigen mening en besluitvorming beïnvloeden. In overeenstemming met deze twee typen bronnen, onderscheidt de argumentatietheorie respectievelijk het autoriteitsargument en het *ad populum*-argument. Alhoewel deze argumenten individueel uitgebreid onderzocht zijn, werden ze pas heel recentelijk voor het eerst in één wetenschappelijk studie met elkaar vergeleken. Hornikx, Harris en Boekema (2018) lieten in deze studie proefpersonen bij verschillende situaties beoordelen hoeveel meningen van leken er nodig waren om op te wegen tegen de mening van een individuele expert. Het onderhavige onderzoek bouwde voort op deze studie en ging na in hoeverre de numerieke voorkeur (bevraagd op basis van Viswanathan, 1993) en de wiskundige vaardigheid (getoetst aan de hand van Zillmann, Callison & Gibson, 2009) van respondenten van invloed zijn op de inschattingen die ze maken van het aantal benodigde meningen van leken. Een aantal van de situaties uit de eerdere studie werden gerepliceerd en er werd eveneens gebruik gemaakt van de Bayesiaanse voorspelling. Dit nieuwe onderzoek biedt echter minimale ondersteuning voor de verwachtingen rondom het replicatieonderzoek en de hypothesen aangaande numerieke voorkeur en wiskundige vaardigheid konden niet overtuigend geverifieerd of gefalsificeerd worden. Opmerkelijk genoeg werd zelfs een omgekeerde trend waargenomen, waarbij mensen met een relatief lage numerieke voorkeur of wiskundige vaardigheid de inschattingen beter in lijn met de Bayesiaanse voorspelling bleken te maken dan mensen met relatief hoge numerieke voorkeur of wiskundige vaardigheid. Verder onderzoek is nodig om meer concrete inzichten te verkrijgen en sterkere conclusies te trekken. In de discussie zijn voorstellen gedaan voor dergelijk onderzoek.

Woord vooraf van de auteur

Nijmegen, 21 juni 2018

Beste lezer,

Voor u ligt de verslaglegging van mijn MA-scriptie met als themaonderwerp *overtuigen met de mening van anderen*. Deze rapportage dient als bewijsvoering voor de Radboud Universiteit te Nijmegen, waarin ik laat zien aan alle eindtermen te voldoen om de master Communicatie en Beïnvloeding succesvol af te ronden.

Dit rapport vergelijkt twee argumentvormen, namelijk het autoriteitsargument en het *ad populum*-argument, en bouwt voort op eerder onderzoek van Hornikx, Harris en Boekema (2018). Nieuw binnen dit onderzoek is dat zowel de invloed van numerieke voorkeur als wiskundige vaardigheid op deze argumentvormen werd beoordeeld.

Eerder opgedane kennis, tijdens cursussen uit de pre-master Communicatie- en Informatiewetenschappen, bleek van grote waarde. Graag wil ik in het bijzonder H.W.M. Giesbers bedanken omdat hij mij heeft geënthousiasmeerd voor dit onderwerp. Eerst als docent bij de cursus Argumentatietheorie en daarna als begeleider van mijn BA-scriptie met als themaonderwerp *argumenten voor het goede doel*.

Deze eerder opgedane kennis heb ik tijdens dit scriptietraject verder kunnen verbreden. Allereerst verdient daarvoor dr. A.P.J.V. van Hooft mijn dank voor zijn kritische blik op het onderzoeksvoorstel. Tot slot ben ik dr. J.M.A. Hornikx heel bijzonder dankbaar voor zijn waardevolle begeleiding, zijn hulp bij diverse statistische evaluaties en de nieuwe vakinhoudelijke informatie die hij mij regelmatig aanreikte.

Ik ben erg blij en bovendien trots dat ik mijn aandachtsgebied heb kunnen versterken en middels het voltooien van deze scriptie ook de laatste stap heb kunnen zetten naar het afstuderen binnen het profiel van *marketingcommunicatie*.

Als lezer wens ik u veel plezier met het lezen van mijn scriptie.

Met vriendelijke groet,

Niek Gorris

Inleiding

Stel, als consument ben je op zoek naar een wasmachine. Zelf heb je weinig verstand van witgoed, maar via Coolblue kom je een optie van Miele tegen. Een technische deskundige in witgoed praat je bij over de diversiteit van programma's, de laadcapaciteit, de geoptimaliseerde mechanische trommelbeweging en de overige functies die zorgen voor een uitstekende was- en droogprestatie. Het lijkt een degelijke optie. Vervolgens bekijk je de gemiddelde beoordeling die eerdere gebruikers het apparaat hebben gegeven en lees je enkele van hun reviews terug. Op een schaal van vijf punten beoordeelden zij de wasmachine slechts met twee sterren. Verder lees je dat de trommel veel lawaai maakt bij het centrifugeren en dat zwarte was regelmatig wit uitgeslagen zeepresten vertoont. Wiens raad neem je het meest ter harte? Luister je naar de deskundige van Coolblue en bestel je deze wasmachine of volg je de menigte gebruikers en besluit je door te zoeken naar een andere wasmachine?

Welk argument de koper het meest beïnvloedt in zijn keuze valt binnen het aandachtsgebied van de argumentatietheorie, welke een onderscheid maakt tussen het autoriteitsargument en het *ad populum*-argument. Bij het autoriteitsargument vormen individuen hun oordeel op basis van de mening van een expert (Walton, Reed & Macagno, 2008), in dit geval de ervaringsdeskundige van Coolblue. Bij het *ad populum*-argument vormen individuen hun oordeel op basis van eerdere meningen van een groep anderen (Walton et al., 2008), in dit geval de recensies en de gemiddelde beoordeling die eerdere gebruikers achter lieten.

In alledaagse oordeelvorming laten mensen zich vaak beïnvloeden door anderen. Bijvoorbeeld door een ervaringsdeskundige op televisie, een beroemdheid in een tijdschrift of door vrienden en familie (Friedman & Friedman, 1979; Erdogan, 1999; Hoeken, Hornikx & Hustinx, 2012). Mensen lijken in dit soort gevallen overstag te gaan doordat anderen al een mening hebben gevormd. De invloed van deze bronnen is bekeken vanuit verschillende disciplines, zoals de sociale psychologie (Harkins & Petty, 1987), de cognitieve psychologie (Hahn, Harris & Corner, 2009) en de consumentenpsychologie (Freling & Dacin, 2010).

Er is veel kennis beschikbaar over de onderliggende acceptatieprocessen, de werking en het effect van dit soort bronnen (Petty, Cacioppo & Goldman, 1981; Maheswaran & Chaiken, 1991; Petty & Cacioppo, 1986; Eagly & Chaiken, 1993). Veruit het meeste onderzoek spitst zich toe op de werking van één van de genoemde bronnen. Zo werd de typische consument (Brock, 1965), de expert (Petty et al., 1981), de beroemdheid (Petty, Cacioppo & Schumann, 1983) en de grote groep (Maheswaran en Chaiken, 1991) al individueel onderzocht.

Friedman en Friedman (1979) waren een van de eersten die dit soort bronnen met elkaar vergeleken. Zij vergeleken de overtuigingskracht van een boodschap bij drie soorten producten, afhankelijk van of ze werden aangeprezen door een beroemdheid, een expert of een typische consument. Er bleek inderdaad een significant verband te bestaan tussen het aangeprezen product en het soort bron. Zo behaalde een beroemdheid het beste resultaat bij het aanprijzen van juwelen, een expert bij een stofzuiger en de typische consument bij koekjes.

Vreemd genoeg is er nog weinig theorie beschikbaar over een vergelijking tussen het autoriteitsargument en het *ad populum*-argument. Recentelijk kwam daar verandering in doordat de twee argumentvormen met elkaar werden vergeleken door Hornikx et al. (2018). Zij lieten proefpersonen oordelen over situatieschetsen of scenario's met daarin telkens één expert met een bepaalde mening die afweek ten opzichte van de mening van een groep minder deskundige leken. Proefpersonen werden in de scenario's geïnformeerd aan de hand van percentages, die gingen over de deskundigheid van de expert (variërend van 75.00% t/m 99.99%) en de deskundigheid per individueel groepslid van de groep minder deskundige leken (variërend van 51.00% t/m 60.00%). Op deze manier ontstonden er vijf versies van vijf scenario's met telkens een wisselende percentagecombinatie tussen de expert en de groep leken. Aan het eind van ieder scenario kregen de proefpersonen steeds dezelfde numerieke inschattingsvraag, namelijk: 'Hoe groot moet, volgens u, de groep leken zijn om op te wegen tegen de mening van de expert?'. Met behulp van het Bayesiaanse voorspellingsmodel hadden de auteurs voor alle 25 scenario's normatief berekend hoe groot de groep leken waarschijnlijk zou zijn. Op basis van de getoetste scenario's met het autoriteitsargument en het *ad populum*-argument, bleken de respondenten de groepsgroottes vrij goed in lijn met de Bayesiaanse voorspelling in te schatten.

Waar Hornikx et al. (2018) op basis van hun onderzoek konden concluderen dát mensen in staat zijn de twee argumentvormen met elkaar te kunnen vergelijken en hier bovendien vrij goed in zijn, bleef de vraag waarom mensen hier vrij goed in zijn onbeantwoord. In het stimulusmateriaal van Hornikx et al. (2018) werden proefpersonen blootgesteld aan numerieke gegevens in de vorm van percentages en bovendien dienden zij steeds een numerieke inschatting te maken aangaande de groepsgroottes. Het lijkt voor de hand te liggen dat de mate van iemands numerieke voorkeur en wiskundige vaardigheid invloed heeft op hoe goed hij of zij in staat is om deze numerieke inschattingen te maken. Mensen verschillen namelijk in de attitude die zij hebben ten opzichte van het rekenen met numerieke gegevens (Viswanathan, 1993) en in wiskundige vaardigheden (Zillmann et al., 2009). Het is denkbaar dat iemand die het leuk vindt om met getallen te werken (dus een hoge numerieke voorkeur heeft) en hier

bovendien goed in is (een hoge wiskundige vaardigheid heeft), beter in staat is om in lijn met de voorspelling de benodigde groeps groottes in te schatten dan iemand die het minder leuk vindt om met getallen te werken (lage numerieke voorkeur) en er bovendien minder goed in is (lage wiskundige vaardigheid).

Het is van maatschappelijk belang om te onderzoeken of de mate van iemands numerieke voorkeur en wiskundige vaardigheid bepalend is voor hoe goed hij of zij de benodigde groeps grootte correct in lijn met de voorspelling inschat. Op deze manier draagt dit onderzoek bij aan het verder ontwikkelen van inzichten in hoe communicatie en argumentatie invloed hebben op meningen en beslissingen van individuen en hoe deze inzichten kunnen helpen in het effectief toepassen ervan in marketingcommunicatie processen en uitingen.

Om verder voort te bouwen op het vergelijkingsonderzoek dat Hornikx et al. (2018) zijn gestart tussen het autoriteitsargument en het *ad populum*-argument, is hun onderzoek in het onderhavige onderzoek deels gerepliceerd en daarnaast verder uitgebouwd door zowel de rol van numerieke voorkeur als die van wiskundige vaardigheid binnen een studie te beoordelen. In het onderstaande theoretische kader worden allereerst beide argumentvormen afzonderlijk in meer detail besproken, om vervolgens de overtuigingskracht van beide type te vergelijken. Dan volgt er uitleg en theoretische onderbouwing van de gebruikte Bayesiaanse benadering en tenslotte worden de individuele verschillen in numerieke voorkeur en wiskundige vaardigheid besproken.

Theoretisch kader

Gezien het feit dat het autoriteitsargument en het *ad populum*-argument tot op heden voornamelijk apart zijn beschouwd in de literatuur, is er weinig bekend over de daadwerkelijke relatie tussen deze twee argumentvormen. Zo valt in argumentatieonderzoek de invloed van de meningen van de leken onder ‘een beroep doen op de publieke opinie’ (Godden, 2008; Walton, 1999; Wreen, 1993), terwijl de mening van een expert valt onder ‘een beroep doen op de opinie van een deskundige’ (Walton, 1997). In de op schema gebaseerde benadering van argumenten, worden beiden argumenten beschouwd als *source-based arguments* (Walton et al., 2018) en, op enkele uitzonderingen na (zie Wreen, 1993), gezien als misleidende argumentvormen (Godden, 2008; Van Eemeren, Garssen & Meuffels, 2009; Walton et al., 2008). Als de twee argumenten samen worden besproken, worden meestal de verschillen tussen de twee soorten argumenten benadrukt. Naast deze verschillen staan onderstaand ook de overeenkomstige kenmerken van de twee argumentvormen beschreven.

Op bron gebaseerde beroepen (*Source-based appeals*)

Als je ergens weinig of geen verstand van hebt dan ben je afhankelijk van informatie van andere bronnen. Aangezien deze informatie bepalend is voor jouw beeld van de werkelijkheid, dien je jezelf af te vragen hoe aannemelijk de informatie is en hoe betrouwbaar de bron is. De combinatie van de aannemelijkheid van de informatie en de betrouwbaarheid van de bron is namelijk bepalend of mensen bereid zijn deze informatie over te nemen (Bovens & Hartmann, 2003).

Walton et al. (2008) onderscheiden een groep argumenten, zogenaamde *Source-Based Arguments*, waarbij er in het argument wordt gerefereerd naar een bron. De bronnen waarnaar gerefereerd worden, zijn meestal één of meerdere personen die in de positie zijn om iets te weten. Het autoriteitsargument en het *ad populum*-argument zijn zoals eerder vermeld onderdeel van deze groep argumenten.

1) *Het autoriteitsargument*

Bij het autoriteitsargument wordt iemands oordeel gevormd op basis van de mening van een deskundige (Walton et al., 2008). Schellens en Verhoeven (1994, p. 130) formuleren argumentatie op basis van autoriteit als volgt: ‘Als X iets op gebied Y zegt, dan is het doorgaans waar’. Een expert ontleent zijn geloofwaardigheid – en daarmee de kracht van dit argument – aan kennis of ervaring binnen zijn specifieke vakgebied.

De overtuigingskracht van een deskundige bron is al gedurende een lange periode uitgebreid onderzocht. Zo lieten Petty et al. (1981) proefpersonen een tekst lezen met daarin sterke óf zwakke argumenten en een bron die wel óf niet deskundig was. Proefpersonen lieten zich in hun oordeel over de tekst leiden door een eenvoudige beslisregel, namelijk ‘Accepteer eerder het standpunt van geloofwaardige mensen dan van ongeloofwaardige mensen’ (Hoeken et al., 2012, p. 152).

Hoeken et al. (2012) concludeerden op basis van eerder onderzoek (Petty et al. 1981; Hovland, Janis & Kelley, 1953; O’Keefe, 2002) dat deskundigheid (bepaald door expertise en ervaring) en betrouwbaarheid (bepaald door belangeloosheid en oprechtheid) de mate van geloofwaardigheid van een bron beïnvloedt. Ook Wilson en Sherrell (1993) erkenden het belang van de geloofwaardigheid van de bron en vatten veel onderzoek samen. Op basis van een meta-analyse van 114 studies concludeerden zij dat bronnen met een hoge geloofwaardigheid over het algemeen overtuigender zijn dan bronnen met een lage geloofwaardigheid (zie ook Pornpitakpan, 2004).

2) *Het ad populum-argument*

Bij het *ad populum*-argument vormen individuen hun oordeel op basis van de mening van een groep personen (Walton et al., 2008). Deze argumentvorm, ookwel *'the bandwagon fallacy'* genoemd, wordt gezien als de 'populistische drogreden'. Het is een redentatie waarbij een beroep wordt gedaan op de mening van de meerderheid (= populariteit) om te bewijzen dat een stelling waar of onwaar is (Van Eemeren & Snoeck Henkemans, 2011). In de consumentenpsychologie gaat er veel aandacht uit naar de invloed van deze argumentvorm (Freling & Dacin, 2010), meestal bestempeld als de consensusvuistregel. Volgens Hoeken et al. (2012, p. 161): 'Als zoveel mensen zeggen dat het een goed product is, dan zal het dat wel zijn'. Neem het voorbeeld van de Miele wasmachine. Hierbij werkt deze argumentvorm overtuigend doordat elk individu uit de groep eerdere gebruikers een beetje gebruikerservaring met de wasmachine heeft en bovendien volledig belangeloos is: zij zijn namelijk niet gebaat bij het succes of falen van Miele.

Verschillende onderzoeken laten zien dat informatie over het aantal mensen dat een bepaald standpunt onderschrijft bijdraagt aan de acceptatie van dat standpunt (Chaiken, 1987; Maheswaran & Chaiken, 1991). Zo lieten Maheswaran en Chaiken (1991) bijvoorbeeld participanten een advertentie voor een antwoordapparaat beoordelen. In de advertentie stond vermeld de meerderheid (80%) óf de minderheid (20%) van de gebruikers tevreden was met het product. De advertentie waarin de meerderheid van de gebruikers tevreden was bleek meer overtuigend dan de advertentie waarin de minderheid van de gebruikers tevreden was.

Vergelijking van de twee argumenten

Kortom, beide argumentsoorten zijn individueel uitgebreid onderzocht en kunnen dienen als een overtuigend argument om mensen mee te krijgen met een standpunt. Om de kwaliteit van een argument te bepalen wordt er vaak gekeken naar het resultaat: overtuigt het argument iemand wél of niet? Hiermee blijft echter de vraag wát het argument wel of niet overtuigend maakt onbeantwoord (Hornikx, 2013). Om dit te achterhalen moet er gekeken worden naar de intrinsieke kenmerken van het argument. Om het argument zelf als vertrekpunt te nemen zou het argumentatieschema-perspectief belangrijk kunnen zijn (Hornikx, 2013).

Het argumentatieschema-perspectief

Om de kwaliteit van argumentatie te analyseren en te evalueren worden argumentatieschema's gebruikt. Het argumentatieschema is 'de typologische karakterisering van het soort rechtvaardiging of ontcrachting dat in de argumentatie wordt nagestreefd' (Van Eemeren,

Grootendorst & Snoeck Henkemans, 1997, p. 24). Voor ieder type argument is er in de theorie over argumentatieschema's een unieke beschrijving. Voor een correcte evaluatie van de argumentatie is het gebruikelijk dat er kritische vragen (evaluatievragen) aan zijn toegevoegd. Een kritische vraag ('V') helpt bij het vaststellen of een argumentatie geldig is of niet (Hornikx, 2013).

Het argumentatieschema voor het autoriteitsargument ziet er als volgt uit.

Bron E is een expert in vakgebied S die propositie A bevat
E stelt dat propositie A (in vakgebied S) waar (onwaar) is
Het is aannemelijk dat A waar (onwaar) is

De zes bijbehorende kritische vragen staan weergegeven in tabel 1 (Walton, 1997; Walton et al., 2008, p. 200). Onderstaand is ook het argumentatieschema voor het *ad populum*-argument (Walton, 1999, p. 200) te zien en de drie hierbij behorende kritische vragen (Walton, 1989, p. 89) staan wederom weergegeven in tabel 1.

Iedereen (in een bepaalde referentiegroep) accepteert A. Daarom is A waar (of zou je A moeten accepteren).
Iedereen (in een bepaalde referentiegroep) verwerpt A. Daarom is A onwaar (of zou je A moeten verwerpen).

In de bovenstaande argumentatieschema's en hun bijbehorende kritische vragen zijn er zowel verschillen als overeenkomsten te zien. Zo verschillen de kritische vragen bijvoorbeeld in het aantal en de concreetheid waarin deze zijn gesteld. Bij het autoriteitsargument worden namelijk zes evaluatievragen gesteld die een concreet antwoord impliceren. Neem bijvoorbeeld V4 ('is E betrouwbaar als bron?'), deze vraag kan alleen beantwoord worden met 'ja' of 'nee'. Terwijl er bij het *ad populum*-argument slechts drie evaluatievragen worden gesteld, die tevens een wat meer open karakter hebben. Neem bijvoorbeeld V3 ('Welke reden is er om te denken dat het oordeel van de grote meerderheid klopt?'), hier zijn veel meer antwoordmogelijkheden dan 'ja' of 'nee'. Maar nog belangrijker dan deze verschillen in aantal of concreetheid, is de aard van de vragen: géén enkele vraag overlapt.

Toch lijkt er wel overlap te zijn tussen de kenmerken en kritische vragen van de twee verschillende argumenten. Harris, Hahn, Madsen en Hsu (2016) noemen (a) betrouwbaarheid en (b) expertise als kenmerken van een autoriteitsargument. Hornikx (2013) noemt drie kenmerken die van belang zijn voor de kwaliteit van een *ad populum*-argument, namelijk (1) het aantal bronnen, (2) de mate van onderlinge afhankelijkheid van de bronnen en (3) de betrouwbaarheid van de afzonderlijke bronnen. Allereerst valt het op dat bij beide argumenten het kenmerk betrouwbaarheid terugkomt. Neem nogmaals V4 bij autoriteitsargumentatie ('Is E persoonlijk betrouwbaar als bron?'), dit komt vrijwel geheel overeen met het derde kenmerk

dat Hornikx (2013) noemt bij het *ad populum*-argument ('de betrouwbaarheid van de afzonderlijke bronnen'). Neem daarnaast als voorbeeld V5 bij autoriteitsargumentatie ('is A consistent met wat andere deskundige beweren?'). Deze kritische vraag heeft veel weg van het eerste kenmerk dat Hornikx (2013) noemt bij het *ad populum*-argument, namelijk 'het aantal bronnen'. Ondanks dat de twee argumenten in eerder onderzoek veelal individueel bestudeerd en besproken zijn, hebben de argumentvormen dus wel degelijk overlap of tenminste overeenkomsten.

Tabel 1. *Kritische vragen voor de evaluatie van het ad populum-argumentatie en het autoriteitsargumentatie.*

Autoriteitsargumentatie (Walton, 1997, p. 223)		<i>Ad populum</i> -argumentatie (Walton, 1989, p. 89)	
V1	Hoe geloofwaardig is E als bron?	V1	Accepteert de grote meerderheid van de referentiegroep A als waar?
V2	Is E een expert op het gebied van A?		
V3	Wat beweerde E dat A impliceert?	V2	Is er ander relevant bewijs beschikbaar dat het standpunt onderbouwt dat A niet waar is?
V4	Is E persoonlijk betrouwbaar als bron?		
V5	Is A consistent met wat andere deskundige beweren?	V3	Welke reden is er om te denken dat het oordeel van de grote meerderheid klopt?
V6	Is E's bewering gebaseerd op evidentie?		

Nu de argumentatieschema's en hun bijbehorende kritische vragen in kaart zijn gebracht, lijken de ingrediënten aanwezig om een vergelijking tussen de twee argumenten te maken. Helaas is dit lastiger dan gedacht, omdat er recentelijk steeds meer kritiek bestaat op het argumentatieschema-perspectief. Onderstaand wordt deze kritiek samengevat om vervolgens een alternatief te geven hoe de vergelijking wél gemaakt kan worden.

Kritische kanttekeningen op het argumentatieschema-perspectief

De kritiek die bestaat op het argumentatieschema-perspectief wordt helder opgesomd in de beschrijving van Hornikx (2013). Zo plaatst hij vraagtekens bij de samenhang tussen de vragen en het standpunt, het aantal kritische vragen en de samenhang tussen de kritische vragen onderling. Daarnaast vraagt hij zich af of de kritische vragen überhaupt wel goed te beantwoorden zijn doordat de vragen vaak binair zijn gesteld: alleen de antwoorden 'ja' of 'nee' zijn mogelijk. Zie bijvoorbeeld V1 bij autoriteitsargumentatie ('hoe geloofwaardig is E als bron?'). Wanneer ben je volledig overtuigd dat iemand wel of niet geloofwaardig is als bron? Een genuanceerdere inschatting, zoals bijvoorbeeld: deze persoon is 'redelijk' of 'in hoge mate' geloofwaardig lijkt realistischer.

Kortom, er zijn behoorlijk wat kritische kanttekeningen te plaatsen bij het argumentatieschema-perspectief en er is meer overlap tussen de twee argumentatievormen dan het argumentatieschema-perspectief in eerste instantie doet vermoeden. Om door te zoeken naar een manier om argumenten van hoge kwaliteit vast te stellen kan mogelijk een Bayesiaanse benadering uitkomst bieden (Hornikx, 2013).

Een Bayesiaanse benadering van op bronnen gebaseerde beroepen

In verschillende vakgebieden, zoals de besliskunde, statistiek en cognitieve psychologie wordt er steeds meer gebruik gemaakt van de stelling van Bayes. Deze uit de achttiende eeuw stammende benadering maakt kwalitatieve en kwantitatieve voorspellingen mogelijk over hoe overtuigend verschillende argumenten zijn (Bovens & Hartmann, 2003; Hahn et al., 2009; Harris et al., 2016). De stelling kan helpen bij het beantwoorden van de vraag hoe de mening van een expert dient te worden afgewogen tegen de mening van een groep minder deskundige leken.

Hahn et al. (2009) hebben een Bayesiaans perspectief voor de betrouwbaarheid van de bron geschetst. Zij bestudeerden hoe het aantal bronnen en de deskundigheid van deze bronnen, de acceptatie van het standpunt beïnvloeden. In twee experimenten manipuleerden zij de inhoud van het argument en de betrouwbaarheid van de afzender. Er bleek een significant effect op de overtuigingskracht van een argument: hoe betrouwbaarder de bron, des te meer bronnen er werden aangehaald en des te overtuigender bleek de boodschap.

De stelling van Bayes

Het cruciale vertrekpunt van deze stelling is dat mensen hun kennis zien als waarschijnlijkheden in plaats van zekerheden. De stelling gaat uit van conditionele waarschijnlijkheid: de inschatting die iemand maakt van een gebeurtenis hangt af van de mate waarin een andere gebeurtenis zich voordoet (Hornikx, 2013). Op het moment dat er één aspect verandert, veranderen de overige elementen automatisch mee. Hornikx (2013, p. 135) stelt: ‘net zoals er kan worden gesteld dat de waarschijnlijkheid van X afhangt van het optreden van Y, kan er ook worden gesteld dat de waarschijnlijkheid van standpunt X afhangt van argument Y’. Hahn en Oaksford (2006, 2007) brachten aan het licht dat deze stelling ook nuttig kan zijn voor onderzoek betreffende argumentkwaliteit. Voor deze doeleinden is het vervolgens ook al met succes toegepast (zie bijvoorbeeld Hornikx et al., 2018).

Tabel 2. *De stelling van Bayes uitgelegd met het standpunt over de wasmachine. In de stelling is P de waarschijnlijkheid, H het standpunt en E het argument (Hornikx, 2013, p.135).*

$$\begin{array}{ccc}
 \text{oordeel wasmachine zonder argument} & & \text{oordeel wasmachine gegeven wasmachine fijn} \\
 | & & | \\
 \hline
 P(h|e) = \frac{P(h)P(e|h)}{P(h)P(e|h) + P(\neg h)P(e|\neg h)} \\
 | & & | \\
 \text{normatief eindoordeel} & & \text{oordeel wasmachine gegeven wasmachine niet fijn}
 \end{array}$$

Een schematische weergave van de stelling van Bayes is te zien in tabel 2. Laat h de hypothese zijn dat de wasmachine geen goede aanschaf is en e de evidentie dat zo goed als iedereen die een recensie heeft achtergelaten zich negatief uitlaat over de kwaliteit van de machine; P staat voor de waarschijnlijkheid (*probability*) en is altijd een getal tussen de 0 en 1. Volgens de stelling van Bayes bestaat het normatieve eindoordeel – het oordeel dat iemand heeft over het standpunt gegeven het argument – uit drie componenten: (a) iemands voorafgaande overtuiging over het standpunt $P(h)$ voordat er evidentie bekend is, (b) iemands overtuiging over de evidentie $P(e|h)$ gegeven dat het standpunt waar is en (c) iemands overtuiging over de evidentie $P(e|\neg h)$ gegeven dat het standpunt niet waar is. Het oordeel dat de hypothese niet waar is ($P(\neg h)$ zie tabel 2), is altijd gelijk aan $1 - P(h)$.

Deze stelling wordt vaak nader toegelicht aan de hand van *belief updating* (Baratgin & Politzer, 2010). Centraal hierbij is het gegeven dat iemand een oordeel heeft over h en hoe dit oordeel wordt aangepast door de nieuwe kennis verkregen vanuit e : $P(h|e)$. De overtuiging $P(h)$ is een persoonlijke kwestie. Als iemand er bijvoorbeeld heilig van overtuigd is dat de elektronische valuta Bitcoin blijft stijgen (h), dan geldt $P(h) = 1$. Het toevoegen van e zal niet leiden tot een ander oordeel ($P(h|e) = 1$). Ook hoe e en h zich tot elkaar verhouden is afhankelijk van de persoon in kwestie. De ene persoon hecht meer waarde aan de groeps mening van de recensieschrijvers dan de andere persoon. Hun normatieve eindoordeel $P(h|e)$ zal hierdoor verschillend zijn. Wél geldt voor beide partijen dat hun normatieve eindoordeel $P(h|e)$ gebaseerd is op de rechterzijde van tabel 2, namelijk de drie eerder genoemde (deel)componenten a, b en c (Hornikx, 2013).

De meest basale conclusie die op grond van deze theorie getrokken kan worden is dat als $P(h|e)$ (evidentie is waar) groter is dan $P(e|\neg h)$ (evidentie is niet waar), dit resulteert in een verhoging van het geloof in h . Op het moment dat het omgekeerde het geval is en $P(h|e) < P(e|\neg h)$, dan zal het leiden tot een daling in het geloof in h . Stel je bijvoorbeeld een ontmoeting voor

met iemand waarvan je op aan kan (Persoon A) en een ontmoeting met iemand die in jouw ogen een leugenaar is (Persoon B). Informatie dat afkomstig is van persoon A komt geloofwaardiger op ons over, terwijl de informatie dat afkomstig is van persoon B juist minder waarschijnlijk is om waar te zijn (Hahn et al, 2009).

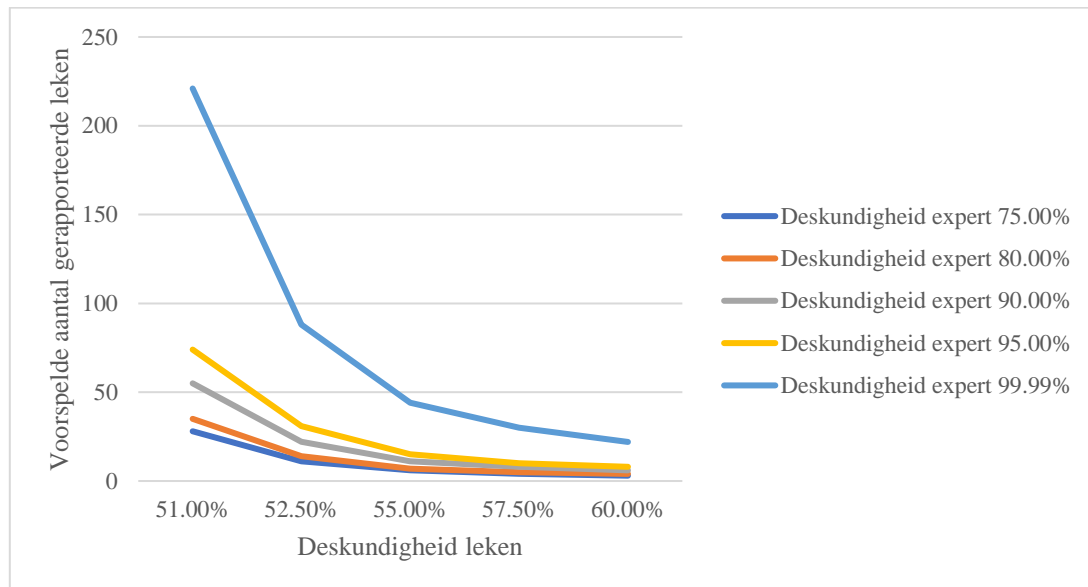
Resultaten van de studie van Hornikx et al. (2018)

Om terug te komen op het eerder besproken onderzoek van Hornikx et al. (2018), waarin er werd bekeken hoeveel meningen van leken er nodig zijn om op te wegen tegen de mening van een expert, volgt hier een uitgebreidere bespreking. Zij vergeleken namelijk voor het eerst, met succes, autoriteitsargument en het *ad populum*-argument met behulp van de stelling van Bayes.

Hornikx et al. (2018) poneerden dat de benodigde groepsgroottes af zouden hangen van de mate van deskundigheid van de expert (meer leken nodig naar mate de expert deskundiger is) en de mate van deskundigheid van de leken (minder leken nodig naar mate de leken deskundiger zijn). In een extreem voorbeeld stellen zij dat wanneer de leken precies even accuraat zouden zijn als de expert, dat de mening van één enkele leek al volstaat om op te wegen tegen de mening van deze expert. Gegeven het feit dat leken per definitie minder accuraat zijn dan experts, ontwierpen Hornikx et al. (2018) stimulusmateriaal opgebouwd uit verschillende scenario's (situatieschetsen) waarbij de groep leken inderdaad telkens minder accuraat (51.00% t/m 60.00%) was dan de betreffende expert (75.00% t/m 99.99%). Alhoewel de leken steeds minder deskundig waren, werd er in de scenario's wel benadrukt dat het een grote groep leken betrof die, onafhankelijk van elkaar, allemaal hetzelfde standpunt onderschreven.

Figuur 1 toont de effecten van vijf verschillende niveaus van leken deskundigheid en vijf verschillende niveaus van expert deskundigheid op het aantal lekenrapporten dat volgens de Bayesiaanse voorspelling nodig was om op te wegen tegen de mening van de expert. Uit Figuur 1 blijkt dat het de verwachting is dat naarmate de expert deskundiger wordt het ingeschatte aantal gerapporteerde leken toeneemt en naarmate de leken deskundiger worden het ingeschatte aantal gerapporteerde leken juist afneemt. Statistische analyse bevestigde dat deze resultaten significant waren (Hornikx et al., 2018, p. 124). Mensen bleken intuïtief best goed in het vergelijken van het autoriteitsargument en het *ad populum*-argument en bovendien maakten ze inschattingen die redelijk in lijn met deze bovenstaande Bayesiaanse voorspellingen waren.

Figuur 1. *Voorspelde aantal gerapporteerde leken dat nodig is om op te wegen tegen de expert in vijf verschillende condities van deskundigheid op basis van de Bayesiaanse voorspelling (zie Figuur 1 in Hornikx et al., 2018, p. 122).*



Hornikx et al. (2018) gaven aan dat de door hun uitgevoerde studie een explorerende studie was op een gebied waar nog relatief weinig onderzoek was uitgevoerd. Daarom was het nuttig dat er een replicatie zou worden uitgevoerd om hun bevindingen te verifiëren. Met het geschetste theoretische kader en de onderzoeksresultaten van Hornikx et al. (2018) in het achterhoofd, kunnen de volgende hypothesen aangaande het replicatieonderzoek worden gesteld (geïnspireerd op de onderzoeksvragen van Hornikx et al., 2018):

- H1: Naarmate de deskundigheid van de expert toeneemt, rapporteren proefpersonen dat er meer leken nodig zijn in de groep die opweegt tegen de mening van de expert.
- H2: Naarmate de deskundigheid van de groep leken toeneemt, rapporteren proefpersonen dat er minder leken nodig zijn in de groep die opweegt tegen de mening van de expert.
- H3: De daadwerkelijke inschattingen van proefpersonen vertonen hetzelfde patroon als de Bayesiaanse voorspellingen.

Hornikx et al. (2018) onderzochten nog een vierde onderzoeksvraag, waarbij werd beoordeeld of de inschattingen van mensen aangaande de benodigde groepsgroottes de Bayesiaanse interactie weerspiegelden tussen de expertdeskundigheid en de lekendeskundigheid. Deze vraag valt echter buiten de scope van dit onderzoek, omdat het voor de onderliggende studie voldoende is om te weten of de inschattingen van mensen kwantitatief in lijn zijn met de Bayesiaanse voorspellingen (H3).

Individuele verschillen in de omgang met numerieke informatie

Proefpersonen van de studie van Hornikx et al. (2018) werd gevraagd om numerieke inschattingen te maken gebaseerd op het lezen van scenario's. In de scenario's lazen proefpersonen situaties waarbij scenario's verschilden in numerieke informatie in de vorm van percentages. Bovendien dienden proefpersonen aan het eind van elk scenario een numerieke inschatting te maken omtrent de door hen verwachte groepsgrootte van leken om op te wegen tegen de expert.

Het ligt voor de hand dat de manier waarop iemand omgaat met de gepresenteerde numerieke informatie van invloed kan zijn op de uiteindelijke kwantitatieve schatting die hij of zij per scenario maakt. Mensen zijn verschillend. Mensen verschillen in bijvoorbeeld numerieke voorkeur (Viswanathan, 1993): de ene persoon vindt het leuker om met getallen te werken dan de andere persoon. Bovendien verschillen mensen in wiskundige vaardigheid (Zillmann et al., 2009): de ene persoon is meer bekwaam met rekenen dan de andere persoon. Het is mogelijk dat de mate waarin iemand het prettig vindt om met getallen te werken en de mate waarin iemand goed is in het werken met getallen, van invloed zijn op hoe goed iemand de numerieke inschattingen bij de scenario's in lijn met de Bayesiaanse voorspelling maakt.

1) Numerieke voorkeur

Viswanathan (1993) stelt dat in veel onderzoek naar numerieke voorkeur de primaire focus ten onrechte ligt op iemands vermogen om praktische, wiskundige problemen op te lossen, zoals bijvoorbeeld 'Als vijf kerstkaarten samen 65 cent kosten, hoeveel kost elke kaart?'. Terwijl volgens hem ook iemands attitude ten opzichte van numerieke informatie een belangrijke rol speelt bij numerieke voorkeur. Zijn publicatie gaat dus meer over iemands gevoelens, zoals zin of tegenzin, bij het omgaan met getallen dan over iemands talent voor getallen.

Om de attitude van iemand ten opzichte van numerieke informatie te meten, zijn er eerder al enkele schalen ontwikkeld. Zo ontwikkelde Aiken (1974) een schaal om de attitude ten opzichte van wiskunde weer te geven en stelde Wise (1985) een schaal op om de attitude ten opzichte van statistiek te meten. Geïnspireerd door hun werk, onderzocht Viswanathan (1993) individuele verschillen in voorkeuren voor numerieke informatie. In zijn bijlage definieert hij *Preference for Numerical Information (PNI)* als 'de neiging om numerieke informatie te gebruiken en zich bezig te houden met denken en betrekken van numerieke informatie' (Viswanathan, 1993, p. 741). Centraal voor de mate van voorkeur voor numerieke informatie stelt hij dus de neiging naar het verkrijgen van numerieke informatie en denken in termen van numerieke informatie.

Op basis hiervan ontwikkelde Viswanathan (1993) een nieuwe schaal die de individuele numerieke voorkeur van personen meet. Voorkeur voor numerieke informatie werd geoperationaliseerd met behulp van een uiteenlopende reeks elementen of aspecten, zoals de mate waarin mensen genieten van numerieke informatie (bijvoorbeeld 'Ik geniet van werk dat het gebruik van getallen vereist'), het leuk vinden van numerieke informatie (bijvoorbeeld 'Ik vind het niet leuk om over problemen met getallen na te denken') en de gepercipieerde behoefte aan numerieke informatie (bijvoorbeeld 'Ik vind dat meer informatie in een numerieke vorm beschikbaar moet zijn'). Andere aspecten waren bruikbaarheid, belangrijkheid, gepercipieerde relevantie, tevredenheid en aandacht of interesse (voor voorbeelden hierbij zie Viswanathan 1993, p. 743). Bovenstaande deelonderwerpen werden bevraagd in 20 items. Bij het selecteren van deze items heeft Viswanathan (1993) geprobeerd om de informatie zo algemeen mogelijk te houden, waardoor de ontwikkelde schaal in een brede context toepasbaar zijn.

Waar Hornikx et al. (2018) op basis van hun onderzoek konden concluderen dát mensen in staat zijn om autoriteitsargument en *ad populum*-argument met elkaar kunnen vergelijken en intuïtief vrij goed inschattingen blijken te maken, gaat het onderhavige onderzoek dieper in op de vraag waarom mensen hier vrij goed in zijn. Door de aanwezigheid van numerieke informatie in de scenario's was de verwachting dat mensen die het leuk vinden om met getallen te werken hun inschattingen beter in lijn met de voorspellingen maken dan mensen met een lagere mate van numerieke voorkeur. Om de invloed van Numerieke voorkeur te beoordelen wordt dan ook de volgende hypothese getoetst:

H4: Proefpersonen met relatief hoge numerieke voorkeur schatten beter in lijn met de Bayesiaanse voorspellingen het aantal leken in dat nodig is om op te wegen tegen de mening van een expert, dan de proefpersonen met relatief lage numerieke voorkeur.

2) *Wiskundige vaardigheden*

Mensen nemen heel snel informatie op die in getallen is aangeboden (Viswanathan & Childers, 1996). Dit is ook niet gek want mensen worden hier voortdurend aan blootgesteld. Neem bijvoorbeeld producten in een supermarkt, de verbale productinformatie gaat altijd gepaard met numerieke informatie zoals de prijs of voedingswaarde. Om in de hedendaagse samenleving te slagen – of om minimaal te functioneren – wordt naast linguïstische kennis ook een bepaalde mate van bekwaamheid met cijfers verwacht (Zillmann et al., 2009). Dit wordt ook wel aangeduid als *numeracy* (= gecijferdheid), een concept dat is ontwikkeld door Lipkus, Samsa en Rimer (2001).

Zillmann et al. (2009) bestudeerden individuele verschillen in het omgaan met getallen uit nieuwsberichten. Zij ontwikkelden een schaal om *arithmetic aptitude* (= rekenkundige bekwaamheid) te meten. Hierbij was zowel aandacht voor numerieke informatie als ook herinnering van numerieke informatie uit deze nieuwsberichten doorslaggevend voor de score van rekenkundige bekwaamheid. Personen die hoog scoorden op rekenkundige bekwaamheid herinnerden zich de genoemde frequenties en ratio's uit de nieuwsberichten beter dan personen die lager scoorden op rekenkundige bekwaamheid. Dat personen die lager scoorden op rekenkundige bekwaamheid minder waarde hechtten aan een correcte herinnering van de numerieke informatie verklaarden Zillmann et al. (2009) doordat zij waarschijnlijk minder aandacht voor numerieke informatie hadden en wellicht hun beoordeling meer op basis van intuïtie hadden gemaakt.

In de zoektocht naar antwoorden op de bevindingen dat a) de daadwerkelijke inschattingen van proefpersonen hetzelfde patroon vertonen als de Bayesiaanse voorspellingen (Horninkx, 2018) en b) wellicht mensen vrij goed intuïtief inschattingen blijken te maken (Zillmann et al., 2009) kan ook de wiskundige vaardigheid van proefpersonen invloed hebben. Vandaar dat de laatste hypothese van de huidige studie is:

H5: Proefpersonen met relatief hoge wiskundige vaardigheden schatten beter in lijn met de Bayesiaanse voorspellingen het aantal leken in dat nodig is om op te wegen tegen de mening van een expert, dan de proefpersonen met relatief lage wiskundige vaardigheden.

Methode

Proefpersonen kregen vijf scenario's waarin telkens één deskundige voorkwam met een tegengestelde mening ten opzichte van een groep leken. In deze scenario's werden proefpersonen geïnformeerd aan de hand van percentages. Deze percentages die gingen zowel over de deskundigheid van een bepaalde expert (variërend van 75.00% t/m 99.99%) als de deskundigheid van een bepaalde groep leken (variërend van 51.00% t/m 60.00%) in een specifieke situatie. Aan het eind van ieder scenario kregen de proefpersonen telkens dezelfde inschattingvraag, namelijk: 'Hoe groot moet, volgens u, de groep leken zijn om op te wegen tegen de mening van de expert?'.

Dit eerste gedeelte van het onderzoek was gebaseerd op de onderzoeksmethode van Hornikx et al. (2018). Vervolgens kregen de proefpersonen een vragenlijst gericht op numerieke voorkeur (gebaseerd op Viswanathan, 1993) en een rekentoets gericht op wiskundige vaardigheid (gebaseerd op Zillmann et al., 2009).

Materiaal

Het stimulusmateriaal bestond, net als bij de studie van Hornikx et al. (2018), uit vijf scenario's waarin telkens één deskundige voorkwam naast een groep van, uiteraard minder technisch deskundige maar mogelijk tot op zekere hoogte ervaringsdeskundige leken. Elk scenario had een ander thema of onderwerp, te weten 1) transport in de stad, 2) dijken versterken, 3) rekeningrijden op de A325, 4) autovrije zones en 5) modern windmolenpark.

De scenario's hadden steeds de vorm van een dialoog tussen persoon A, persoon B en persoon C (geïnspireerd op eerder materiaal aangaande Bayesiaanse argumentatie van o.a. Oaksford & Hahn, 2004; Harris, Hsu en Madsen, 2012; Hornikx et al., 2018), waardoor het leek op een alledaags gesprek. De opbouw van de dialoog was in elk scenario hetzelfde. Het gesprek begon doordat persoon A vraagt of persoon B gehoord heeft van een plan en wat zijn mening hierover is. Persoon B antwoordt dat hij van het plan heeft gehoord en dat hij het een goed plan vindt. Nadat persoon A vraagt waarom hij deze mening heeft, onderbouwt persoon B zijn standpunt door te refereren aan een mening van een expert. Op dit punt wordt de eerste onafhankelijke variabele 'deskundigheid van de expert' gemanipuleerd, namelijk door het toevoegen van de zin: *'en hij kan het weten omdat hij in ...% van de gevallen gelijk heeft als het om dit onderwerp gaat'* (percentages variëren hierbij van 75.00% t/m 99.99%). Vervolgens gaat persoon C deelnemen aan de conversatie, omdat hij de mening van een groep leken heeft gehoord en deze mening tegengesteld is aan de mening van de expert. Op dit punt wordt de tweede onafhankelijke variabele 'deskundigheid van de leken' gemanipuleerd, namelijk door het toevoegen van de zin: *'de leken zijn minder deskundig over dit onderwerp; over dit onderwerp heeft elke leek in ...% van de gevallen gelijk, echter is het wel een grote groep leken die dit beweert.'* (percentages variëren daarbij van 51.00% t/m 60.00%).

Aan het einde van elke dialoog werden de proefpersonen gevraagd om een numerieke inschatting te maken van het benodigde aantal leken in de groep om op te wegen tegen de verkondigde mening van de expert. Onderstaand is als voorbeeld een dialoog uitgelicht over 'transport in de stad' (scenario 1; versie 1). De manipulatie van de twee onafhankelijke variabelen is in het onderstaande voorbeeld dikgedrukt weergegeven. In Appendix A3 zijn de overige scenario's versie 1 te lezen.

Persoon A zegt: Heb je gehoord dat ze van het kruispunt op de Jachtlaan in Apeldoorn een rotonde willen maken om het verkeer beter te laten doorstromen?

Persoon B zegt: Ja, maar waarom stel je deze vraag?

Persoon A zegt: Ik vraag dit omdat ik niet weet of dit wel een goed plan is.

Persoon B zegt: Ik denk van wel.

Persoon A zegt: En waarom denk je dit dan?

Persoon B zegt: Omdat een wegebouwkundige van de BAM aangeeft dat de rotonde zorgt voor een betere doorstroming van het verkeer. **En hij kan het weten omdat hij in 99.99% van de gevallen gelijk heeft als het om verkeersdoorstroming gaat.**

(Persoon C hoort de conversatie tussen persoon A en persoon B en spreekt vervolgens persoon B tegen.)

Persoon C zegt: Ik heb echter gehoord dat sporters, van een sportschool in Apeldoorn, denken dat de rotonde niet zal zorgen voor een betere doorstroming van het verkeer.

De sporters zijn niet deskundig op het gebied van verkeersdoorstroming; over dit onderwerp heeft elke sporter in 55.00% van de gevallen gelijk, echter is het wel een grote groep sporters die dit beweert.

Hoe groot moet de groep sporters zijn om op te wegen tegen de mening van de wegebouwkundige van de BAM?
..... sporters

Proefpersonen

In totaal hebben 175 Nederlandse proefpersonen (leeftijd: $M = 31.66$, $SD = 15.40$, tussen de 19-80 jaar, 54.1% man) het onderzoek afgerond. Het opleidingsniveau varieerde van lager beroepsonderwijs (3.2%) tot aan wetenschappelijk onderwijs (58.1%). Wetenschappelijk onderwijs vormde tevens de grootste categorie. Negen van de oorspronkelijk 184 proefpersonen zijn uitgesloten. Vier proefpersonen omdat zij méér dan de helft van de vragen niet hadden ingevuld. Bovendien was het onduidelijk of het kleine gedeelte gegevens dat wel was ingevuld, dan wel serieus was ingevuld. Daarnaast werden vier proefpersonen uitgesloten omdat zij de laatste vraag (wat is, volgens u, het doel van het onderzoek?) enigszins correct hadden beantwoord.¹ Omdat het op basis van hun beginscores niet duidelijk is of deze proefpersonen hun eigen numerieke voorkeur en rekenvaardigheid hebben aangepast rekening houdend met het doel van het onderzoek, zijn ook zij uit de dataset verwijderd. Tot slot is één Duitse

¹ Twee voorbeelden van omschrijvingen die te veel raakvlak hadden met het daadwerkelijke onderzoeksdoel:
Omschrijving 1: 'Hoeveel niet-experts er nodig zijn om de mening van een expert te overtreffen zonder de wiskundige berekening te maken. Wellicht dat de rekenkwaliteiten die later getest worden daar invloed op hebben.'
Omschrijving 2: 'Redeneren mensen volgens het bayesiaanse model?'

proefpersoon verwijderd, omdat dit onderzoek Nederlandse scenario's betref en bestemd was voor Nederlandse proefpersonen.

Onderzoeksontwerp

In deze studie werd gebruik gemaakt van een 5 (deskundigheid van de expert: 99.99%, 95.00%, 90.00%, 80.00%, 75.00%) x 5 (deskundigheid van de leken: 60.00%, 57.50%, 55.00%, 52.50%, 51.00%) ontwerp. Elke proefpersoon is aan ieder niveau van de twee onafhankelijke variabelen blootgesteld, dus er is sprake van een binnenproefpersoonontwerp. Omdat het een gedeeltelijke replicatie betreft, is het onderzoekontwerp in deze studie hetzelfde als in Hornikx et al. (2018).

In tabel 3 is het ontwerp van de studie te zien. De deskundigheid van de expert en de deskundigheid van de leken werden gemanipuleerd door proefpersonen te informeren over het percentage waarin deze bronnen gelijk hadden. De deskundigheid van de leken was verschillend per versie. De deskundigheid van de expert was gebonden aan de vijf scenario's: transport in de stad (99.99%), dijken versterken (75.00%), rekeningrijden op de A325 (95.00%), autovrije zones (80.00%) en modern windmolenpark (90.00%). Proefpersonen zagen de scenario's altijd in deze volgorde.

Tabel 3. *De onderzoeksopzet en de deskundigheid van de leken voor elk scenario in de verschillende versies (geïnspireerd op de opzet van Hornikx et al., 2018, p. 123).*

Versie	deskundigheid van de expert				
	99.99%	95.00%	90.00%	80.00%	75.00%
	Transport in de stad (scenario 1)	Rekeningrijden op A325 (scenario 3)	Modern windmolenpark (scenario 5)	Autovrije zones (scenario 4)	Dijken versterken (scenario 2)
1	55.00%	51.00%	60.00%	52.50%	57.50%
2	60.00%	57.50%	52.50%	51.00%	55.00%
3	52.50%	55.00%	51.00%	57.50%	60.00%
4	51.00%	60.00%	57.50%	55.00%	52.50%
5	57.50%	52.50%	55.00%	60.00%	51.00%

Instrumentatie

De afhankelijke variabele is de inschatting van het aantal leken, dat nodig is om even overtuigend te zijn als de expert. De proefpersonen kregen vijf scenario's voorgelegd, waarin de inschatting van het aantal leken steeds letterlijk werd bevestigd: *'Hoe groot moet, volgens u,*

de groep leken zijn om op te wegen tegen de mening van de expert?'. De proefpersonen werden hierbij niet beperkt aan een minimum of maximum van het te noemen aantal (vrije numerieke respons).

Via Google Scholar is er gezocht naar een verkorte schaal om numerieke voorkeur te meten. Het originele format van de schaal van Viswanathan (1993; 20 items), bleek al diverse keren met succes ingekort. Zie bijvoorbeeld de schaal met 11 items bij Ackerman & Beier (2006; $\alpha = .89$), de schaal met 8 items bij Fernandes, Lynch en Netemeyer (2014; $\alpha = .90$) en bij Camanho & Fernandes (2016; $\alpha = .89$). Een *Native Speaker* heeft de gereduceerde vragenlijst met acht Engelse items van Fernandes et al. (2014) en Camanho en Fernandes (2016) vertaald naar het Nederlands. Zo werden bijvoorbeeld de stellingen 'ik vind werk waarbij je met getallen moet werken leuk' en 'het geeft mij voldoening om alledaagse problemen op te lossen waarbij getallen betrokken zijn' bevraagd aan de hand van een 7-punts Likertschaal, van 'zeer mee oneens' tot 'zeer mee eens'. Om concentratieverslapping bij de proefpersonen tegen te gaan zijn de items bij vraag 4, 5 en 7 omgekeerd gesteld en achteraf gehercodeerd. Ook in dit onderzoek was de betrouwbaarheid van de verkorte schaal van Viswanathan (1993), bestaande uit acht items, goed ($\alpha = .82$). De vragenlijst om numerieke voorkeur te meten is bijgevoegd in Appendix A4.

Om wiskundige vaardigheid te meten is de *Arithmetic Aptitude Test* van Zillmann et al. (2009, p. 398) gebruikt. Aan de hand van tien meerkeuze items meet deze schaal hoe capabel de proefpersoon was in het maken van elementaire rekenkundige opgaven, met bijvoorbeeld percentages en waarschijnlijkheden. De moeilijkheidsgraad verschilde per opgave: variërend van gemakkelijke opgaven die basiskennis bevroegen, tot aan moeilijke vragen die vergevorderde vaardigheden toetsten. Als voorbeeld zijn onderstaand twee items uitgelicht (A = relatief makkelijk item, B = relatief moeilijk item; juiste antwoorden zijn dikgedrukt). Wederom werden de items vertaald door de *native speaker*. In Appendix A5 zijn alle tien items te zien.

Stelling: 'De vorst heeft 5/16^e deel van de mandarijnenooft vernietigd.'

Vraag A: Is dit minder, meer of gelijk aan een kwart?

- minder dan een kwart **meer dan een kwart** gelijk aan een kwart

Stelling: 'Van de 2,5 miljoen sparren die jaarlijks in deze bergen sterven, schijnen er 11.212 te sterven door besmetting met de zogenaamde Edelspar kever.'

Vraag B: Hoe groot is de kans voor een willekeurige spar in deze bergen om dood te gaan als gevolgd van Edelspar kever besmetting (geef de kans weer als een getal tussen 0 en 1)?

- 0.045 **0.0045** 0.00045 0.000045 0.0000045
-

Een goed antwoord werd gecodeerd als *juist* (1), alle overige (foutieve) antwoordmogelijkheden werden gecodeerd als *onjuist* (0). Het lijkt voor de hand te liggen om de juiste antwoorden op de tien items bij elkaar op te tellen, om een score te krijgen variërend van 0 (laagste) tot 10 (hoogste). Maar om meer spreiding tussen de proefpersonen te krijgen is er een gewogen ‘rapportcijfer’ voor wiskundige vaardigheid toegekend, op dezelfde manier als Zillmann et al. (2009) dit deden. Allereerst hield dit cijfer rekening met de gokkans, omdat de rekentoets bestond uit meerkeuzevragen. Daarnaast werden de antwoorden ook gecorrigeerd voor de moeilijkheidsgraad. De moeilijkheidsgraad van de vragen werd niet vooraf bepaald, maar vastgesteld op basis van de daadwerkelijk gegeven antwoorden van de proefpersonen. De scores zijn aangepast naar aanleiding van de antwoordpatronen.

De verhouding (N=175) van de juiste antwoorden versus onjuiste antwoorden per testitem waren als volgt: (1) 168/7, (2) 159/16, (3) 166/9, (4) 174/1, (5) 171/4, (6) 165/10, (7) 166/9, (8) 97/78, (9) 107/68, (10) 94/81. De items 1 t/m 7 werden gekenmerkt als relatief eenvoudig; items 8 t/m 10 konden gekenmerkt worden als relatief moeilijk. Voor de tien items verschilden de juiste versus onjuiste antwoorden significant van gelijkheid door binomiale test (voor item 1 t/m 7, $p < .001$; item 9, $p = .004$). Alleen item 8 en 10 waren niet significant ($p = .173$ en $p = .364$, respectievelijk).

Samenvattend beoordeelden proefpersonen dus eerst hun eigen mate van numerieke voorkeur op basis van de schaal van Viswanathan (1993), waarna hun numerieke vaardigheid daadwerkelijk werd getoetst door de *Arithmetic Aptitude Test* van Zillmann et al. (2009, p. 398). Uit een correlatie voor numerieke voorkeur en wiskundige vaardigheid bleek een significant, positief verband te bestaan ($r(175) = .36, p < .001$). Naarmate de proefpersonen aangaven over hogere numerieke voorkeur te beschikken, scoorden zij ook hoger op de wiskundige vaardigheidstest. Als dit niet het geval was geweest, zou dit een reden zijn om de scores op de twee schalen in twijfel te trekken. Maar het feit dat de numerieke voorkeur in dit onderzoek correleert met de wiskundige vaardigheid draagt bij aan de validiteit van de schalen.

Tot slot vulden de proefpersonen hun demografische gegevens in en verklaarden zij wat zij dachten dat het uiteindelijke doel van de studie was. De manier waarop dit bevraagd werd is te zien in Appendix A6.

Procedure

De proefpersonen werden uitgenodigd om deel te nemen aan het online onderzoek via het programma *Qualtrics*. Het onderzoek werd individueel afgenomen en de procedure was voor iedereen gelijk. Na een welkomstboodschap werden de proefpersonen geïnstrueerd aan de hand

van een voorbeeldcasus. De welkomstboodschap en instructie zijn respectievelijk terug te vinden in Appendix A1 en A2. Het doel van het onderzoek werd voorafgaand niet medegedeeld, op specifiek verzoek van enkele proefpersonen heeft de onderzoeker dit achteraf wel toegelicht. Afname van het onderzoek duurde gemiddeld 15 minuten en proefpersonen ontvingen geen beloning voor deelname. Er waren verder geen storende factoren.

Statistische toetsing

Uiteraard is een robuuste procedure voor statistische toetsing essentieel voor elk wetenschappelijk onderzoek om analyse en conclusie goed te onderbouwen. In het onderhavige geval bleek dit een wat lastige puzzel, waarbij de onderzoeker ervoor heeft gekozen om twee ingrepen te doen om zo solide mogelijke bevindingen te kunnen rapporteren.

(1) Ten eerste is het aantal in de analyse betrokken proefpersonen teruggebracht van 175 naar 134. Bij het bekijken van de ruwe data bleek namelijk dat 41 proefpersonen extreem hoge numerieke schattingen ingevuld hadden, die geheel buiten het algemeen geldende patroon van de overige 134 proefpersonen vielen. Onderstaand wordt dit in meer detail besproken en de gekozen oplossing verder toegelicht (zie 1, *Extreem hoge numerieke inschattingen*).

(2) Ten tweede is ervoor gekozen om bij de statistische evaluaties van de proefpersonen met betrekking tot numerieke voorkeur en wiskundige vaardigheid ook niet alle 134 proefpersonen mee te nemen. Om een duidelijkere scheiding te krijgen in proefpersonen met een relatief ‘hoge’ dan wel ‘lage’ numerieke voorkeur en wiskundige vaardigheid, werden door middel van een dubbele mediaansplit de 25% proefpersonen met hoogste voorkeur of vaardigheid afgezet tegen de 25% proefpersonen met de laagste voorkeur of vaardigheid. Onderstaand wordt ook dit verder in detail besproken en toegelicht (zie 2, *Spreiding tussen ‘hoge’ of ‘lage’ numerieke voorkeur en wiskundige vaardigheid*).

Voor beide gevallen geldt dat er werd gestreefd naar een situatie waarin de ruwe data van zoveel mogelijk proefpersonen werden behouden, maar dat deze data wel betekenisvol te interpreteren was.

1. Extreem hoge numerieke inschattingen

De proefpersonen (N=175) werden gevraagd om per scenario een numerieke inschatting te maken over de benodigde groeps grootte. Tegen de verwachting in noemden sommige proefpersonen (N=41) extreem hoge aantallen, namelijk aantallen die soms ver boven de duizend waren. Als deze hoge ‘uitbijters’ allemaal zouden worden meegenomen in de analyse,

dan zouden deze schattingen een enorme invloed hebben op de gemiddelden scores. De onderzoeker was dan ook van mening dat bij het gebruik van alle ruwe data (N=175) deze niet meer betekenisvol te interpreteren zouden zijn, zoals ook te zien is in tabel 4. Als de ruwe data niet te interpreteren zijn, zal ook het trekken van relevante conclusies niet mogelijk zijn. De vraag is dan, bij welke extreme score wordt een proefpersoon het beste niet meer meegenomen, met andere woorden wat is het grensgetal?

Bij het bepalen van het grensgetal is ernaar gestreefd om zo weinig mogelijk proefpersonen te verliezen en daarmee het bestand aan ruwe data zoveel mogelijk intact te houden. De onderzoeker heeft een gewogen en beredeneerde keuze gemaakt voor het grensgetal ‘999’. In de optiek van de onderzoeker is dit een getal waarboven schattingen als extreem en weinig realistisch kunnen worden aangemerkt. Het getal voldoet tevens aan de ambitie om de ruwe data van zoveel mogelijk proefpersonen te behouden in de dataset. Er is bewust niet gekozen voor het getal 1000, oftewel om de grens te leggen bij “tot en met duizend”, maar voor “kleiner dan duizend” (dus: *maximaal* 999). Dit heeft als reden dat het getal ‘1000’ veelvuldig voorkomt in het databestand, en dus expliciet gevolgen heeft voor de normaalverdeling van de responsen/scores.

Tabel 4. *De gemiddelden inschattingen over de benodigde groepsgroottes van alle proefpersonen (N=175) vergeleken met de gemiddelden inschattingen van proefpersonen die aantallen hebben genoemd kleiner dan duizend (N=134).*

		Gemiddelden van de inschatting	
		N = 175	N = 134
		(alle proefpersonen)	(proefpersonen <1000)
Deskundigheid leken	51.00%	7792	132
	52.50%	6393	117
	55.00%	60596	115
	57.50%	7291	120
	60.00%	6347	95
Deskundigheid expert	99.99%	16075	138
	95.00%	1557	119
	90.00%	64175	131
	80.00%	6254	91
	75.00%	355	97

2. *Spreading tussen 'hoge' of 'lage' numerieke voorkeur en wiskundige vaardigheid*

Om een duidelijkere scheiding te krijgen tussen proefpersonen met een 'hoge' of 'lage' numerieke voorkeur en vaardigheid, zijn proefpersonen door middel van een dubbele mediaansplit verdeeld in groepen. De gegevens hierover staan respectievelijk in tabel 5 en 6.

De acht items van Viswanathan (1993) om numerieke voorkeur te meten werden bevraagd op 7-puntsschaal. Als eerste stap werden de proefpersonen (N=134) op basis van hun numerieke voorkeur door middel van een mediaansplit (mediaan = 4.88) verdeeld in twee groepen, namelijk een groep met relatief lage numerieke voorkeur (scoorde < 4.88, gecodeerd als groep 1) en een groep met relatief hoge numerieke voorkeur (> 4.88, gecodeerd als groep 2). Deze enkele mediaansplit was echter nog niet voldoende, omdat in deze situatie proefpersonen die *nét* onder of *nét* boven de mediaan scoorden toch onderdeel gingen uitmaken van de groep met respectievelijk lage numerieke voorkeur (Groep 1) of hoge numerieke voorkeur (Groep 2). Om deze reden werd er gezocht naar een manier om zoveel mogelijk proefpersonen te behouden en toch zorg te dragen voor een zo'n groot mogelijk verschil tussen 'laag' en 'hoog' op de twee relevante schalen.

Om een duidelijkere scheiding te creëren tussen 'laag' en 'hoog', werd er gekozen om de 25% proefpersonen met de laagste numerieke voorkeur af te zetten tegen de 25% proefpersonen met de hoogste numerieke voorkeur. Een mediaansplit bij Groep 1, de groep proefpersonen met relatief lage numerieke voorkeur, wees uit welke proefpersonen tot de laagste 25% behoorden (scoorde < 4.14, gecodeerd als Groep 1A) en welke niet (scoorde > 4.14, gecodeerd als Groep 1B). Een mediaansplit bij Groep 2, de groep proefpersonen met relatief hoge numerieke voorkeur, wees uit welke proefpersonen tot de hoogste 25% behoorden (scoorde > 5.49, gecodeerd als Groep 2B) en welke niet (scoorde < 5.49, gecodeerd als Groep 2A).

De statistische analyses met betrekking tot numerieke voorkeur werden vervolgens uitgevoerd over de proefpersonen die deel uitmaakten van Groep 1A (N=35) en Groep 2B (N=36). Door deze twee groepen met elkaar te vergelijken was het verschil tussen proefpersonen met een lage numerieke voorkeur en hoge numerieke voorkeur voldoende gecreëerd. Bovendien werden de statistische toetsen alsnog gedaan over meer dan de helft van de 134 proefpersonen, namelijk N = 71.

Tabel 5. *Groepsindeling van de proefpersonen (N=134) door middel van een dubbele mediaansplit op basis van op 7-puntsschaal bevraagde numerieke voorkeur (Viswanathan, 1993). Groep 1A (laagste numerieke voorkeur) en Groep 2B (hoogste numerieke voorkeur) zijn gebruikt voor de statistische toetsing.*

	Groep 1 (N = 62)		Groep 2 (N = 72)	
	Groep 1A (N = 35)	Groep 1B (N = 27)	Groep 2A (N = 36)	Groep 2B (N = 36)
Numerieke voorkeur	< 4.14	> 4.14	< 5.49	> 5.49

Op dezelfde manier werden de 134 proefpersonen verdeeld in groepen op basis van hun wiskundige vaardigheid. De proefpersonen hadden een gewogen rapportcijfer (zie *Instrumentatie*) gebaseerd op hun correcte antwoorden op de tien items van de *Arithmetic Aptitude Test* van Zillmann et al. (2009). Op basis van dit rapportcijfer werden de proefpersonen wederom door middel van een mediaansplit (mediaan = 7.16) verdeeld in twee groepen, namelijk een groep met relatief lage wiskundige vaardigheid (scoorde <7.16, gecodeerd als groep 1) en een groep met relatief hoge wiskundige vaardigheid (> 7.16, gecodeerd als groep 2).

Een volgende mediaansplit bij Groep 1, de groep proefpersonen met relatief lage wiskundige vaardigheid, wees uit welke proefpersonen tot de laagste 25% behoorden (scoorde < 4.39, gecodeerd als Groep 1A) en welke niet (scoorde > 4.39, gecodeerd als Groep 1B). Een mediaansplit bij Groep 2, de groep proefpersonen met relatief hoge wiskundige vaardigheid, wees uit welke proefpersonen tot de hoogste 25% behoorden (scoorde > 7.58, gecodeerd als Groep 2B) en welke niet (scoorde < 7.58, gecodeerd als Groep 2A).

De statistische analyses met betrekking tot wiskundige vaardigheid werden daarop uitgevoerd over de proefpersonen die deel uitmaakten van Groep 1A (N=32) en Groep 2B (N=37). Tussen deze twee groepen was het verschil tussen proefpersonen met een lage wiskundige vaardigheid en hoge wiskundige vaardigheid maximaal. Wederom werden de statistische toetsen over meer dan de helft van de 134 proefpersonen gedaan, namelijk N = 69.

Tabel 6. *Groepsindeling van de proefpersonen (N=134) door middel van een dubbele mediaansplit op basis van op 10-puntsschaal bevraagde wiskundige vaardigheid (Zillmann et al., 2009). Groep 1A (laagste wiskundige vaardigheid) en Groep 2B (hoogste wiskundige vaardigheid) zijn gebruikt voor de statistische toetsing.*

	Groep 1 (N = 62)		Groep 2 (N = 72)	
	Groep 1A (N = 32)	Groep 1B (N = 30)	Groep 2A (N = 35)	Groep 2B (N = 37)
Wiskundige vaardigheid	< 4.39	> 4.39	< 7.58	> 7.58

Samenvattend, proefpersonen die op 7-puntsschaal lager scoorden dan 4.14 werden gecodeerd als ‘iemand met een lage numerieke voorkeur’ (N=35) en proefpersonen die hoger scoorden dan 5.49 werden gecodeerd als ‘iemand met een hoge numerieke voorkeur’ (N=36). Proefpersonen die op 10-puntsschaal lager scoorden dan 4.39 werden gecodeerd als ‘iemand met een lage wiskundige vaardigheid’ (N=32) en proefpersonen die hoger scoorden dan 7.58 werden gecodeerd als ‘iemand met een hoge wiskundige vaardigheid’ (N=37). Bij de statistische analyses werden de 35 proefpersonen met lage numerieke voorkeur afgezet tegen de 36 proefpersonen met hoge numerieke voorkeur (N=71) en werden de 32 proefpersonen met lage wiskundige vaardigheid afgezet tegen de 37 proefpersonen met hoge wiskundige vaardigheid (N=69).

Toetsen van hypothesen en onderzoeksvragen

De eerste drie hypothesen werden op dezelfde manier getoetst als in Hornikx et al. (2018). Om H1 en H2 te toetsen werden er variantieanalyses met herhaalde metingen uitgevoerd, met de inschatting omtrent de benodigde groeps grootte als afhankelijke variabelen en de mate van deskundigheid van de expert (75.00% t/m 99.99%) of leken (51.00% t/m 60.00%) als onafhankelijke variabelen. Om H3 te beoordelen en daarmee te beoordelen of er een verband was tussen de Bayesiaanse voorspelde benodigde groeps groottes en de daadwerkelijk ingeschatte benodigde groeps groottes, zijn regressieanalyses uitgevoerd. Tot slot werd er ook aan de hand van regressieanalyses beoordeeld of proefpersonen met hoge numerieke voorkeur, beter in lijn met de Bayesiaanse voorspellingen inschatten dan proefpersonen met lage numerieke voorkeur (H4). En of proefpersonen met hoge numerieke vaardigheid, beter in lijn met de Bayesiaanse voorspellingen inschatten dan proefpersonen met lage wiskundige vaardigheid (H5).

Resultaten

Deskundigheid van de expert

Om de eerste hypothese te toetsen is er gekeken of er een effect bestaat van de vijf verschillende niveaus van deskundigheid van de expert op de daadwerkelijke inschatting van de proefpersonen. Mauchly's test liet zien dat er niet was voldaan aan de assumptie van sphericiteit, daarom zijn de vrijheidsgraden Huynh-Feldt ($\epsilon = .84$) gecorrigeerd. Uit een variantie-analyse met herhaalde metingen bleek er een significant hoofdeffect van deskundigheid van de expert op de daadwerkelijke inschatting van het aantal benodigde leken ($F(3.35, 445.37) = 5.70, p < .001$).

Een paarsgewijze vergelijking (met Sidak-aanpassing) liet zien dat een expert met een deskundigheid van 99.99% niet significant afweek van een expert met een deskundigheid van 95.00% en een expert met een deskundigheid van 90.00% (beide p 's $> .611$), maar wel significant afweek van een expert met een deskundigheid van 80.00% ($p = .024$) en een expert met een deskundigheid van 75.00% ($p = .037$). Bij een expert met een deskundigheid van 99.99% ($M = 138.41, SD = 161.55$) schatten proefpersonen significant meer leken nodig te hebben, dan bij een expert met een deskundigheid van 80.00% ($M = 90.28, SD = 138.42$) of 75.00% ($M = 96.85, SD = 128.75$). Tot slot bleek ook dat een expert met een deskundigheid van 90.00% significant afweek van een deskundigheid van 80.00% ($p = .005$) en van een expert met een deskundigheid van 75.00% ($p = .023$). Bij een expert met een deskundigheid van 90.00% ($M = 130.01, SD = 155.62$) schatten proefpersonen significant meer leken nodig te hebben, dan bij een expert met een deskundigheid van 80.00% ($M = 90.28, SD = 138.42$) of 75.00% ($M = 96.85, SD = 128.75$). De overige vergelijkingen waren niet significant (alle p 's $> .165$).

Deskundigheid van de leken

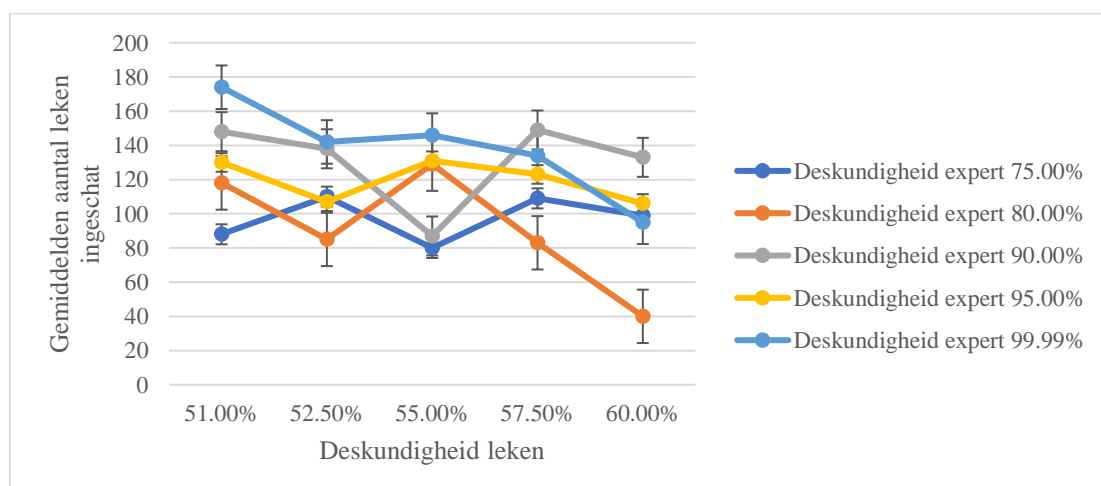
Om de tweede hypothese te toetsen is er gekeken of er een effect bestaat van de vijf verschillende niveaus van lekendeskundigheid op de daadwerkelijke inschatting van de proefpersonen. Mauchly's test liet zien dat er niet was voldaan aan de assumptie van sphericiteit, daarom zijn de vrijheidsgraden Huynh-Feldt ($\epsilon = .95$) gecorrigeerd. Uit een variantie-analyse met herhaalde metingen bleek er een significant hoofdeffect van deskundigheid van de leken op de daadwerkelijke inschatting van het aantal benodigde leken ($F(3.79, 503.41) = 2.51, p = .044$).

Een paarsgewijze vergelijking (met Sidak-aanpassing) liet zien dat een lekendeskundigheid van 51.00% significant afweek met een lekendeskundigheid van 60.00%

($p = .007$). Bij een lekendeskundigheid van 51.00% ($M = 131.33$, $SD = 151.09$) schatten proefpersonen meer leken nodig, dan bij een lekendeskundigheid van 60.00% ($M = 93.37$, $SD = 120.54$). Verder bleek een lekendeskundigheid van 51.00% niet significant af te wijken met een lekendeskundigheid van 52.50%, 55.00% en 57.50% (alle p 's $> .903$). De overige vergelijkingen waren ook niet significant (alle p 's $> .212$).

Er is beperkte ondersteuning voor de verwachtingen. Hypothese 1 en 2 dienden om te bepalen of de inschattingen van proefpersonen kwalitatief in lijn zijn met de Bayesiaanse voorspelling. Figuur 2 liet zien dat, hoewel niet perfect, het algemene patroon van de voorspellingen in figuur 1 lijkt te worden waargenomen. Zo was een van de algemene voorspellingen hier dat hoe hoger het deskundigheidsniveau van de expert, hoe groter de ingeschatte benodigde groepsgrootte zou zijn. In sommige gevallen blijkt dit inderdaad te kloppen. De gemiddelde ingeschatte benodigde groepsgroottes zijn bijvoorbeeld bij een expert met een deskundigheidsniveau van 99.99% structureel groter dan bij een expert met een deskundigheidsniveau van 80.00%. Een andere assumptie uit figuur 1 was dat hoe hoger het deskundigheidsniveau van de leken zou zijn, hoe kleiner de ingeschatte benodigde groepsgrootte zou zijn. Ook dit blijkt gedeeltelijk te kloppen. Met uitzondering van de expert met een deskundigheidsniveau van 75.00% (de lichtblauwe lijn), is iedere gemiddeld ingeschatte benodigde groepsgrootte bij een lekendeskundigheid van 60.00% lager dan bij een lekendeskundigheid van 51.00%. In de tabel in Appendix B staan de gegevens uit Figuur 2 in een vereenvoudigde tabelvorm met daarbij telkens de standaardafwijkingen weergegeven.

Figuur 2. *Gemiddelde inschattingen van het aantal leken dat nodig is om op te wegen tegen de expert verdeeld over de condities. Error bars zijn plus en min 1 standaardfout ($N=134$).*



Kwantitatieve fit met de Bayesiaanse voorspellingen

Doordat de vijf scenario's getoetst werden in vijf verschillende versies, ontstonden er in het databestand 25 waarden. Elke waarde stond voor een andere deskundigheidscombinatie van de expert en de groep leken (zie tabel 3). Om H3 te toetsen dienden, per deskundigheidscombinatie, de kwantitatieve fit tussen de Bayesiaanse voorspelling en de daadwerkelijke antwoorden van proefpersonen te worden beoordeeld.

Hornikx et al. (2018) berekenden aan de hand van de stelling van Bayes de 25 'voorspelde waarden' (zie Figuur 1). Vervolgens werden, per deskundigheidscombinatie, de gemiddelden responsen van proefpersonen berekend. Deze vormden de 'daadwerkelijke waarden'. Uiteindelijk werd het verband geanalyseerd tussen de 25 voorspelde waarden en de 25 daadwerkelijke waarden. Uit een enkelvoudige lineaire regressie bleek dat de daadwerkelijke inschatting van proefpersonen voor 32% te verklaren was door de Bayesiaanse voorspelling ($F(1, 23) = 10.69, p = .003$). De Bayesiaanse voorspelling bleek een significante voorspeller voor de daadwerkelijke inschatting van proefpersonen ($\beta = .56, p = .003$). Tabel 7 geeft de gegevens van de regressieanalyse weer.

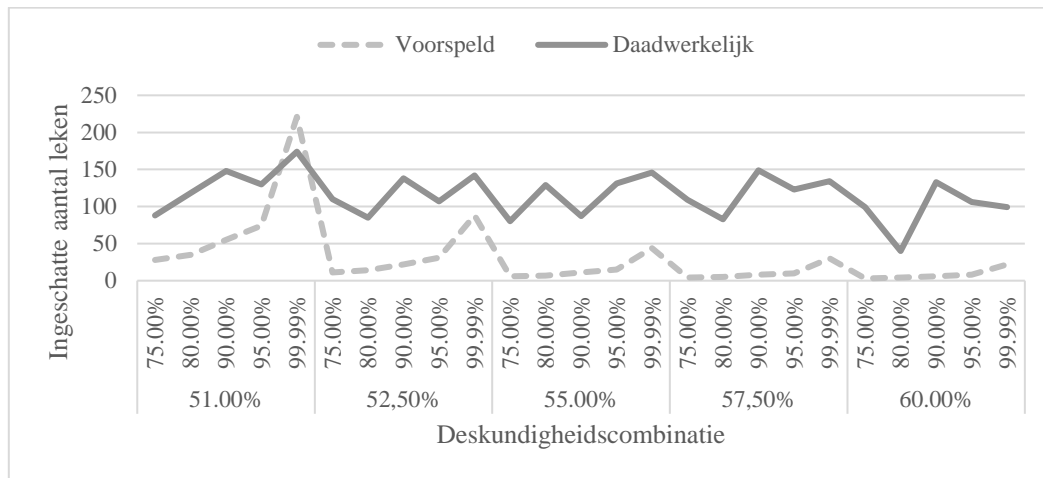
Tabel 7. *Regressie-analyse voor Bayesiaanse inschattingen als voorspeller van daadwerkelijke inschatting (N=25).*

Variabele	B	SE B	β
Intercept	98.65	6.91	
Bayesiaanse voorspelling	.42	.13	.56*
R ²	.29		
F	10.69*		

* $p = .003$.

Het verband tussen de Bayesiaanse voorspelde inschattingen en de daadwerkelijke inschattingen van proefpersonen (N=134) staan weergegeven in figuur 3. Op de x-as zijn de combinaties van deskundigheid van de expert en deskundigheid van de leken te zien. Op de y-as is de Bayesiaanse voorspelde inschatting per combinatie te zien. Allereerst is te zien dat de voorspelde- en daadwerkelijke inschattingen niet exact overeenkomen, over het algemeen liggen de gemiddelden van de daadwerkelijke inschattingen hoger dan de voorspelde inschattingen. Het valt wel op dat bijna overal waar het Bayesiaanse model pieken omhoog verwacht, deze pieken ook volgen in de daadwerkelijke scores. Kanttekening bij dit resultaat is dat er ook meer pieken aanwijsbaar zijn in de daadwerkelijke scores.

Figuur 3. Een grafische weergave van het verband tussen de Bayesiaanse voorspelde inschatting van het benodigde aantal leken en het door de proefpersonen (N=134) inschatte benodigde aantal leken.



Koppeling met numerieke voorkeur en wiskundige vaardigheid

Om de vierde hypothese te toetsen werd gekeken of de inschattingen van proefpersonen met hoge numerieke voorkeur beter in lijn waren met de Bayesiaanse voorspelde inschattingen, dan de inschattingen van proefpersonen met lage numerieke voorkeur. Hiervoor werd het verband geanalyseerd tussen de 25 voorspelde waarden en de 25 gemiddelden van de daadwerkelijke waarden van de proefpersonen met hoge numerieke voorkeur. In tegenstelling tot de verwachting werd er geen significant resultaat gevonden voor proefpersonen met hoge numerieke voorkeur ($F(1, 23) = 2.39, p = .136; \beta = .31$). Wel bleek uit een enkelvoudige lineaire regressie dat de daadwerkelijke inschatting van proefpersonen met lage numerieke voorkeur voor 15% te verklaren was door de Bayesiaanse voorspelling ($F(1, 23) = 5.37, p = .030, \beta = .44$). Tabel 8 geeft de gegevens van de regressieanalyse weer.

Tabel 8. Regressie-analyse voor Bayesiaanse voorspelling als voorspeller van daadwerkelijke inschattingen van proefpersonen met lage of hoge numerieke voorkeur (N=71).

Variabele	Numerieke voorkeur					
	Laag			Hoog		
	B	SE B	β	B	SE B	β
Intercept	108.65	16.36		95.70	10.60	
Voorspelling	.70	.30	.44*	.30	.20	.31
R ²	.15			.06		
F	5.37*			2.39		
	$p = .030^*$			$p = .136$		
	$p = .517$			$p = .003^*$		

Om de vijfde en laatste hypothese te toetsen werd gekeken of de inschattingen van proefpersonen met hoge wiskundige vaardigheid beter in lijn waren met de Bayesiaanse voorspelde inschattingen, dan de inschattingen van proefpersonen met lage wiskundige vaardigheid. Hiervoor werd het verband geanalyseerd tussen de 25 voorspelde waarden en de 25 gemiddelden van de daadwerkelijke waarden van de proefpersonen met hoge wiskundige vaardigheid. In tegenstelling tot de verwachting werd er wederom geen significant resultaat gevonden voor proefpersonen met hoge wiskundige vaardigheid ($F(1, 23) = .43, p = .517, \beta = .14$). Wel bleek, middels enkelvoudige lineaire regressie, dat de daadwerkelijke inschatting van proefpersonen met lage wiskundige vaardigheid voor 30% te verklaren was door de Bayesiaanse voorspelling ($F(1, 23) = 11.22, p = .003, \beta = .573$). Tabel 9 geeft de gegevens van de regressieanalyse weer.

Tabel 9. *Regressie-analyse voor Bayesiaanse voorspelling als voorspeller van daadwerkelijke inschattingen van proefpersonen met lage of hoge wiskundige vaardigheid (N=69).*

Variabele	Wiskundige vaardigheid					
	Laag			Hoog		
	B	SE B	β	B	SE B	β
Intercept	B	SE B	β	76.79	8.00	
Voorspelling	111.95	10.77		.496	.148	.57*
R ²	.13	.20	.14	.30		
F	-.024			11.22*		
	$p = .517$			$p = .003^*$		

Conclusie

Het onderzoek had twee doelen, namelijk replicatie van eerder onderzoek en uitbreiding met persoonlijkheidskenmerken. Het eerste deel was gebaseerd op de studie van Hornikx et al. (2018). In het tweede deel werd de koppeling met numerieke informatie gemaakt. Zo werd de invloed van de mate van numerieke voorkeur (Viswanathan, 1993) en wiskundige vaardigheid (Zillmann et al., 2009) van proefpersonen beoordeeld. De onderzoeksresultaten geven aanleiding tot een aantal conclusies, die zijn opgedeeld in A) conclusies met betrekking tot het replicatieonderzoek en B) conclusies met betrekking tot de koppeling met numerieke voorkeur en wiskundige vaardigheid.

A) Conclusies met betrekking tot het replicatieonderzoek

In dit onderzoek werd minimale ondersteuning gevonden voor de verwachtingen aangaande het replicatieonderzoek. Alhoewel de resultaten minder sterk waren dan bij de studie van Hornikx et al. (2018), werd het verwachte patroon nog steeds waargenomen.

De verwachting was dat naarmate de deskundigheid van de expert toe zou nemen de proefpersonen de benodigde groepsgrootte zouden vergroten (H1) en dat naarmate de deskundigheid van de leken toe zou nemen de proefpersonen de benodigde groepsgrootte zouden verkleinen (H2). Deze hypothesen werden beide door de resultaten van dit onderzoek gedeeltelijk bevestigd. Zowel bij hypothese 1 als hypothese 2, waren er een aantal van de vergelijkingen tussen de vijf groepen significant. Daarnaast was de verwachting dat er een verband zou zijn tussen de daadwerkelijke opgegeven waarden van proefpersonen op de scenario's en de vooraf berekende voorspelde waarden volgens het Bayesiaanse model (H3). Mensen bleken niet in staat om in absolute zin het aantal leken dat nodig is accuraat in te schatten (zie figuur 3), maar in het algemeen komt het patroon van hun inschattingen wel overeen met de voorspellingen die gedaan zijn aan de hand van het Bayesiaanse model. De Bayesiaanse voorspelling bleek een significante voorspeller voor de daadwerkelijke inschatting van proefpersonen. Wat verder opviel was wel dat de gemiddelden van de daadwerkelijke inschattingen beduidend hoger waren dan de voorspelde inschattingen.

B) Conclusies met betrekking tot de koppeling met numerieke voorkeur en wiskundige vaardigheid

In dit onderzoek werd geen ondersteuning gevonden voor de verwachtingen aangaande de koppeling met numerieke voorkeur en wiskundige vaardigheid.

Als eerste werd getoetst of de groep proefpersonen met een relatief hoge numerieke voorkeur beter in lijn met de Bayesiaanse voorspellingen de benodigde groepsgroottes inschatten, dan de groep proefpersonen met een relatief lage numerieke voorkeur (H4). Dit bleek niet het geval te zijn. Tegen de verwachting in werd er geen significant resultaat gevonden voor de groep proefpersonen met relatief hoge numerieke voorkeur, terwijl er wél een significant resultaat gevonden werd voor de groep proefpersonen met relatief lage numerieke voorkeur. Dit wil zeggen dat de groep proefpersonen met relatief lage numerieke voorkeur de groepsgroottes beter in lijn met de Bayesiaanse voorspelling hadden ingeschat, dan dat de groep proefpersonen met relatief hoge numerieke voorkeur dit deden. De verwachte koppeling tussen hoge numerieke voorkeur en het goed schatten van groepsgroottes is dus niet waargenomen, en het is opmerkelijk te noemen dat zelfs het omgekeerde waargenomen werd.

Als tweede werd getoetst of de groep proefpersonen met een relatief hoge wiskundige vaardigheid de benodigde groeps groottes beter in lijn met de Bayesiaanse voorspellingen inschatten, dan de groep proefpersonen met een relatief lage wiskundige vaardigheid (H5). Ook dit bleek niet het geval te zijn geweest. Wederom werd er tegen de verwachting in wél een significant resultaat gevonden voor de groep proefpersonen met relatief lage wiskundige vaardigheid en niet voor de groep proefpersonen met relatief hoge wiskundige vaardigheid. De groep proefpersonen met relatief lage wiskundige vaardigheid bleken de groeps groottes beter in te schatten in lijn met de Bayesiaanse voorspelde groeps groottes, dan dat de groep proefpersonen met relatief hoge wiskundige vaardigheid. Wederom werd dus het omgekeerde van de verwachte koppeling waargenomen.

Discussie

Ook onderstaande discussie is opgedeeld in A) discussie met betrekking tot het replicatieonderzoek en B) discussie met betrekking tot de koppeling met numerieke voorkeur en wiskundige vaardigheid.

A) Discussie met betrekking tot het replicatieonderzoek

De statistische analyses zijn niet gedaan over alle 175 deelnemers die het onderzoek succesvol hadden afgerond, maar over een verkleinde groep bestaande uit 134 proefpersonen (zie *statistische toetsing*). De reden hiervoor was dat er proefpersonen waren die buitenproportioneel hoge inschattingen hadden gemaakt, waardoor een vergelijking van de gemiddelden zinloos zou gaan worden. Er volgde een zoektocht naar een eerlijke methode om het effect van deze hoge uitbijters te elimineren, zodat de data betekenisvol geïnterpreteerd konden worden. Toen is gekozen om bij de analyses alle data mee te nemen van proefpersonen die inschattingen hadden gemaakt tot aan het grensgetal van 999 toe. In de optiek van de onderzoeker, is dit namelijk het getal waarboven schattingen als extreem en weinig realistisch kunnen worden aangemerkt en bovendien voldeed het aan de ambitie om de ruwe data van genoeg proefpersonen te behouden. De onderzoeker had ook andere opties overwogen om de situatie op te lossen, tevens opties die mogelijk nóg beter voldeden aan de ambitie om het databestand zoveel mogelijk intact te houden. Zo had er een logaritme van alle inschattingen genomen kunnen worden, waardoor de hele grote getallen relatief sterker verkleind werden dan de kleine getallen. Op deze manier zouden mogelijk alle data meegenomen worden in de statistische toetsing. Er is echter niet voor deze optie gekozen omdat analyse van de ruwe data voorkeur had boven het analyseren van bewerkte data.

In eerste instantie dient er gekeken te worden naar de oorzaak waarom sommige proefpersonen extreem hoge inschattingen hebben gemaakt en waardoor het nou komt dat de inschattingen zover uit elkaar zijn komen te liggen. Navraag onder proefpersonen gaf te kennen dat het te maken kan hebben gehad met de gekozen scenario's en een verscheidenheid in persoonskenmerken van de proefpersonen. Neem als voorbeeld het scenario bij de instructietekst (onderwerp: zwemmen in IJssel kan leiden tot verdrinking) en het eerste scenario (onderwerp: rotonde in Apeldoorn leidt tot betere verkeersdoorstroming). In het geval van verdrinking gaat het om leven en dood, mensen blijken het moeilijk te vinden om dit in te schatten. Ze vertrouwen blindelings de Adjudant van de Waterpolitie (expert), omdat zij denken dat de wandelaars uit de buurt (leken) mogelijk niet objectief genoeg naar de situatie kijken. Terwijl het in het andere geval gaat over een minder ingrijpende maatregel, namelijk of er ergens in Apeldoorn wel of geen rotonde moet komen. Op dat moment vinden mensen de mening van een wegebouwkundige van de BAM (expert) minder relevant, want sporters bij een sportclub in de buurt fietsen er waarschijnlijk wekelijks overheen en zijn daardoor in een even goede positie om dit te weten. Mogelijk is het onderwerp (ernst) van belang voor de inschatting die mensen maken en kan dit een factor zijn die een rol speelt waardoor de inschattingen ver uit elkaar zijn komen te liggen.

Een mogelijke oplossing, om te voorkomen dat de antwoorden ver uit elkaar komen te liggen, kan zijn het aanreiken van een ankerpunt of het beperken van de antwoordmogelijkheden. Deze oplossing is gebaseerd op het '*Mississippi counting problem*'. Als je iemand onvoorbereid vraagt hoelang hij of zij denkt dat de Mississippi rivier is, dan vinden mensen het over het algemeen lastig om deze vraag te beantwoorden. Als je echter aan iemand een ankerpunt biedt en een gevoel van richting geeft, dan ervaren mensen het als veel gemakkelijker een antwoord te formuleren. Dit zou je bijvoorbeeld kunnen doen door de vraag als volgt te stellen: 'ik denk dat de Mississippi rivier 5000 kilometer lang is, hoelang denk jij dat de rivier is?'. De ontvanger kan zichzelf dan afvragen of de rivier langer of juist korter is en in welke mate langer of korter. In dit onderzoek werden proefpersonen bij inschattingen op de vijf scenario's vrijgelaten voor wat betreft hun numerieke respons. Voor vervolgonderzoek zou de auteur adviseren om proefpersonen in de introductie een kader te geven, door bijvoorbeeld te stellen: 'per scenario mag u minimaal 1 en maximaal 999 leden toewijzen aan de groep leken die nodig is om op te wegen tegen de mening van de expert'. Op deze manier krijgen mensen een ankerpunt aangereikt waardoor zij makkelijker antwoorden kunnen formuleren.

De inschattingen van proefpersonen bij Hornikx et al. (2018) lagen veel minder ver uit elkaar dan bij deze studie, terwijl hetzelfde onderzoekontwerp (met deels andere scenario's) is gebruikt. Er kan geconcludeerd worden dat de uiteenlopende inschattingen dus niet liggen aan een fout in het ontwerp. Wel kan er voor vervolgonderzoek nagedacht worden over een ontwerp waarbij de ecologische validiteit hoger ligt. Denk nogmaals aan het voorbeeld van de wasmachine. In het dagelijks leven komen mensen steeds vaker dergelijke visuele situaties tegen waarin de mening van een expert tegenover die van een groep leken wordt gezet. In de huidige studies werden de scenario's echter aangeboden in tekstuele vorm, terwijl een meer visuele manier van het stimulusmateriaal presenteren wellicht de ecologische validiteit zou verbeteren. In vervolgonderzoek zou de onderzoeker bijvoorbeeld advertenties van Coolblue of Bol.com kunnen laten zien, met hierin een expert en een groep leken met een tegengestelde mening. Het kan dan bijvoorbeeld vijf advertenties betreffen van witgoed (denk aan koelkast, diepvries, afwasmachine, wasmachine en droger) of bruingoed (denk aan televisie, radio, cd-spelers, geluidsboxen en PlayStation). Bij advertenties voor deze producten kan vervolgens de deskundigheid van de expert worden aangegeven. In de vorm van onderstaande recensies kan ook de hoeveelheid óf de deskundigheid van de leken gemanipuleerd worden. Alle overige randzaken zoals een productfoto, de prijs, etc. dienen hetzelfde te zijn bij elk product. Op het moment dat mensen deze visuele advertenties beoordelen herkennen ze mogelijk meer de 'echte' situatie.

Voorafgaand aan het onderzoek was de auteur van mening dat een pre-test van het experiment niet nodig zou zijn, omdat het een replicatie van eerder onderzoek betrof. Terugkijkend op deze situatie, met de alle uiteenlopende inschattingen in de ruwe data in gedachten, bleek dit geen juiste inschatting. Mogelijk had een pre-test dit voorafgaand kunnen uitwijzen of had de onderzoeker voorafgaand al kunnen concluderen dat het nuttig was geweest om samen met de proefpersonen de instructie door te nemen en de voorbeeldcasus te behandelen.

Verder had een pre-test mogelijk geholpen om de doeltreffendheid van de scenario's te verifiëren en met name hoe extreem de situaties zijn die werden beschreven. De twee gerepliceerde scenario's, scenario 1 (Ronde in Apeldoorn) en scenario 5 (Autovrije zones) betroffen vrij neutrale scenario's, terwijl de drie nieuw ontwikkelde scenario's minder neutraal waren. Zo is scenario 2 (Dijken versterken) in het kader van *Global Warming* een voorbeeld van een meer extreem scenario, net zoals scenario 5 (Modern windmolenpark) een meer emotioneel geladen onderwerp is: iedereen weet rationeel dat windmolens nodig zijn voor alternatieve energieopwekking, maar niemand wil deze in zijn of haar 'achtertuin' hebben.

Scenario 3 (Rekeningrijden op de A325) is vanwege een andere reden minder neutraal, namelijk doordat het een voorbeeld is van een scenario waarbij de groep leken een groter belang heeft in het onderwerp van discussie. Het scenario speelt zich af in de buurt van Nijmegen (Nijmegen-Lent), tevens de plaats waar het onderzoek is afgenomen. Dit alles kan invloed hebben gehad op de inschattingen van proefpersonen. Proefpersonen werden echter niet verzocht om hun geboorteplaats en woonplaats bij de demografische gegevens in te vullen, als dit wel was gedaan kon de invloed hiervan worden nagegaan. Het pre-testen of de scenario's daadwerkelijk gezien worden door proefpersonen als neutraal of extreem, had gedaan kunnen worden aan de hand van de mate van betrokkenheid van de proefpersonen bij de scenario's. Als mensen bijvoorbeeld bij het scenario met betrekking tot het moderne windmolenpark meer betrokkenheid laten zien dan bij een neutraler scenario zoals het scenario met betrekking tot een rotonde in Apeldoorn, dan kan dit helpen met verklaren waarom de inschattingen (ver) uit elkaar lagen.

Kortom, een pre-test had kunnen uitwijzen dat de proefpersonen moeite hadden met het beoordelen van de scenario's in de huidige vorm van het stimulusmateriaal. Door het stimulusmateriaal in meerdere (visuele aspecten) of mindere mate (beperken van antwoordmogelijkheden) aan te passen, was de opdracht wellicht duidelijker geweest en mogelijkwjs hadden zo de (ver) uit elkaar liggende aantallen kunnen worden voorkomen of op z'n minst kunnen worden beperkt.

B) Discussie met betrekking tot de koppeling met numerieke voorkeur en wiskundige vaardigheid

Bij de statistische analyse, met betrekking tot numerieke voorkeur en wiskundige vaardigheid, werden ook niet alle data van de 134 proefpersonen meegenomen (*zie statistische analyses*). Om een duidelijk verschil te hebben tussen 'hoog' en 'laag' wees een dubbele mediaansplit uit welke proefpersonen tot de relatief hoogste of laagste groep behoorden, voor zowel numerieke voorkeur als wiskundige vaardigheid. De analyses omtrent numerieke voorkeur werden gedaan over 71 van de 134 proefpersonen: de 25% proefpersonen met de relatief laagste numerieke voorkeur (N=35) werden afgezet tegen de 25% proefpersonen met de relatief hoogste numerieke voorkeur (N=36). Voor de analyses omtrent wiskundige vaardigheid werd de data van 69 van de 134 proefpersonen geanalyseerd: de 25% proefpersonen met relatief de laagste wiskundige vaardigheid (N=32) werden afgezet tegen de 25% proefpersonen met de hoogste wiskundige vaardigheid (N=37).

Ook hier had de onderzoeker andere opties overwogen, wederom opties die beter voldeden aan de ambitie om het databestand zoveel mogelijk intact te houden. De keuze voor een dubbele mediaansplit bleek echter wel een verstandige. Het verband was anders dan de auteur intuïtief kon bedenken. Aangezien er door de dubbele mediaansplit sprake was van een erg duidelijke splitsing tussen ‘hoog’ en ‘laag, kan worden uitgesloten dat dit een reden was dat het verband anders was dan vooraf werd verwacht.

Het gevonden verband was tegenovergesteld aan de verwachtingen van de hypothesen. De eerder uitgesproken twijfel over het materiaal en de inschattingen, geldt natuurlijk ook voor dit onderzoeksdoel. Wetenschappelijk gezien is het wel interessant waarom de inschattingen beter werden gemaakt door mensen met lage numerieke voorkeur of wiskundige vaardigheid. Hornikx et al. (2018) stelden al dat mensen intuïtief vrij goed zijn in het maken van een dergelijke inschatting. Een mogelijke verklaring kan zijn dat mensen met lage numerieke voorkeur en wiskundige vaardigheid de inschatting meer maken op basis van intuïtie, terwijl mensen met relatief hoge numerieke voorkeur of wiskundige vaardigheid de scenario's benaderen als een op te lossen rekensom. In toenemende mate dienen mensen in de 'echte wereld' dergelijke afwegingen te maken tussen een expert met een tegengestelde mening met een grotere groep leken. Doordat mensen dit steeds vaker moeten doen wordt hun intuïtieve gedrag meer automatisch en minder rationeel. Doordat zij er minder rationeel over nadenken laten zij zich mogelijk ook minder leiden door de ernst en impact van de scenario's.

Samenvattend, in dit onderzoek werd minimale ondersteuning gevonden voor de verwachtingen rondom het replicatieonderzoek. Alhoewel de resultaten minder sterk waren dan bij de studie van Hornikx et al. (2018), was het verwachte patroon nog steeds zichtbaar. Uit dit onderzoek bleek echter niet dat er een koppeling was tussen relatief hoge numerieke voorkeur of wiskundige vaardigheid en het goed inschatten van de groepsgroottes. Opmerkelijk genoeg bleek zelfs het tegenovergestelde waar: mensen met relatief lage numerieke voorkeur of wiskundige vaardigheid maakten beter in lijn met de Bayesiaanse voorspelling inschattingen omtrent de groepsgroottes. Een mogelijke verklaring hiervoor kan zijn dat mensen met relatief lage numerieke voorkeur of wiskundige vaardigheid minder rationeel hadden nagedacht over de ernst en impact van de scenario's en hun inschatting meer hadden gebaseerd op basis van intuïtie. Verder onderzoek is nodig om meer concrete inzichten hierover te verkrijgen en sterkere conclusies te trekken.

Literatuur

Ackerman, P.L., & Beier, M.E. (2006). Determinants of Domain Knowledge and Independent Study Learning in an Adult Sample. *Journal of Educational Psychology*, 98 (2), 366-381. doi: 10.1037/0022-0663.98.2.366

Aiken, L. R. (1974). *Assessment of personality*. Needham Heights, MA: Allyn & Bacon.

Baratgin, J., & Politzer, G. (2010). Updating: A psychologically basic situation of probability revision. *Thinking & Reasoning*, 16 (4), 253-287. doi: 10.1080/13546783.2010.519564

Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford: Oxford University Press.

Brock, T.C. (1965). Communicator-Recipient similarity and decision change. *Journal of Personality and Social Psychology*, 1 (6), 650-654.

Camanho, N., & Fernandes, D. (2016) The Mortgage Illusion. doi: 10.2139/ssrn.1856325

Chaiken, S. (1987). *The heuristic model of persuasion*. In M. P. Zanna, J. M. Olson, & C. P.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Eemeren, F.H. van, Garssen, B., & Meuffels, B. (2009). *Fallacies and judgments of reasonableness: Empirical research concerning the pragma-dialectical discussion rules*. Dordrecht: Springer.

Eemeren, F. H. van, Grootendorst, R., & Snoeck Henkemans, F. (1997). *Handboek argumentatietheorie*. Groningen: Martinus Nijhoff.

Eemeren, F. H. van, & Snoeck Henkemans, A. F. (2011). *Argumentatie. Inleiding in het identificeren van meningsverschillen en het analyseren, beoordelen en houden van betogen*. Groningen: Noordhoff, 137-144 & 234-237.

Erdogan, B. Z. (1999). Celebrity endorsement: A literature review. *Journal of Marketing Management*, 15 (4), 291-314. doi: 10.1362/026725799784870379

Fernandes, D., Lynch Jr. J.G., & Netemeyer, R.G. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Management Science*, 60 (8), 1861-1883. doi: 10.1287/mnsc.2013.1849

Freling, T. H., & Dacin, P. A. (2010). When consensus counts: Exploring the impact of consensus claims in advertising. *Journal of Consumer Psychology*, 20, 163-175. doi: 10.1016/j.jcps.2009.12.001

Friedman, H., & Friedman, L. (1979). Endorser effectiveness by product type. *Journal of Advertising Research*, 19 (5), 63-71.

Godden, D.M. (2008). On common knowledge and ad populum: Acceptance as grounds for acceptability. *Philosophy and Rhetoric*, 51, 101-129.

Hahn, U., Harris, A. J. L., & Corner, A. (2009). Argument content and argument source: An exploration. *Informal Logic*, 29 (4), 337-367.

Hahn, U., & Oaksford, M. (2006). A Bayesian approach to informal argument fallacies. *Synthese*, 152, 207-236. doi: 10.1007/s11229-005-5233-2

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review*, 114, 704-732. doi: 10.1037/0033-295X.114.3.704

Harkins, S. G., & Petty, R. E. (1987). Information Utility and the Multiple Source Effect. *Journal of Personality and Social Psychology*, 52 (2), 260-268.

Harris, A.J.L., Hahn, U., Madsen, J.K., & Hsu, A.S. (2016). The Appeal to Expert Opinion: Quantitative Support for a Bayesian Network Approach. *Cognitive Science*, 40 (6), 1496-1533. doi: 10.1111/cogs.12276

Harris, A.J.L., Hsu, A.S., & Madsen, J.K. (2012). Because Hitler did it! Quantitative tests of Bayesian argumentation using ad hominem. *Thinking & Reasoning*, 18 (3), 311-343. doi: 10.1080/13546783.2012.670753

Hoeken, H., Hornikx, J., & Hustinx, L. (2012). *Overtuigende teksten. Onderzoek en ontwerp*. Bassum: Coutinho.

Hornikx, J. (2013). Een Bayesiaans perspectief op argumentkwaliteit. Het ad populum-argument onder de loep. *Tijdschrift voor Taalbeheersing*, 35 (2), 128-143. doi: 10.5117/TVT2013.2.HORN

Hornikx, J., Harris, A. J. L., & Boekema, J. (2018). How many laypeople holding a popular opinion are needed to counter an expert opinion? *Thinking & Reasoning*, 24 (1), 117-128. doi: 10.1080/13546783.2017.1378721

Hovland, C.I., Janis, I.L., & Kelly, H.H. (1953). *Persuasion and communication*. New Have: Yale university press.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, 21 (1), 37-44.

Maheswaran, D., & Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effect of incongruent information on processing and judgment. *Journal of Personality and Social Psychology*, 61 (1), 13-25.

Oaksford, M., & Hahn, U. (2004). A Bayesian Approach to the Argument From Ignorance. *Canadian Journal of Experimental Psychology*. 58 (2), 75-85. doi: 10.1037/h0085798

O'Keefe, D.J. (2002). *Persuasion: Theory and research* (2e editie). Thousand Oaks, CA: Sage.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer. doi: 10.1007/9781-4612-4964-1

Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, 41 (5), 847-855.

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, 10 (2), 135-146. doi: 10.1086/208954

Pornpitakpan, C. (2004). The Persuasiveness of Source Credibility: A Critical Review of Five Decades' Evidence. *Journal of Applied Social Psychology*, 34 (2), 243-281. doi: 10.1111/j.1559-1816.2004.tb02547.x

Schellens, P. J., & Verhoeven, G. (1994). *Argument en tegenargument. Een inleiding in de analyse en beoordeling van betogende teksten*. Groningen: Martinus Nijhoff.

Viswanathan, M. (1993). Measurement of Individual Differences in Preference for Numerical Information. *Journal of Applied Psychology*, 78 (5), 741–752. doi: 10.1037/0021-9010.78.5.741

Viswanathan, M., & Childers T.L. (1996). Processing of Numerical and Verbal Product Information. *Journal of Consumer Psychology*, 5 (4), 359-385. doi: 10.1207/S15327663jcp0504_03

Walton, D. N. (1989). *Informal logic: a handbook for critical argumentation*. Cambridge: Cambridge University Press.

Walton, D. N. (1997). *Appeal to expert opinion: Arguments from authority*. University Park, PA: Penn State Press.

Walton, D. N. (1999). *Appeal to popular opinion*. University Park, PA: Penn State Press.

Walton, D. N., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. Cambridge: Cambridge University Press.

Wilson, E. J., & Sherrell, D. L. (1993). Source effects in communication and persuasion research: A meta-analysis of effect size. *Journal of the Academy of Marketing Science*, 21 (2), 101-112. doi: 10.1177/009207039302100202

Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement*, 45, 401-405. doi: 10.1177/001316448504500226

Wreen, M. (1993). Jump with common spirits: Is an ad populum argument fallacious? *Metaphilosophy*, 24, 61-75.

Zillmann, D., Calison, C., & Gibson, R. (2009). Quantitative Media Literacy: Individual Differences in Dealing with Numbers in the News. *Media Psychology*, 12 (4), 394 – 416. doi: 10.1080/15213260903287275

Appendix A. Stimulusmateriaal (versie 1)

Appendix A1. Welkomboodschap

Fijn dat u wilt meewerken aan het onderzoek voor mijn masterscriptie!

Het blijkt dat op sommige mobiele telefoons de tekst onduidelijk te lezen is. In dit geval is het raadzaam om de vragenlijst in te vullen op een computer of tablet. Als u op de rode button onderaan dit scherm klikt, dan volgt eerst een instructie waarna het onderzoek begint.

Het invullen van de vragenlijst duurt ongeveer 10 minuten. De gegeven antwoorden zullen anoniem en vertrouwelijk worden behandeld. Mocht u nog vragen of opmerkingen hebben over het onderzoek kunt u contact met mij opnemen via n.gorris@student.ru.nl.

Hartelijk dank voor uw medewerking!
Niek Gorris
Radboud Universiteit Nijmegen



Appendix A2. Instructie (inclusief voorbeeldcasus)

Onderzoek over gemeentelijke plannen

Graag willen we uw mening horen over een aantal gemeentelijke plannen. Steeds worden deze gepresenteerd in een gespreksvorm tussen persoon A, B en C. Lees eerst de onderstaande voorbeeldcasus en instructie, voordat u met de vragenlijst begint.

Instructie

- Persoon A zegt: Heb je gehoord dat men van plan is om het zwemmen in de IJssel bij Olst te verbieden, omdat de stroming van langsvarende boten tot verdrinking kan leiden?
- Persoon B zegt: Ja, maar waarom stel je deze vraag?
- Persoon A zegt: Ik vraag dit omdat ik niet weet of dit wel een goed plan is.
- Persoon B zegt: Ik denk van wel.
- Persoon A zegt: En waarom denk je dit dan?

Persoon B zegt: Omdat een Adjudant van de waterpolitie aangeeft dat de stroming van langsvarende boten in de IJssel bij Olst tot verdrinking kan leiden. En hij kan het weten omdat hij in 95.00% van de gevallen gelijk heeft als het om de veiligheid op het water gaat.

(Persoon C hoort de conversatie tussen persoon A en persoon B en spreekt vervolgens persoon B tegen.)

Persoon C zegt: Ik heb echter gehoord dat wandelaars, die weleens in de omgeving van Olst wandelen, aangeven dat de stroming van langsvarende boten in de IJssel bij Olst niet tot verdrinking kan leiden.

De wandelaars zijn niet deskundig op het gebied van veiligheid op het water; over dit onderwerp heeft elke wandelaar in 60.00% van de gevallen gelijk, echter is het wel een grote groep wandelaars die dit beweert. Hoe groot moet de groep wandelaars zijn om op te wegen tegen de mening van de Adjudant van de waterpolitie?

..... wandelaars

Vul hierboven het aantal personen in dat volgens u nodig is om tegen de expert op te kunnen wegen. Aangezien het uw persoonlijke mening is, zijn er geen foute antwoorden mogelijk.

Schrijf uw geschatte aantal personen op in cijfers.

Let op: Einde instructie. Als u op de rode button klikt begint het onderzoek.

Hartelijk dank voor uw medewerking!

Niek Gorris

Radboud Universiteit Nijmegen



Appendix A3. Scenario 's (percentages specifiek voor versie 1)

Scenario 1: Transport in de stad

Persoon A zegt: Heb je gehoord dat ze van het kruispunt op de Jachtlaan in Apeldoorn een rotonde willen maken om het verkeer beter te laten doorstromen?

Persoon B zegt: Ja, maar waarom stel je deze vraag?

Persoon A zegt: Ik vraag dit omdat ik niet weet of dit wel een goed plan is.

Persoon B zegt: Ik denk van wel.

Persoon A zegt: En waarom denk je dit dan?

Persoon B zegt: Omdat een wegenbouwkundige van de BAM aangeeft dat de rotonde zorgt voor een betere doorstroming van het verkeer. En hij kan het weten omdat hij in 99.99% van de gevallen gelijk heeft als het om verkeersdoorstroming gaat.

(Persoon C hoort de conversatie tussen persoon A en persoon B en spreekt vervolgens persoon B tegen.)

Persoon C zegt: Ik heb echter gehoord dat sporters, van een sportschool in Apeldoorn, denken dat de rotonde niet zal zorgen voor een betere doorstroming van het verkeer.

De sporters zijn niet deskundig op het gebied van verkeersdoorstroming; over dit onderwerp heeft elke sporter in 55.00% van de gevallen gelijk, echter is het wel een grote groep sporters die dit beweert. Hoe groot moet de groep sporters zijn om op te wegen tegen de mening van de wegenbouwkundige van de BAM?

..... sporters



Scenario 2: Dijken versterken

Persoon A zegt: Heb je gehoord dat ze de dijken in Friesland willen versterken, omdat de bodemdaling en de zeespiegelstijging een steeds grotere bedreiging vormt voor Nederland?

Persoon B zegt: Ja, maar waarom stel je deze vraag?

Persoon A zegt: Ik vraag dit omdat ik niet weet of dit wel een goed plan is.

Persoon B zegt: Ik denk van wel.

Persoon A zegt: En waarom denk je dit dan?

Persoon B zegt: Omdat prof. dr. Toine Smits, die vandaag de leerstoel ‘duurzaam waterbeheer’ krijgt aan de Erasmus Universiteit in Rotterdam, aangeeft dat dit noodzakelijk is omdat de huidige dijken niet sterk genoeg zijn. En hij kan het weten omdat hij in 75.00% van de gevallen gelijk heeft als het om dijkenbouw gaat.

(Persoon C hoort de conversatie tussen persoon A en persoon B en spreekt vervolgens persoon B tegen.)

Persoon C zegt: Ik heb echter gehoord dat bezoekers van een restaurant in Friesland denken dat de strijd tegen water niet te winnen is met het versterken van dijken.

De restaurantbezoekers zijn niet deskundig op het gebied van dijkenbouw; over dit onderwerp heeft elke restaurantbezoeker in 55.00% van de gevallen gelijk, echter is het wel een grote groep restaurantbezoekers die dit beweert. Hoe groot moet de groep restaurantbezoekers zijn om op te wegen tegen de mening van de dijkenbouwkundige Toine Smits?

..... restaurantbezoekers



Scenario 3: Rekeningrijden op de A325

Persoon A zegt: Heb je in de krant gelezen dat men van plan is om rekeningrijden op de A325 te introduceren, zodat het aantal files tussen Arnhem en Nijmegen wordt teruggedrongen?

Persoon B zegt: Ja, maar waarom stel je deze vraag?

Persoon A zegt: Ik vraag dit omdat ik niet weet of dit wel een goed plan is.

Persoon B zegt: Ik denk van wel.

Persoon A zegt: En waarom denk je dit dan?

Persoon B zegt: Omdat een hoogleraar Geografie, Planologie en Milieu, werkzaam bij de Radboud Universiteit Nijmegen, aangeeft dat het aantal files in de regio hierdoor vermindert. En hij kan het weten omdat hij in 95.00% van de gevallen gelijk heeft als het om verkeersdoorstroming gaat.

(Persoon C hoort de conversatie tussen persoon A en persoon B en spreekt vervolgens persoon B tegen.)

Persoon C zegt: Ik heb echter gehoord dat bezoekers van de CineMec Nijmegen, een bioscoop gelegen aan de A325, denken dat rekeningrijden niet zal zorgen voor een betere doorstroming van het verkeer.

De bioscoopbezoekers zijn niet deskundig op het gebied van verkeersdoorstroming; over dit onderwerp heeft elke bioscoopbezoeker in 60.00% van de gevallen gelijk, echter is het wel een grote groep bioscoopbezoekers die dit beweert. Hoe groot moet de groep bioscoopbezoekers zijn om op te wegen tegen de mening van de hoogleraar Geografie, Planologie en Milieu?

..... bioscoopbezoekers



Scenario 4: Autovrije zones

Persoon A zegt: Heb je gehoord dat men de Nieuwstraat in Apeldoorn autovrij wil maken, omdat er verwacht wordt dat dit zal zorgen voor meer klandizie in de winkels in de straat?

Persoon B zegt: Ja, maar waarom stel je deze vraag?

Persoon A zegt: Ik vraag dit omdat ik niet weet of dit wel een goed plan is.

Persoon B zegt: Ik denk van wel.

Persoon A zegt: En waarom denk je dit dan?

Persoon B zegt: Omdat een hoogleraar Retailmarketing, werkzaam bij de Universiteit Utrecht, aangeeft dat dit zal helpen tot meer klandizie in de winkels in de straat. En hij kan het weten omdat hij in 80.00% van de gevallen gelijk heeft als het om consumentenkoop gaat.

(Persoon C hoort de conversatie tussen persoon A en persoon B en spreekt vervolgens persoon B tegen.)

Persoon C zegt: Ik heb echter gehoord dat stratenmakers, die bezig zijn in de buurt, aangeven dat dit niet zal helpen tot meer klandizie in de winkels in de straat.

De stratenmakers zijn niet deskundig op het gebied van consumentenkoop; over dit onderwerp heeft elke stratenmaker in 52.50% van de gevallen gelijk, echter is het wel een grote groep stratenmakers die dit beweert. Hoe groot moet de groep stratenmakers zijn om op te wegen tegen de mening van de hoogleraar retailmarketing?

..... stratenmakers



Scenario 5: Modern windmolenpark

Persoon A zegt: Heb je in de krant gelezen dat men van plan is om een modern windmolenpark in Den Bosch aan te leggen, omdat dit type park genoeg energie opwekt om alle basisscholen in Den Bosch van stroom te voorzien?

Persoon B zegt: Ja, maar waarom stel je deze vraag?

Persoon A zegt: Ik vraag dit omdat ik niet weet of dit wel een goed plan is.

Persoon B zegt: Ik denk van wel.

Persoon A zegt: En waarom denk je dit dan?

Persoon B zegt: Omdat Sible Schöne, de nieuwe algemeen directeur van het HIER klimaatbureau, aangeeft dat dit type park genoeg energie opwekt om alle basisscholen in Den Bosch van stroom te voorzien. En hij kan het weten omdat hij in 90.00% van de gevallen gelijk heeft als het om alternatieve energiebronnen gaat.

(Persoon C hoort de conversatie tussen persoon A en persoon B en spreekt vervolgens persoon B tegen.)

Persoon C zegt: Ik heb echter gehoord dat overblijfoeders, die werkzaam zijn in de buurt van Den Bosch, denken dat dit park niet genoeg energie opwekt voor alle basisscholen in Den Bosch.

De overblijfoeders zijn niet deskundig op het gebied van alternatieve energiebronnen; over dit onderwerp heeft elke overblijfoeder in 57.50% van de gevallen gelijk, echter is het wel een grote groep overblijfoeders die dit beweert. Hoe groot moet de groep overblijfoeders zijn om op te wegen tegen de mening van de algemeen directeur van het HIER klimaatbureau?

..... overblijfoeders



Appendix A4. Vragenlijst numerieke voorkeur (Viswanathan, 1993)

Geef aan in welke mate u het eens bent met de volgende stellingen:

	zeer mee oneens	mee oneens	enigszins mee oneens	neutraal	enigszins mee eens	mee eens	zeer mee eens
Ik vind werk waarbij je met getallen moet werken leuk.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Het geeft mij voldoening om alledaagse problemen op te lossen waarbij getallen betrokken zijn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Numerieke informatie is erg nuttig in het dagelijks leven.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik geef liever geen aandacht aan informatie waarbij getallen betrokken zijn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik denk liever niet na over zaken waar getallen bij betrokken zijn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik vind het leuk om berekeningen te maken met numerieke informatie.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In de meeste situaties vind ik numerieke informatie niet relevant.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ik vind het leren en gebruiken van numerieke informatie belangrijk om weloverwogen beslissingen te maken.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

>>

Appendix A5. Arithmetic Aptitude Test (Zillmann, Callison & Gibson, 2009, p. 398)

Tot slot worden er tien *multiple choice* vragen gesteld om uw rekenvaardigheid te toetsen. Neem de tijd voor uw antwoorden. U mag absoluut geen rekenmachine gebruiken! Als u wilt mag u wel pen en papier gebruiken. Bent u zelf ook benieuwd naar het resultaat? Noteer dan uw antwoorden of maak er een foto van, want de correcte antwoorden zijn achteraf beschikbaar.

Interpretaties van kwalitatieve informatie

Wij bestuderen hoe mensen omgaan met complexe en vaak verwarrende feiten en getallen die in de media worden gerapporteerd. De onderstaande voorbeelden komen uit het nieuws en u wordt verzocht het correcte antwoord te geven of in te schatten:

Vraag 1: "De vorst heeft $\frac{5}{16}$ deel van de mandarijnenooft vernietigd." Is dit minder, meer of gelijk aan een kwart?

- minder dan een kwart **meer dan een kwart** gelijk aan een kwart

Vraag 2: "Op de middelbare school in deze kleine gemeenschap slagen maar zeven van de 20 studenten voor hun eindexamen" Over welk percentage slagende studenten wordt hier gesproken?

- 7% 20% 27% **35%** 63%

Vraag 3: "Doug McKinney kan rondkomen van een inkomen van €7.500. Hij kan de beloofde loonsverhoging van 20% nog steeds niet geloven." Hoeveel gaat hij in de toekomst verdienen?

- €7.750 €8.500 €8.750 **€9.000** €9.500

Vraag 4: "De storm (een tornado) beschadigde 80% van alle gebouwen in deze kleine gemeenschap" Welk deel van de gebouwen werd beschadigd?

- $\frac{3}{4}$ **$\frac{4}{5}$** $\frac{8}{12}$ $\frac{10}{18}$ $\frac{20}{30}$

Vraag 5: "Afgelopen zondag kreeg 30% van de 1200 patiënten in het ziekenhuis bezoek." Hoeveel patiënten ontvingen bezoek?

- 180 300 **360** 420 680

Vraag 6: "Twee van de acht biggen met ondergewicht overleefden de eerste twee weken." Wat is het percentage biggen dat overleeft?

- 10% 12% 18% 20% **25%**

Vraag 7: "Het vergif bleek verrassend effectief. Van de 20.000 bruine ratten die werden getest, gingen er 18.000 binnen 24 uur dood." Wat is het percentage dode ratten?

- 90%** 80% 70% 60% 50%

Vraag 8: "Slechts 10% van de mensen die aan dit virus worden blootgesteld krijgt daadwerkelijk de ziekte en slechts 5% van de geïnfecteerden sterft eraan." Als de blootgestelde populatie een miljoen mensen betreft, hoeveel mensen sterven er dan aan de ziekte?

- 5 50 500 **5.000** 50.000

Vraag 9: "Volgens onderzoek van het RIVM (Rijksinstituut voor Volksgezondheid en Milieu), is het risico voor het ontwikkelen van deze ziekte na een beet 0,00003." Als 1 miljoen mensen gebeten worden, hoeveel ontwikkelen er dan die ziekte?

- 3 **30** 300 3.000 30.000

Vraag 10: "Van de 2,5 miljoen sparren die jaarlijks in deze bergen sterven, schijnen er 11.212 te sterven door besmetting met de zogenaamde Edelspar kever." Hoe groot is de kans voor een willekeurige spar in deze bergen om dood te gaan als gevolg van Edelspar kever besmetting (geef de kans weer als een getal tussen 0 en 1)?

- 0.045 **0.0045** 0.00045 0.000045 0.0000045
-

Noot: de correcte antwoorden van vraag 1 tot 10 zijn respectievelijk 2,4,4,2,3,5,1,4,2,2.



Appendix A6. Persoonlijke gegevens

Geslacht

Man

Vrouw

Leeftijd

Nationaliteit

Huidige of hoogstgenoten opleiding

Wat is, volgens u, het doel van dit onderzoek?

Dit is het einde van de vragenlijst.

Let op: alleen als u op rode button onderaan deze pagina klikt, verzendt u uw antwoorden definitief.

Bent u nog benieuwd naar uw resultaat bij de rekentoets? [Klik hier](#) voor de correcte antwoorden.
Vergeet daarna dan alstublieft niet om alsnog op de rode button te klikken om uw antwoorden definitief in te zenden.

Nogmaals bedankt voor uw deelname,
Niek



Appendix B. Gemiddelden daadwerkelijke inschattingen per deskundigheidscombinatie

Tabel *Overzicht van het aantal leken dat per deskundigheidscombinatie nodig is om even overtuigend te zijn als de expert (N=134), inclusief de standaardafwijkingen.*

Expert:	99.99%		95.00%		90.00%		80.00%		75.00%		Totaal
Leken:	M	SD	M	SD	M	SD	M	SD	M	SD	Gemiddeld
60.00%	94.69	124.57	106.18	120.68	132.55	154.68	40.30	35.35	98.96	133.74	95.54
57.50%	133.67	158.70	122.81	154.30	148.57	193.16	83.25	109.24	109.23	136.06	119.51
55.00%	145.91	175.80	131.18	142.42	86.70	79.22	129.07	203.61	79.81	109.76	114.53
52.50%	142.43	162.37	107.27	110.64	138.42	184.71	84.98	76.73	110.29	169.89	116.68
51.00%	174.18	182.97	130.00	132.96	148.07	148.96	118.23	179.73	88.03	89.87	131.70
Totaal gemiddeld	138.18		119.48		130.86		91.16		97.26		