

Literatuurscan sociale media en democratie met speciale aandacht voor anonimiteit

Rens Vliegenthart
Sanne Kruikemeier
Emma Turkenburg
Alice Hamilton



WAGENINGEN
UNIVERSITY & RESEARCH

Literatuurscan sociale media en democratie met speciale aandacht voor anonimiteit

Rens Vliegenthart
Sanne Kruikemeier
Emma Turkenburg
Alice Hamilton

Strategische Communicatie groep

Rapport geschreven in opdracht van Ministerie
van Binnenlandse Zaken en Koninkrijksrelaties

Wageningen University & Research
Wageningen, november 2023 / februari 2024



WAGENINGEN

UNIVERSITY & RESEARCH

Referentienummer H7-101206-34307

© 2023 Wageningen, Wageningen University & Research

De verantwoordelijkheid voor de inhoud berust bij de auteurs van dit rapport.

Samenvatting

In deze notitie beantwoorden we de vraag wat de recente empirische wetenschappelijke inzichten zijn over de impact van verschillende aspecten van sociale media op democratie. – we schetsen een algemeen beeld en naast anonimiteit richten we ons ook het gebruik van algoritmen en onderliggende verdienmodellen van sociale mediaplatformen op verschillende aspecten van de democratie, met name polarisatie, radicalisering, bedreigingen, desinformatie en verharding van publieke debat.

Ons onderzoek is tweeledig:

- We doen een systematisch literatuuronderzoek naar sociaal-wetenschappelijke studies over de relatie tussen sociale media en democratie gepubliceerd in de jaren 2022 en 2023. Deze artikelen worden systematisch geanalyseerd op een aantal inhoudskarakteristieken, waaronder de onderzochte platformen en karakteristieken, en de uitkomstvariabelen (welke aspecten van democratie worden in beschouwing genomen?);
- Voor onderzoek specifiek naar anonimiteit op sociale media en onderzoek dat zich specifiek op Nederland richt doen we een uitgebreidere verkenning waarin ook studies uit eerdere jaren worden meegenomen.

De belangrijkste bevindingen van onze studie zijn de volgende:

- Opvallend is de sterke dominantie van het onderzoek in de Verenigde Staten, waarvan resultaten niet één op één vertalen naar de Nederlandse context. De specifieke verschillen tussen het politieke en mediasysteem van beide landen dwingt tot terughoudendheid in het generaliseren van Amerikaanse bevindingen naar de Nederlandse context;
- Het onderzoek bevestigt niet een éénduidig beeld van een negatieve impact van sociale media op democratie en het aantal bevindingen dat een positieve invloed van sociale media op democratische uitkomsten is zelfs groter dan het aantal negatieve bevindingen.
- Het onderzoek naar anonimiteit is beperkt en leidt niet tot een éénduidige conclusie. In recente studies wordt wel vaker gewezen op de negatieve gevolgen van anonimiteit, met name voor de kwaliteit van online discussies;
- Onderzoek naar sociale media en democratie in de Nederlandse context laat zien dat sociale media een belangrijke nieuwsbron zijn voor veel mensen. Er is weinig bewijs voor de aanwezigheid van grote hoeveelheden misinformatie op sociale media of sterke polariserende effecten van het gebruik van sociale media.
- Er is nog veel dat we niet weten, onder andere over de verspreiding van haatdragende communicatie en de effecten van sociale media inhoud op (gevoelens van) polarisatie.

Inleiding

De potentiële risico's van (het gebruik van) sociale media op allerlei aspecten van de Nederlandse democratie worden veelvuldig bediscussieerd, zowel in het publieke debat alsook in de politiek. De motie Gündogan c.s. (kamerstuk 30821-192) reflecteert deze zorgen en verzoekt om een onderzoek naar de mogelijke risico's van verschillende aspecten van sociale media voor onze democratie, met speciale aandacht voor de rol die anonimiteit speelt.

De motie reflecteert de inherent normatieve afwegingen die uiteindelijk ten grondslag liggen aan beleidskeuzes en het feit dat anonimiteit online een groot goed is, maar ook tot excessen kan leiden.

Veronderstelde mechanismen

De mogelijke gevolgen van sociale media voor onze democratie zijn velerlei. Grofweg zijn er drie argumentatielijnen te onderscheiden, die eigenlijk los van de specifieke onderzoeksvraag gebruikt worden.

Er zijn wetenschappers die de nadruk leggen op de negatieve invloed van sociale media op allerlei democratisch gedrag, zoals politieke betrokkenheid en participatie. Het oorspronkelijke argument dat daarbij gemaakt wordt is dan dat internet en sociale media gebruik ten koste gaan van zaken als vrijwilligerswerk en andere sociale activiteiten (Putnam, 2000). Andere wetenschappers beargumenteren dat sociale media een 'mobilisatie' effect kunnen hebben en juist achtergestelde groepen in een samenleving de kans geven om burgers te activeren. Een derde stroming wijst op het 'versterkende' effect van sociale media, waarmee bestaande verhoudingen en gedrag ondersteund en versterkt worden.

Empirisch onderzoek laat zien dat afhankelijk van de specifieke invulling van het onderzoek al deze drie mechanismes kunnen optreden, maar door de tijd zien we verschillende accenten. In eerste instantie heeft het onderzoek zich sterk gericht op de vraag hoe op individueel niveau het gebruik van sociale media een invloed had op allerlei vormen van politieke betrokkenheid en gedrag (Boulianne, 2015). In de afgelopen jaren zijn de vragen meer specifiek geworden, waarbij de specifieke eigenschappen en inhoud van sociale media een meer centrale rol hebben gekregen. Daarnaast zien we dat het onderzoek zich op een bredere waaier aan politieke uitkomsten gaat richten, waarbij de aandacht verschuift van 'klassieke' maten van politieke attitudes en gedrag naar polarisering, radicalisering en verharding van het publieke debat. Dit zijn zaken die meer direct een risico vormen voor onze democratische rechtsorde.

Belangrijk is om hierbij op te merken dat de context van cruciaal belang is en dat effecten afhankelijk zijn van onder andere het politieke systeem waarin zij onderzocht worden. Opvallend is

de sterke dominantie van het onderzoek in de Verenigde Staten, waarbij resultaten niet één op één vertalen naar de Nederlandse context.

Aanpak

In drie databases¹ van onderzoek uit de sociale wetenschappen hebben we gezocht naar relevante literatuur met betrekking tot sociale media en democratie die recent verschenen is, namelijk in 2022 en 2023. De zoekactie omvatte alle recente, peer-reviewed artikelen geschreven in het Nederlands of Engels die verwezen naar "sociale media" binnen vijf woorden van één van de volgende sleutelwoorden: democratie, politiek, polarisatie, anonimiteit of verdienmodellen, inclusief variaties op deze sleutelwoorden in de titel, het abstract en de *keywords* van het artikel. Na verwijdering van dubbele en ontoegankelijke zoekresultaten, werden de overgebleven 355 artikelen op basis van titel, verschijningsjaar en abstract beoordeeld als relevant of irrelevant. Hierbij sluiten we studies naar niet-Westerse landen, waarbij gesteld kan worden dat deze sowieso maar zeer spaarzaam aanwezig waren. Wij kozen voor deze relatief korte onderzoeksperiode om twee redenen. Ten eerste is er een aantal literatuurstudies die eerdere perioden reeds in beschouwing nemen. Ten tweede is er een pragmatisch argument – de hoeveelheid studies is zeer groot.

Ter aanvulling op deze selectie werden nog artikelen gezocht die specifiek betrekking hebben op anonimiteit, algoritmen, verdienmodellen of surveillancekapitalisme, onder andere via een vierde database (Web of Science). Na deze stappen zijn 140 artikelen meegenomen in de uiteindelijke analyse. Deze analyse bestond uit het systematisch coderen van relevante informatie uit elk artikel, onder andere het land waarin het onderzoek werd uitgevoerd, welk sociaal medium werd onderzocht, welke methode werd gehanteerd, welke afhankelijke en onafhankelijke variabelen werden meegenomen en een inventarisatie van mogelijk relevante bevindingen. Omdat in deze populatie het aantal studies over anonimiteit beperkt bleef, hebben we specifiek hiervoor een extra aantal artikelen verzameld waarbij we een langere periode in beschouwing nemen. Ook wordt er een vergelijkbaar uitgebreider overzicht van studies die specifiek betrekking hebben op Nederland gepresenteerd. Deze worden in de resultatensectie apart besproken.

In de onderstaande sectie schetsen we een algemeen beeld van de focus en bevindingen en kijken specifiek naar het onderzoek dat binnen Nederland is uitgevoerd. Daarnaast bespreken we meer uitgebreid het onderzoek dat zich richt op anonimiteit, algoritmen en verdienmodellen.

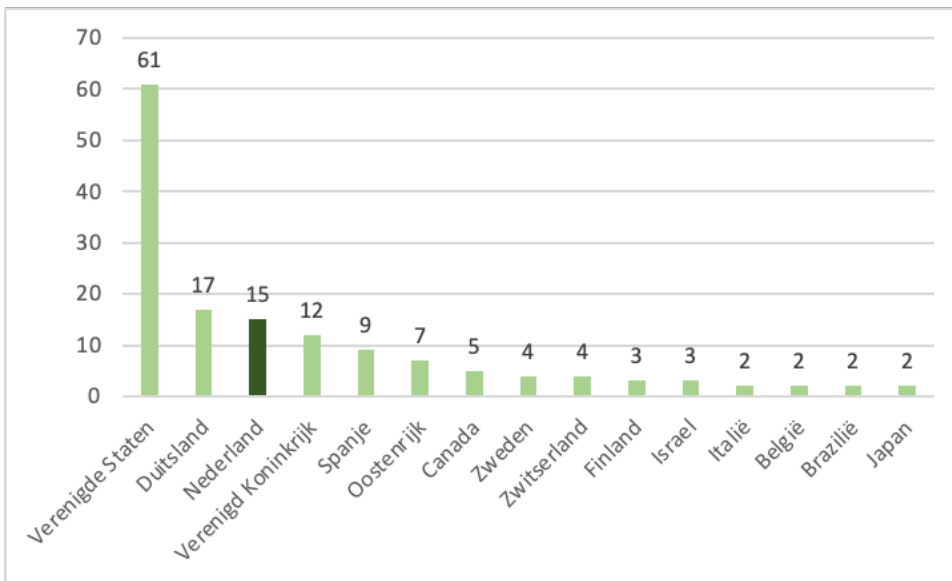
Bevindingen

Geografische focus

Het empirische onderzoek naar sociale media en democratie wordt sterk gedomineerd door studies die in de Verenigde Staten (61) zijn uitgevoerd. Figuur 1 presenteert een overzicht van de landen

¹ PsychInfo; Communication and Mass Media Complete; Worldview Political Science Abstracts

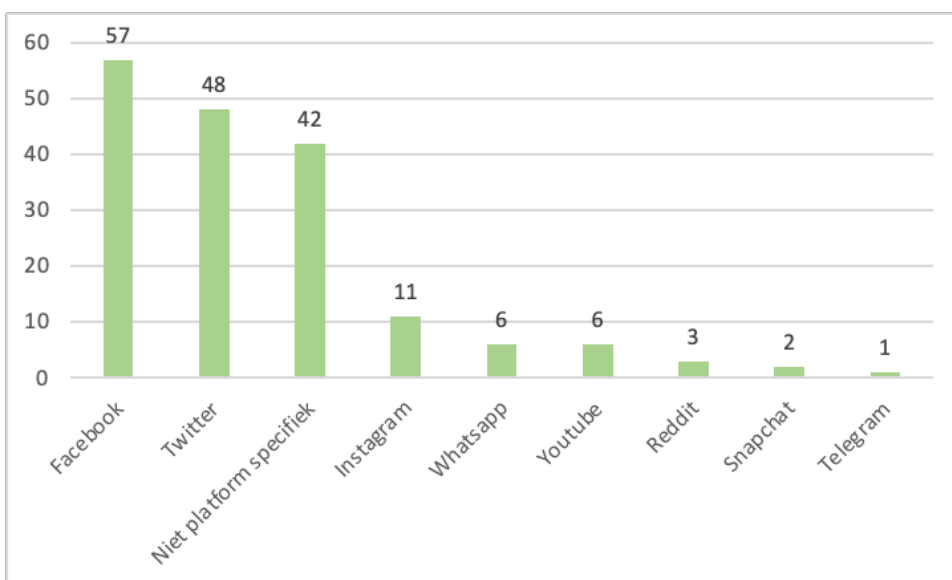
die meer dan één keer onderwerp van studie zijn. De dominantie van de Verenigde Staten valt op, maar ook Nederland is goed vertegenwoordigd, in 15 artikelen staat de Nederlandse context centraal.



Figuur 1. Aantal keren dat landen bestudeerd worden

Platformen

De studies richten zich sterk op Facebook (57) en Twitter (48) (zie Figuur 2). Veelgebruikte platformen als Instagram en YouTube worden veel minder bestudeerd en platformen waarvan vaak gezegd wordt dat zij plekken van radicalisering zijn (bijvoorbeeld Telegram) krijgen nauwelijks aandacht.

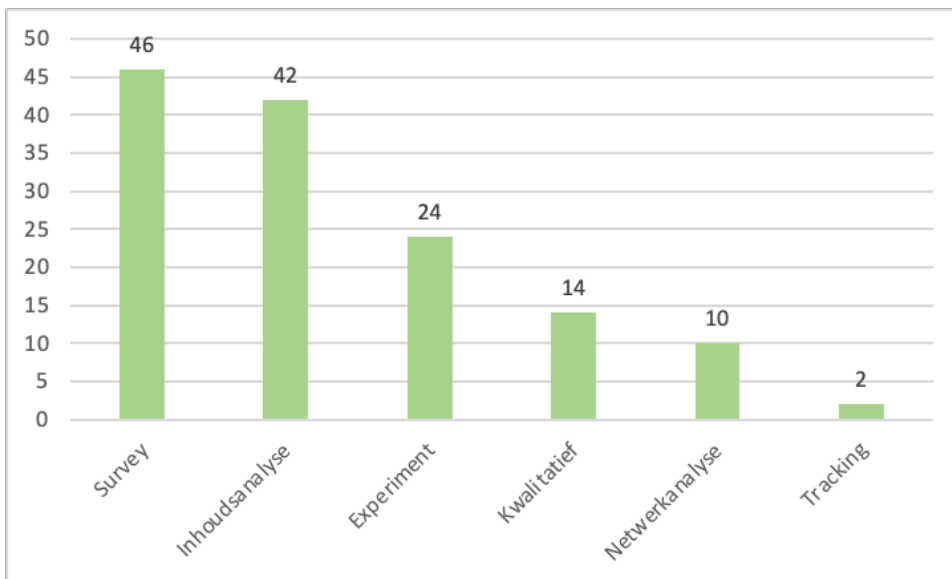


Figuur 2. Aantal keren dat verschillende platformen bestudeerd worden

Methoden

Er is sprake van een variëteit aan gebruikte methoden. Surveys (46), inhoudsanalyses (42) en experimenten (24) worden het meeste gebruikt, maar ook kwalitatieve methoden en

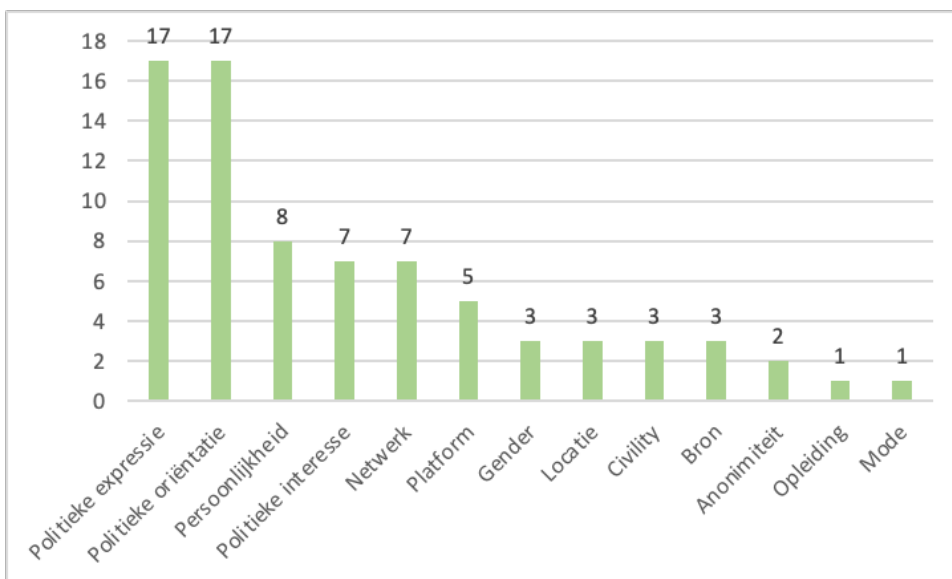
netwerkanalyses komen voor (Figuur 3). In slechts twee gevallen wordt gebruik gemaakt van tracking methoden, waarmee het daadwerkelijke gedrag van sociale mediagebruikers kan worden vastgelegd.



Figuur 3. Aantal keren dat verschillende platformen bestudeerd worden

Inhoudelijke focus

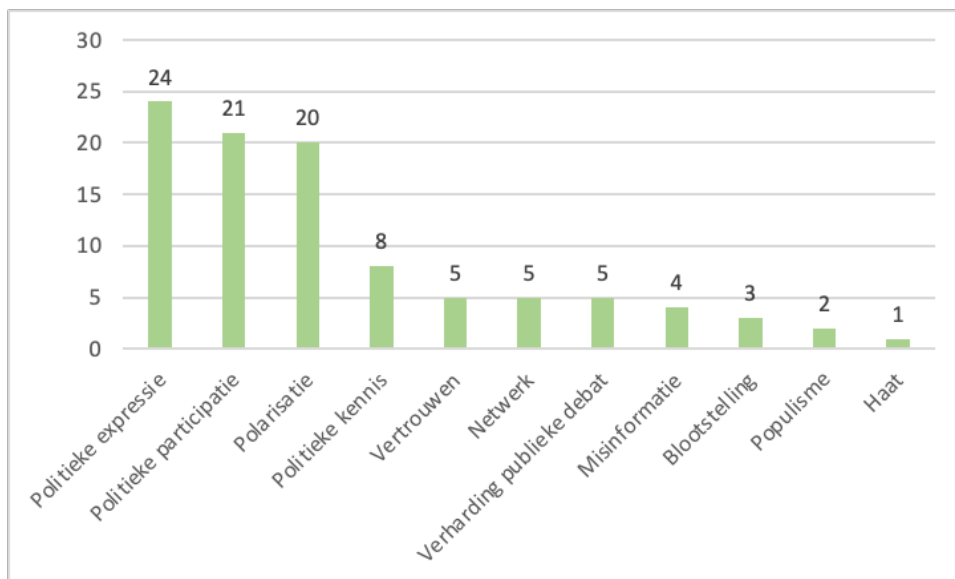
Het onderzoek benoemt een aantal risico's – waarbij polarisatie en radicalisering het meeste voorkomen (21). De verharding van het maatschappelijke debat wordt door een deel van de onderzoeken geadresseerd (7) en ook de verspreiding van desinformatie krijgt aandacht (6).



Figuur 4. Onafhankelijke variabelen centraal in de geanalyseerde studies

Conceptueel kan een onderscheid gemaakt worden in variabelen die een invloed uitoefenen (onafhankelijke variabelen) en uitkomstvariabelen, dus variabelen die beïnvloed worden

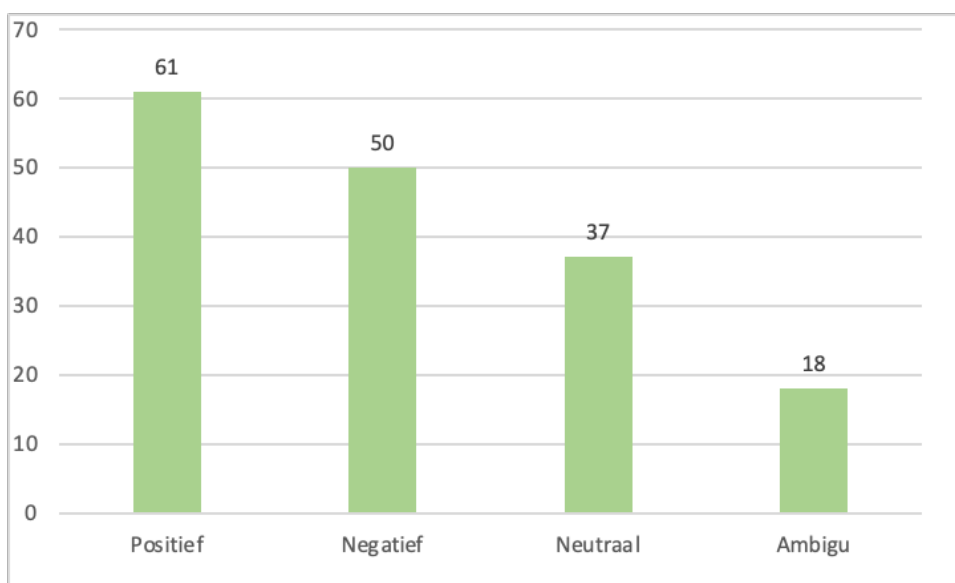
(afhankelijke variabelen). Voor beiden vinden we een grote variëteit, waarbij het onderzoek zich concentreert op de invloed van politieke expressie (17) en oriëntatie (17), persoonlijkheid (8) en politieke interesse (7) (Figuur 4). Uitkomsten die veel bekeken worden zijn politieke expressie (24) en participatie (21), maar ook polarisatie (20) en politieke kennis (8).



Figuur 5. Afhankelijke variabelen centraal in de geanalyseerde studies

Algehele bevindingen

Een belangrijke vraag is of de effecten van sociale media op democratie op basis een bepaalde studie als positief of negatief gekenmerkt kunnen worden (zie Figuur 6).



Het aantal 'positieve' bevindingen (61) is hoger dan het aantal 'negatieve' bevindingen (50). In een aantal andere gevallen kunnen de resultaten als neutraal (37) of ambigu (18) worden

gekenmerkt. Het onderzoek bevestigt dus duidelijk niet een éénduidig beeld van een negatieve impact van sociale media op democratie. Vaak wordt verondersteld dat de politieke context een invloed uitoefent en dat effecten van sociale media in een meerpartijstelsel minder negatief zijn dan in een tweepartijstelsel, waar de institutionele context tweedeling en polarisatie faciliteert. Uit de literatuurréview komt dit verschil niet duidelijk naar voren. De verhouding tussen positieve en negatieve bevindingen verschilt niet tussen beide typen systemen.

We richten ons in onderstaande bespreking op relaties die relevant zijn voor de benoemde risico's in de motie.

Ten eerste concepten die direct centraal stonden in de motie. Anonimiteit bespreken we in een aparte sectie hieronder. In onze inventarisatie komt deze maar in twee studies aan de orde, waardoor we een aanvullende inventarisatie over een lange tijdsperiode hebben gemaakt.

In slechts één van de artikelen komt het verdienmodel van platformen expliciet terug. Coscia en Rossi (2022) gebruiken een agent based model om te laten zien hoe contentmoderatie polarisatie kan beïnvloeden. Zij argumenteren dat platformen het delen van berichten als betrokkenheid zien, die vanuit een adverteerdersperspectief financieel aantrekkelijk is. Door de verspreiding van problematische inhoud te verminderen en dus niet alleen financiële overwegingen mee te nemen worden positieve effecten in termen van depolarisatie bereikt.

Over algoritmen en algoritmische selectie is in de afgelopen periode meer onderzoek verschenen. Drie van deze studies komen uit een groot project dat een groep Amerikaanse wetenschappers samen met Meta heeft uitgevoerd. Nyhan et al. (2023) concluderen dat het reduceren van "echo chambers", situaties waarin mensen alleen in aanraking komen met gelijkgestemden, zorgt voor minder blootstelling aan onbeschofte informatie, maar geen effect heeft op affectieve polarisatie², ideologische extremiteit en geloof in misinformatie. González-Baílón et al. (2023) vinden een sterke segregatie tussen Conservatieven en Democraten in nieuwsconsumptie via Facebook. Deze segregatie neemt toe als niet naar potentiële blootstelling maar naar daadwerkelijke blootstelling gekeken wordt en nog meer als er naar engagement wordt gekeken. Verder is er een duidelijke Conservatieve echo chamber te onderscheiden en komen Conservatieven veel meer in aanraking met misinformatie dan Democraten. Guess et al. (2023) onderzoeken in een veldexperiment hoe algoritmisch gedreven ordening van tijdlijnen op Facebook en Instagram en chronologische tijdlijnen andere effecten hebben. Ze vinden dat chronologisch geordende tijdlijnen de tijd die gebruikers op de platforms doorbrachten en hun activiteit aanzienlijk beperkten. De hoeveelheid politieke en onbetrouwbare inhoud nam toe op beide platformen, de hoeveelheid inhoud die als onbeschoft of beledigend werd geclassificeerd nam af op Facebook, en de hoeveelheid inhoud van politiek gematigde bronnen nam toe op Facebook. Ondanks deze aanzienlijke veranderingen beïnvloedde de chronologische ordening de niveaus van issue polarisatie, affectieve polarisatie,

² "Gevoelspolarisatie": Negatieve gevoelens en zelfs afkeer jegens "de andere groep", mensen met andere politieke ideeën en voorkeuren (en positievere gevoelens jegens de "eigen groep").

politieke kennis of andere belangrijke attitudes van gebruikers niet.

Carter en Alford (2023) geven een theoretische verhandeling over hoe de bestrijding van disinformatie door Facebook heeft gezorgd voor een beweging weg van publieke delen van Facebook naar private groepen waar extremisme, complottheorieën en oproepen tot geweld hoogtij vieren. Ibrahim et al. (2023) kijken naar het aanbevelingssysteem van YouTube. Ze construeren zes archetypische Amerikaanse personen en onderzoeken hoe hun kijkgedrag aanbevelingen beïnvloedt. Ze concluderen dat algoritmes mensen van politieke extreme inhoud 'weglokken', maar dat dit sterker geldt voor extreem-rechts dan extreem-links. De eerder genoemde studie van Coscial en Rossi (2022) laat zien dat gerichte pogingen om (algoritmisch) conflict te reduceren kan zorgen voor polarisatie. Ludwig et al. (2022) doen een experiment in Duitsland waarbij ze kijken of aanbevelingssystemen gedreven door negatief sentiment zorgen voor meer affectieve polarisatie dan systemen gebaseerd op neutraal sentiment. De verschillen blijken klein te zijn en alleen op te treden als participanten veel tijd besteden op het aanbevelingssysteem.

Anonimiteit

In deze paragraaf richten we ons specifiek op de impact van anonimiteit op de kwaliteit van inhoud en interacties op sociale media. Als we kijken naar de wetenschappelijke studies die zijn uitgevoerd naar dit onderwerp dan is er een sterke focus op de Verenigde Staten en zijn conclusies vaak gebaseerd op observationele eenmalige studies (dus niet over de tijd heen) die vooral het gedrag van de gebruiker tussen verschillende sociale en digitale media vergelijken (Jaidka et al., 2023). Bijvoorbeeld, in deze onderzoeken wordt vaak een vergelijking gemaakt van de kwaliteit van discussies op Facebook versus Twitter (Oz et al., 2018), Facebook versus YouTube (Halpern & Gibbs, 2013), of Facebook versus nieuwssites (Rowe, 2015). Oz et al. (2018) vinden in een inhoudsanalyse dat reacties op tweets van het Witte Huis significant onbeschaafder en onbeleefder waren en minder deliberatief zijn dan reacties op Facebookberichten van het Witte Huis. Een vervolgonderzoek toont aan dat mensen bedachtzamer reageerden op Facebook-posts van het Witte Huis in vergelijking met tweets van het Witte Huis, maar er werden geen verschillen gevonden in onbeschaafdheid en onbeleefdheid. Halpern en Gibbs (2013) vinden dat Amerikaanse sociale mediagebruikers beschaafder zijn op Facebook (dit platform zien zij als minder anoniem) ten opzichte van YouTube (zien zij als meer anoniem). Fichman en Rathi (2023) vonden in Amerikaans onderzoek dat trolling (provocatie, ontsporing, persoonlijke aanvallen en belediging) vaak voorkomt op Facebook, Instagram en Twitter en dat er soms ook verschillen zijn tussen platformen, maar dat die niet lijken te komen door anonimiteit, maar meer door andere 'features' van de platformen (bijvoorbeeld hoe zichtbaar trolling is voor anderen in je netwerk). Bovendien is het in deze gevallen lastig om een goede vergelijking te maken omdat de gebruikers van deze platformen verschillen (bijvoorbeeld in demografische kenmerken) en ook omdat de 'features' van de platformen verschillen (op Twitter zijn er bijvoorbeeld minder tekens beschikbaar dan op Facebook. Ook zijn er verschillen in hoe zichtbaar je commentaar is voor anderen in je netwerk). We benoemen daarom later in dit stuk ook experimenteel onderzoek met zelfontworpen

platformen dat een sterker causaal verband laat zien tussen anonimiteit en de kwaliteit van onlinegesprekken.

Barnidge (2017) onderzoekt de aard van politiek meningsverschillen op sociale media in vergelijking met *persoonlijke* interacties en *anonieme* online omgevingen (zoals Reddit). Op basis van online vragenlijsten vindt Barnidge dat socialemediagebruikers in de Verenigde Staten meer politieke meningsverschillen ervaren dan in persoonlijke (face-to-face) interacties en in anonieme online omgevingen. Een ander onderzoek stelt een vergelijkbare vraag, maar gebruikt een experimenteel onderzoeksdesign (Combs et al., 2023). Zij kijken of anonieme online gesprekken tussen mensen met verschillende politieke opvattingen polarisatie kan vergroten of juist verkleinen. Combs en collega's (2023) ontwerpen een mobiel chatplatform en vroegen Republikeinen en Democraten in de VS om deel te nemen aan de chatsessies. Zij ontdekken dat mensen die deelnamen aan anonieme gesprekken over politieke onderwerpen tussen mensen die een verschillende partij steunden minder gepolariseerd raken in vergelijking met een controlegroep die een opstel schreef aan de hand van dezelfde gespreksaanwijzingen. Ook Jaidka et al. (2023) ontwerpen een op maat gemaakt experimenteel platform om te onderzoeken of identificeerbaarheid online de kwaliteit van online discussies beïnvloedt. Ook zij vinden geen verschil in discussiekwaliteit tussen situaties waarin deelnemers anoniem of identificeerbaar waren. Daarentegen is de kwaliteit van de discussies hoger wanneer alleen hun sociale identiteit (bijvoorbeeld als wel getoond op welke partij zij stemden) werd getoond. Bovendien hangen depolariserende effecten samen met de beleefdheid van de gesprekken tussen de deelnemers aan het onderzoek. Tot slot laat onderzoek in Duitsland uitgevoerd naar alle reacties op online petities die zijn gepubliceerd op het Duitse sociale mediaplatform www.openpetition.de zien dat niet-anonieme individuen agressiever zijn dan anonieme individuen als het gaat om online brandhaarden (Rost et al., 2016).

In hun korte overzicht van onderzoek naar de effecten van anonimiteit zien Kuehn en Salter (2020) een ontwikkeling waarin aanvankelijk de positieve aspecten van anonimiteit en de daarbij horende vrijheid werden benadrukt (Baym, 2010), maar legt meer recent onderzoek sterker de nadruk op met name de relatie tussen anonimiteit en verhoogde niveaus van onbeschoftheid. Zo hangt "trollen" over het algemeen samen met anonieme gebruikers met "nep" accounts (Edstrom, 2016; Galán-García, de la Puerta, Gómez, Santos, & Bringas, 2017). Onderzoek laat zien dat het verbieden van anonieme reacties misbruik vermindert (Rowe, 2015; Santana, 2014). Gedwongen registratie als gebruiker van een sociaal mediaplatform blijkt kwalitatief betere, maar kwantitatief minder reacties op te leveren, doordat extra inspanning vereist is en/of verantwoordingsplicht mogelijk wordt (Bakker, 2010; Santana, 2014). Rowe (2015) ontdekt dat anonieme reacties relatief onbeleefder en persoonlijker beledigend waren dan geverifieerde Facebook-reacties op dezelfde nieuwsverhalen. Ook andere studies vinden dat anonimiteit zelfregulering en verantwoordelijkheid in online discussies vermindert (Chen, 2017; Oz et al., 2018; Theocharis et al., 2016). Hille en Bakker (2014) vinden echter dat anonieme nieuwsreacties van relatief hogere kwaliteit waren dan niet-anonieme reacties op de overeenkomstige Facebook-pagina's van

nieuwsorganisaties.

Studies gericht op Nederland

In deze sectie bespreken we specifiek het onderzoek naar sociale media dat (deels) werd gedaan in de Nederlandse context. We weten dat sociale media een belangrijke rol spelen in de Nederlandse politieke sfeer. Vergelijkend onderzoek door Graham et al. (2016) toont bijvoorbeeld dat Nederlandse politici al in aanloop naar de verkiezingen van 2010 meer en vaker gebruik maakten van Twitter (nu X) dan hun Britse tegenhangers, en het platform ook frequenter gebruikten om de interactie met burgers te zoeken. De toon en mate van emotionaliteit van dat soort sociale mediabijdragen door politici en politieke partijen kunnen zowel de gevoelens als de opvattingen van burgers beïnvloeden. Onderzoek door Marquart et al. (2022) naar berichten over de EU laat zien dat effecten ook afhankelijk zijn van de politieke oriëntatie van burgers en dat niet alleen negatieve, maar juist ook positieve emoties een belangrijke, activerende rol kunnen spelen.

Niet alle politici in Nederland gebruiken sociale media op dezelfde manier. Meer populistische politici lijken bijvoorbeeld op Twitter (X) in campagnetijd veel gebruik te maken van kritische en antagonistische taal (zoals beledigingen en sarcasme) en dat soort berichten worden vervolgens vaker gedeeld en verspreid dan hun minder controversiële berichten (Gonawela et al., 2018). Populistische partijen en politici lijken daarnaast vaak emoties in te zetten op sociale media (bijvoorbeeld tijdens de COVID-crisis; Widmann, 2022), hetgeen dan weer invloed kan hebben op de emoties en attitudes van burgers (Sandberg et al., 2022).

Sociale media fungeren ook in Nederland als belangrijke (politieke) informatiebronnen voor burgers: men leert door het consumeren van politiek nieuws op sociale media (Ohme et al., 2022; dit effect is overigens sterker bij blootstelling via een desktop, dan via een smartphone). Jeroense et al. (2022) laten zien dat het niet alleen de Nederlandse burgers met meer populistische of postmaterialistische³ attitudes zijn die sociale media gebruiken om politiek nieuws te consumeren of de connectie met politici op te zoeken. Alhoewel deze groepen wel meer geneigd zijn dan anderen om ook daadwerkelijk te reageren op politieke content, lijken sociale media geen sterker effect te hebben op hun stemintenties dan het geval is bij hun minder populistische of postmaterialistische landgenoten.

Tegen de achtergrond van sociale media als belangrijke nieuwsbron, bestaan zorgen over de verspreiding van mis- en disinformatie op sociale media. Onderzoek uitgevoerd in Nederland focust veelal niet zo zeer op de daadwerkelijke verspreiding van onwaarheden, maar meer op *beschuldigingen* daarvan, bijvoorbeeld door het gebruik van labels als "fake news" en het in twijfel trekken van de betrouwbaarheid van publieke media. Zowel politici als burgers kunnen op sociale media berichten verspreiden over vermeend bedrog en onwaarheden en ze doen dit vaak in lijn

³ Mensen met sterke postmaterialistische attitudes hechten veel belang aan non-materiele onderwerpen (bijvoorbeeld zelfexpressie, politieke participatie), dan aan meer materiele onderwerpen (bvb. het voorkomen van prijsstijgingen).

met hun reeds bestaande overtuigingen. Dergelijke inhoud wordt vaker verspreid door meer populistische politici en burgers, en kunnen wantrouwen jegens de media en experts versterken en twijfel zaaien over de betrouwbaarheid van informatie (Hameleers, 2022; Hameleers & Minihold, 2020). Hoe vatbaar iemand is voor dat soort labels en beschuldigingen kan verschillen. Recent onderzoek laat zien dat de betrouwbaarheid van berichten die nieuws in twijfel trekken op zichzelf een redelijk beperkt effect hebben op hoe geloofwaardig burgers dat nieuws vinden: de bron en inhoud van het bericht spelen een grotere rol dan de beschuldiging. De effectiviteit van een dergelijke beschuldiging is afhankelijk van of het desbetreffende nieuwsbericht wel of juist niet in lijn is met bestaande ideeën en overtuigingen (Bos et al., 2023; Janssen & van Gog, 2023).

Het aankaarten van misinformatie kan natuurlijk ook terecht zijn en worden ingezet als maatregel om het geloof in en de verspreiding van foutieve informatie tegen te gaan. Dit soort interventies kunnen echter ook ongewenste effecten hebben, zoals een overschatting van de hoeveelheid misinformatie, en verminderd vertrouwen in informatie die wel accuraat is (Hameleers, 2023).

Met betrekking tot polarisatie laat onderzoek zien dat er parallellen te trekken zijn tussen verdelend taalgebruik van politici op Twitter (X) en vergelijkbare sentimenten in posts door burgers in Facebookgroepen. De voornaamste tweedeling wordt in Nederland echter niet -zoals in de Verenigde Staten- getrokken op basis van partijvoorkeur, maar eerder op basis van het gevoel wel of juist niet te behoren tot een groep achtergestelde, niet vertegenwoordigde burgers (Hameleers, 2020). Onderzoek naar Tweets over één controversieel onderwerp (Zwarte Piet) tussen 2011 en 2019, laat echter zien dat verdeeldheid op basis van identiteit ("wij tegen zij") niet de boventoon voert. Tegenstellingen op basis van het issue zelf (voorstander of tegenstander) bleken minstens even belangrijk, wat aangeeft dat de inhoud er in gepolariseerde discussies nog steeds toe lijkt te doen (Wijngaert, 2022). Hoewel polarisatie en polariserende berichten inderdaad terug te vinden zijn op sociale media, is het nog maar zeer de vraag of het sociale media zijn die polarisatie aanwakkeren, of dat polarisatie misschien het gebruik van sociale media in de hand werkt. Nordbrandt (2023) onderzocht het causale verband tussen het gebruik van sociale media en affectieve polarisatie in Nederland. Zij toont aan dat het gebruik van sociale media affectieve polarisatie niet versterkt (of vermindert), maar dat hoe gepolariseerd iemand is wel kan beïnvloeden hoe vaak die persoon sociale media gebruikt (afhankelijk van hoe actief die persoon al was).

Conclusie

Het onderzoek naar de invloed van sociale media op democratie is veelomvattend en er kan – misschien niet verrassend – geen éénduidig antwoord gegeven worden op de vraag of deze invloed positief of negatief is. Ons overzicht is in lijn met andere overzichten van de literatuur, die ook niet éénduidig de positieve of negatieve aspecten benadrukken (bijvoorbeeld Lorenz-Spreen, 2023). Het antwoord hangt af van *welke* aspecten van sociale media en democratie in beschouwing worden genomen en in welke context. Het feit dat veel onderzoek in de Amerikaanse context is

uitgevoerd maakt een vertaling naar Nederland niet eenvoudig.

Toch kunnen er wel een aantal voorzichtige conclusies worden getrokken, specifiek voor risico's als algoritmische selectie (en bijbehorende verdienmodellen) en anonimiteit. Aanpassingen in algoritmische selectie kunnen enerzijds een positieve invloed hebben op de kwaliteit van informatie en discussie, maar ook informatieconsumptie en online participatie verminderen. Het onderzoek naar anonimiteit is in het bijzonder Amerikaans geïntereerd en berust gedeeltelijk voor een vergelijking tussen platformen, waarbij verschillen niet altijd met zekerheid aan verschillen in anonimiteit kunnen worden toegeschreven. Toch wijst het onderzoek vaker op de negatieve effecten van anonimiteit die in zijn algemeenheid niet zorgt voor meer beleefde en beschaafde gesprekken.

Hoewel er in kwantitatieve termen veel onderzoek is naar de relatie tussen sociale media en democratie ontbreken op een aantal punten nog duidelijke inzichten. Zeker in de Nederlandse context ontbreekt nog een systematisch beeld van de aanwezigheid en verspreiding van *hate speech* en andere schadelijke vormen van communicatie. Daarnaast is ook de nog veel onbekend over de invloed van sociale media op processen van polarisatie en hoe kenmerken van het platform (die bijvoorbeeld anonimiteit mogelijk maken) en met name de inhoud invloed uitoefenen op (gevoelens van) verdeeldheid in de samenleving.

Literatuur

Barnidge, M. (2017). Exposure to political disagreement in social media versus face-to-face and anonymous online settings. *Political Communication*, 34(2), 302-321.

<https://doi.org/10.1080/10584609.2016.1235639>

Bos, L., Laura Egelhofer, J., & Lecheler, S. (2023). Short but Critical?: How "Fake News" and "Anti-Elitist" Media Attacks Undermine Perceived Message Credibility on Social Media. *Communication Research*, 50(6), 695-719. <https://doi.org/10.1177/00936502231178432>

Boulianne, S. (2015). Social media use and participation: A meta-analysis of current research. *Information, Communication & Society*, 18(5), 524-538.

<https://doi.org/10.1080/1369118X.2015.1008542>

Carter, J. S., & Alford, C. (2023). Adoxastic publics: Facebook and the loss of civic strangeness. *Quarterly Journal of Speech*, 109(2), 176-198. <https://doi.org/10.1080/00335630.2022.2139856>

Chen, G. M. (2017). Online incivility and public debate: Nasty talk. *Springer*.

Combs, A., Tierney, G., Guay, B., Merhout, F., Bail, C. A., Hillygus, D. S., & Volfovsky, A. (2023). Reducing political polarization in the United States with a mobile chat platform. *Nature Human*

Behaviour, 7(9), 1454-1461. <https://doi.org/10.1038/s41562-023-01655-0>

Coscia, M., & Rossi, L. (2022). How minimizing conflicts could lead to polarization on social media: An agent-based model investigation. *PloS One*, 17(1), e0263184. <https://doi.org/10.1371/journal.pone.0263184>

Darwish, K., Alexandrov, D., Nakov, P., & Mejova, Y. (2017). Seminar users in the Arabic Twitter sphere. In *Social Informatics: 9th International Conference, SocInfo 2017, Oxford, UK, September 13-15, 2017, Proceedings, Part I* (pp. 91-108). Springer International Publishing.

Edstrom, M. (2016). The trolls disappear in the light: Swedish experiences of mediated sexualized hate speech in the aftermath of Behring Breivik. *International Journal for Crime, Justice and Social Democracy*, 5(2), 96-106.

Fichman, P., & Rathi, M. (2023). Trolling CNN and Fox News on Facebook, Instagram, and Twitter. *Journal of the Association for Information Science and Technology*, 74(5), 493-505. <https://doi.org/10.1002/asi.24753>

Galán-García, P., de la Puerta, J. G., Gómez, C. L., Santos, I., & Bringas, P. G. (2017). Supervised machine learning for the detection of troll profiles in Twitter social networking. *Logic Journal of the IGPL*, 24(1), 42-53. https://doi.org/10.1007/978-3-319-01854-6_43

González-Bailón, S., Lazer, D., Barberá, P., Zhang, M., Allcott, H., Brown, T., ... & Tucker, J. A. (2023). Asymmetric ideological segregation in exposure to political news on Facebook. *Science*, 381(6656), 392-398. <https://doi.org/science.ade7138>

Guess, A. M., Malhotra, N., Pan, J., Barberá, P., Allcott, H., Brown, T., ... & Tucker, J. A. (2023). Reshares on social media amplify political news but do not detectably affect beliefs or opinions. *Science*, 381(6656), 404-408. <https://doi.org/10.1126/science.add8424>

Halpern, D., & Gibbs, J. (2013). Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Computers in Human Behavior*, 29(3), 1159-1168. <https://doi.org/10.1016/j.chb.2012.10.008>

Hille-Knoester, S., & Bakker, P. (2014). Engaging the social news user: Comments on news sites and Facebook. *Journalism Practice*, 8(5), 563-572.

Ibrahim, H., AlDahoul, N., Lee, S., Rahwan, T., & Zaki, Y. (2023). YouTube's recommendation algorithm is left-leaning in the United States. *PNAS nexus*, 2(8), pgad264. <https://doi.org/10.1093/pnasnexus/pgad264>

- Jaidka, K., Mukerjee, S., & Lelkes, Y. (2023). Silenced on social media: The gatekeeping functions of shadowbans in the American Twitterverse. *Journal of Communication*, 73(2), 163-178. <https://doi.org/10.1093/joc/jqac050>
- Kuehn, K. M., & Salter, L. A. (2020). Assessing digital threats to democracy, and workable solutions: A review of the recent literature. *International Journal of Communication*, 14, 22.
- Lorenz-Spreen, P., Oswald, L., Lewandowsky, S., & Hertwig, R. (2023). A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature human behaviour*, 7(1), 74-101. <https://doi.org/10.1038/s41562-022-01460-1>
- Ludwig, K., Grote, A., Iana, A., Alam, M., Paulheim, H., Sack, H., ... & Müller, P. (2023). Divided by the algorithm? The (limited) effects of content- and sentiment-based news recommendation on affective, ideological, and perceived polarization. *Social Science Computer Review*. <https://doi.org/10.1177/08944393221149290>
- Marquart, F., Brosius, A., & de Vreese, C. (2022). United Feelings: The Mediating Role of Emotions in Social Media Campaigns for EU Attitudes and Behavioral Intentions. *Journal of Political Marketing*, 21(1), 85–111. <https://doi.org/10.1080/15377857.2019.1618429>
- Nordbrandt, M. (2023). Affective polarization in the digital age: Testing the direction of the relationship between social media and users' feelings for out-group parties. *New Media & Society*, 25(12), 3392–3411. <https://doi.org/10.1177/14614448211044393>
- Nyhan, B., Settle, J., Thorson, E., Wojcieszak, M., Barberá, P., Chen, A. Y., ... & Tucker, J. A. (2023). Like-minded sources on Facebook are prevalent but not polarizing. *Nature*, 620(7972), 137-144. <https://doi.org/10.1038/s41586-023-06297-w>
- Ohme, J., Maslowska, E., & Mothes, C. (2022). Mobile News Learning—Investigating Political Knowledge Gains in a Social Media Newsfeed with Mobile Eye Tracking. *Political Communication*, 39(3), 339–357. <https://doi.org/10.1080/10584609.2021.2000082>
- Oz, M., Zheng, P., & Chen, G. M. (2018). Twitter versus Facebook: Comparing incivility, impoliteness, and deliberative attributes. *New Media & Society*, 20(9), 3400-3419. <https://doi.org/10.1177/1461444817749516>
- Putnam, R. D. (2000). Bowling alone: The collapse and revival of American community. *Simon and Schuster*.
- Rost, K., Stahel, L., & Frey, B. S. (2016). Digital social norm enforcement: Online firestorms in social media. *PLoS One*, 11(6), e0155923. <https://doi.org/10.1371/journal.pone.0155923>

Rowe, I. (2015). Civility 2.0: A comparative analysis of incivility in online political discussion. *Information, Communication & Society, 18*(2), 121-138.

<https://doi.org/10.1080/1369118X.2014.940365>

Sandberg, L., Jacobs, K., & Spierings, N. (2022). Populist MPs on Facebook: Adoption and emotional reactions in Austria, the Netherlands, and Sweden. *Scandinavian Political Studies, 45*(4), 504–528. <https://doi.org/10.1111/1467-9477.12239>

Santana, A. D. (2014). Virtuous or vitriolic. *Journalism Practice, 8*(1), 18–33.

<https://doi.org/10.1080/17512786.2013.813194>

Theocharis, Y., Barberá, P., Fazekas, Z., Popa, S. A., & Parnet, O. (2016). A bad workman blames his tweets: The consequences of citizens' uncivil Twitter use when interacting with party candidates. *Journal of Communication, 66*(6), 1007-1031. <https://doi.org/10.1111/jcom.12259>

Widmann, T. (2022). Fear, Hope, and COVID-19: Emotional Elite Rhetoric and Its Impact on the Public During the First Wave of the COVID-19 Pandemic. *Political Psychology, 43*(5), 827–850.

<https://doi.org/10.1111/pops.12831>

Wijngaert, L. van de. (2022). Understanding polarization: A case study of Black Pete in the Netherlands. *Psychology of Language and Communication, 26*(1), 399–414.

<https://doi.org/10.2478/plc-2022-19>



Correspondentie adres voor dit rapport:
Wageningen University & Research
Strategic Communication Group
Postbus 8130
6700 EW Wageningen
T 0317 48 43 10

De missie van Wageningen University & Research is 'To explore the potential of nature to improve the quality of life'. Binnen Wageningen University & Research bundelen Wageningen University en gespecialiseerde onderzoeksinstituten van Stichting Wageningen Research hun krachten om bij te dragen aan de oplossing van belangrijke vragen in het domein van gezonde voeding en leefomgeving. Met ongeveer 30 vestigingen, 7.600 medewerkers (6.700 fte) en 13.100 studenten en ruim 150.000 Leven Lang Leren-deelnemers behoort Wageningen University & Research wereldwijd tot de aansprekende kennisinstellingen binnen haar domein. De integrale benadering van de vraagstukken en de samenwerking tussen verschillende disciplines vormen het hart van de unieke Wageningen aanpak.