

De Vrije Universiteit Amsterdam, School of Business and Economics, staat aan de start van een samenwerking met bedrijven om wetenschappelijke inzichten uit digitale gedragsinzichten te verkrijgen. Deze hebben als doel om beter zicht te krijgen op het werkende mechanisme achter gedragsinterventies en de generaliseerbaarheid van deze interventies in verschillende contexten over tijd. In dit artikel licht Eline van Baal toe waarom deze samenwerking gestart wordt en hoe die eruit ziet.

POTENTIE VOOR WETENSCHAP

GEBRUIK VAN COMMERCIEËLE GEDRAGS- EXPERIMENTEN

Wie zich online op een website begeeft, bevindt zich al snel in een experiment. Zo kan het voorkomen dat u een andere weergave van een en dezelfde website ziet dan ik. Denk hierbij aan een andere geschreven boodschap, een andere titel of een andere prijsweergave. Bedrijven registreren wat het verschil in weergave doet met ons gedrag. Het op deze manier uitzetten van een versie A tegen een versie B wordt AB-testen genoemd.

AB-testen heeft als doel om het effect van kleine veranderingen op gedrag te meten. Op deze manier helpen AB-testen bedrijven om een website te ontwikkelen die 'goed' functioneert en leren ze bovendien wat wel en niet werkt. Het doel van een AB-test is doorgaans het verhogen van de verkoop, ook wel conversie-ratio-optimalisatie genoemd. Naarmate AB-testen echter omvangrijker wordt en bedrijven steeds meer inzichten verzamelen, neemt de interesse in gedrag en gedragsinzichten ook toe. Psychologie, en met name keuzegedrag van de consument, is voor bedrijven ineens concreet en meetbaar geworden.

Niet geheel verrassend neemt daarmee ook de interesse in het vakgebied (sociale) psychologie toe. Deze verhoogde

interesse is niet geheel onomstreden. Zo weten we in de sociale psychologie nog altijd niet goed wat het effect van verschillende gedragsinterventies in verschillende omgevingen is. Laat staan dat we een goed zicht hebben op de langetermijneffecten van gedragsinterventies. We hebben te maken met een groeiende interesse enerzijds, maar een gat in de wetenschap anderzijds. Zijn beide werelden verenigbaar?

AB-testen worden thans op grote schaal uitgevoerd. Boodschappen worden steeds vaker online besteld bij een supermarkt, kleding wordt in grote mate verscheept naar de Nederlandse huizen (CBS, 2020) en ook sociale contacten worden steeds meer onderhouden via allerlei sociale media op het virtuele net (CBS, 2019). In al deze omgevingen vinden volop AB-testen en dus digitale experimenten plaats. Met de replicatiecrisis (Makel, 2012) nog in ons achterhoofd en het tekort aan veldstudies in de hedendaagse wetenschap (Gneezy, 2017) biedt het uitvoeren van AB-testen kansen voor de psychologie als wetenschap.

KEUZEARCHITECTUUR

Of we nu thuis in alle rust op de bank onze boodschappen bij elkaar klikken in een online supermarkt of met een mond-

kapje rond zes uur 's avonds in de supermarkt staan: welke keuzes we maken en hoe we deze keuzes maken, wordt in belangrijke mate beïnvloed door de omgeving waarin we ons bevinden. Nobelprijswinnaar Richard Thaler refereert met de term 'choice architecture' aan deze invloed van de omgeving op een keuze. Een tweede term die hij hierbij introduceerde was *nudge*. *Nudge* duidt op een verandering in de omgeving om tot ander gedrag te komen (Thaler, 2008). Belangrijk hierbij is dat *nudging* zou leiden tot het gewenste gedrag zonder dat de keuzevrijheid beperkt wordt of er sprake is van straffen, belonen of andere regelgevingen (Thaler, 2008). De terminologie '*nudge*' en het '*nudgen van gedrag*' is sinds de introductie populair bij zowel overheden als bedrijven. Zo zijn er verschillende *behavioural insights teams* opgericht en is er zelfs een Europees samenwerkingsverband tussen verschillende landen op het gebied van overheid en "*nudgen*" (Behavioural Insight Network: <https://www.binnl.nl/>).

Een bekend wetenschappelijk voorbeeld van een succesvolle *nudge* is het veranderen van de standaardoptie in een pensioenplan: van standaard niet meedoen naar standaard wel meedoen. Deze simpele verandering behoudt de keuzevrijheid, staat los van straffen en belonen en zorgt voor een grotere deelname aan het pensioenplan (Madrian & Shea, 2001). Hiermee in lijn heeft de Dienst Uitvoering Onderwijs (DUO) geëxperimenteerd met het veranderen van de keuze architectuur voor het aanvragen van de studentleningen. Voorheen was de standaardoptie 'maximaal lenen' makkelijk aan te vinken. Door deze optie te veranderen naar een leeg invulveld waarin studenten zelf het te lenen bedrag in moeten vullen nam de hoogte van de leningen onder Nederlandse studenten aanzienlijk af.

Deze voorbeelden illustreren mooi waar *nudgen* in eerste instantie voor bedoeld is: het veranderen van de omgeving om mensen te stimuleren andere of 'betere' keuzes te maken. Andere domeinen hierin zijn het verbeteren van de gezondheid (Li & Chapman, 2013), duurzaam gedrag te stimuleren (Ebeling & Lotz, 2015), pro-sociaal gedrag te stimuleren (Goswami & Urminsky, 2016) en financieel gezond gedrag te stimuleren (Choi et al., 2017).

WETENSCHAPPELIJKE TEKORTKOMINGEN

Ondanks de groeiende populariteit van *nudging* is het gebruik ervan niet geheel onomstreden. Zo laat een meta-analyse over meerdere *nudging*-studies zien dat 74% ervan voornamelijk is uitgevoerd om de effectiviteit van de interventies te onderzoeken in één bepaalde setting (Szasz et al., 2017). Daarmee weten we nog weinig over de werking van verschillende gedragsinterventies in verschillende settings. Daarnaast richt slechts 24% van het onderzoek zich op het verklaren van de moderaties. Het gevonden effect is in veel gevallen nog moeilijk te koppelen aan een verklaring. Verder is er voor slechts 7% van de studies een *power analyse* uitgevoerd waarmee de validiteit en generaliseerbaarheid van de bevindingen nog verder in het geding komen.

Op dit moment is het onderzoeksveld dat zich bezighoudt met *nudging* nog te beperkt om met overtuigend bewijs te komen. Er is meer aandacht nodig voor de werkzaamheid van *nudges* en voor het verklaren en voorspellen van de werking en effectiviteit van bepaalde *nudges* (Szasz et al., 2017). De conclusie is in lijn met een meer algemene bevinding in de gedragswetenschap: er is een grote behoefte aan meer veldonderzoek om besluitvormingsprocessen beter te begrijpen en om externe validiteit voor gedragingen te verhogen (Gneezy, 2017).

Daar komt nog bij dat er in grote mate beperkt onderzoek is gedaan naar de langetermijngevolgen van *nudges*. Onderzoek richt zich op de uitkomsten van de gedragsinterventies direct na de introductie van de interventie. Daarbij wordt niet in acht genomen wat en hoe het effect van deze gedragsinterventies veranderen over tijd (Frey & Rogers, 2014).

Een tekenend voorbeeld van een gedragsinterventie die op de korte termijn positief lijkt maar op de langere termijn negatief is, komt naar voren in onderzoek van Etkin (2016). Etkin onderzocht het effect van het kwantificeren van gedrag om tot betere gedragsinzichten te komen. Denk daarbij aan het meten van de eigen activiteit, zoals het meten van een *smart-watch meten* van het aantal genomen stappen op een dag. In zes verschillende experimenten toonde Etkin aan dat het meten van activiteiten ervoor zorgt dat de activiteit vaker

'AB-testen kunnen helpen om wetenschappelijke bevindingen die op het eerste gezicht robuust lijken te repliceren in het veld en in verschillende contexten'

‘Bij het horen van de termen ‘data’, ‘experimenten’ en ‘bedrijven’ gaan bij menigeen nogal eens de nekharen overeind staan’

ondernomen wordt op de korte termijn (lopen, lezen), maar ook dat het plezier dat mensen aan de activiteit beleven afneemt en op de langere termijn ook de activiteit. Als mogelijke verklaring geeft Etkin het ondermijnen van de intrinsieke motivatie. Door de aandacht te richten op de output kan het gebeuren dat de activiteit gaat voelen als werk of verplichting. Dat doet het plezier afnemen dat mensen uit het ondernemen van de taak halen.

Al bij al is er meer onderzoek nodig waarin niet alleen grotere steekproeven worden gebruikt, maar waarin ook aandacht is voor de precieze werking van de nudge en de daarmee gepaard gaande langetermijneffecten.

GROTERE STEEKPROEVEN

Commerciële AB-testen kunnen de grotere steekproeven bieden die nu nog ontbreken. AB-testen kunnen helpen om wetenschappelijke bevindingen die op het eerste gezicht robuust lijken te repliceren in het veld en in verschillende contexten. Door resultaten van eerder uitgevoerde AB-testen te analyseren in een systematische meta-analyse ontstaat er beter zicht op de generaliseerbaarheid en contextafhankelijkheid van gedragsinterventies. Daarnaast dienen de opgedane inzichten als input voor weer nieuwe hypothesen. Door nieuwe hypothesen zorgvuldig te onderzoeken met zowel veldstudies (nieuwe AB-testen) als gecontroleerde laboratorium studies wordt elke hypothese onderzocht met meerdere experimenten in verschillende contexten. Op deze manier worden er inzichten verworven voor het waargenomen gedragseffect na verloop van tijd en in verschillende contexten. Daarnaast is er op deze manier meer aandacht voor het mechanisme achter de gedragsinterventie en daarmee het potentiële langetermijneffect.

ETHISCHE BEZWAREN

Ondanks de potentie voor wetenschappelijk onderzoek laat het uitvoeren van AB-testen niet iedereen onberoerd. Laat staan het uitvoeren van AB-testen door psychologen en voor wetenschappelijk onderzoek. Bij het horen van de termen ‘data’, ‘experimenten’ en ‘bedrijven’ gaan bij menigeen nogal

eens de nekharen overeind staan. Al helemaal als mensen erachter komen dat ze zelf vaak participant zijn in deze experimenten en daar geen weet van hebben. Een begrijpelijke reactie op een fenomeen waarin we zelf (ongevraagd) deelnemer zijn. Het belang van een grotere bekendheid van deze experimenten is daarmee ook groot. Zodat we als maatschappij en als politiek een grotere invloed krijgen op (de uitvoer van) deze experimenten.

Wanneer de kennis over de aanwezigheid van AB-testen wel aanwezig is, gaan de gedachten al snel uit naar privacy gerelateerde issues. Privacy is een groot goed en de zorgen zijn niet onterecht. Psychologen worden opgeleid om participanten te betrekken bij het onderzoekstraject en op de hoogte te stellen van het doel van het onderzoek en de resultaten. Bij AB-testen gebeurt dit allemaal niet. Toch is het belang van actief betrokken psychologen bij het uitvoeren en analyseren van AB-testen groot. Net als het belang van een brug slaan tussen het bedrijfsleven en de wetenschap. De experimenten vinden op steeds grotere schaal plaats en zonder gedragskennis is het gissen naar het effect op mensen. Laat staan het langetermijneffect op bijvoorbeeld welzijn en maatschappelijke vraagstukken als duurzaamheid. Zolang we niet weten wat precies het effect op gedrag is van de keuzearchitectuur *nudging* en *nudges* precies op gedrag is, weten we ook niet hoe we beleid zo kunnen inrichten dat mensen beschermd of geholpen kunnen worden om daadwerkelijk ‘de juiste’ keuze te maken.

Wanneer bedrijven geen AB-testen afnemen, doen ze geen inzichten op en dat voelt wellicht beter waar het gaat om privacy-issues. Maar dan zit iedereen effectief alsnog in de A-conditie van een experiment, maar dan zonder het leereffect. We zullen dan niet weten wat het effect is van de setting van de keuzearchitectuur op gedrag. Bedrijven bouwen geen kennis op, maar zullen op deze manier evenmin weten of ze een mogelijk negatief effect veroorzaken met de vormgeving van de website en berichtgeving op het welzijn van mensen.

Een pakkend voorbeeld hierbij is de banner ‘Let op, geld lenen kost geld’. In het leven geroepen door de Autoriteit

‘Wie zich online op een website begeeft, bevindt zich al snel in een experiment’

Financiële Markten (AFM) met als doel om consumenten te weerhouden van het afsluiten van een lening. AB-testen die in de financiële sector werden uitgevoerd, gaven indicatie van mogelijk averechtse effecten van deze banner (Online Dialogue). De AFM heeft een gedragsonderzoek uitgevoerd naar de effectiviteit van deze banner en vond geen significante verschillen (AFM, 2016). Een beleidsmaatregel, helaas zonder het gewenste effect, misschien zelfs het averechtse effect. Op dit moment zorgt het beperkte zicht op het mechanisme achter de werking van gedragsinterventies voor een hoop onduidelijkheid omtrent wat ‘goed’ is en wat mensen helpt om ‘betere’ keuzes te maken.

Zonder gedragskennis voeren bedrijven AB-testen uit voor een hogere verkoop en worden websites steeds beter in het verleiden van mensen tot het maken van keuzes die ze misschien helemaal niet willen maken. Hierin zijn we mogelijk de persoonlijke privacy-gerelateerde issues voorbij.

Het zijn de data van alle individuen samen die bedrijven inzicht geven in onze besluitvormingsprocessen. Het zijn onze gezamenlijke data die voor bedrijven interessant zijn. Om van invloed te zijn op het effect van AB-testen en de behaalde doelen ermee zijn een goed beleid en duidelijke kaders noodzakelijk. Psychologen en de wetenschap kunnen hierin een duidelijke rol vervullen door gedragskennis uit te breiden en over te brengen. Zodat AB-testen aansluiten bij waar interventies in de keuzearchitectuur oorspronkelijk voor bedoeld zijn: het verbeteren van menselijk welzijn op financieel-, gezondheids- en duurzaamheidsgebied.

Wordt vervolgd!

OVER DE AUTEUR

Eline van Baal, Msc, is promovenda aan de Vrije Universiteit Amsterdam. Daarvoor werkte ze bij Online Dialogue waar ze zich bezighield met het uitvoeren van online gedragsexperimenten.

Ook is zij lid van de redactieraad van De Psycholoog.

E-mail: elinevbaal@gmail.com.

Literatuur

- Autoriteit Financiële Markten (2016). ‘Let op! Geld lenen kost geld’ geen onmiddellijk effect in verkoopomgeving. <https://www.afm.nl/nl-nl/professionals/nieuws/2016/dec/geld-lenen-geen-effect>
- Centraal Bureau voor de Statistiek (2020). Recordbestedingen bij Europese webwinkels. <https://www.cbs.nl/nl-nl/nieuws/2020/43/recordbestedingen-bij-europese-webwinkels>
- Centraal Bureau voor de Statistiek (2019). Meisjes vaker op sociale netwerken dan jongens. <https://www.cbs.nl/nl-nl/nieuws/2019/22/meisjes-vaker-op-sociale-netwerken-dan-jongens>
- Choi, J.J., Haisley, E., Kurkoski, J. & Massey, C. (2017). Small cues change savings choices. *Journal of Economic Behavior & Organization*, 142, 378-395.
- Ebeling, F. & Lotz, S. (2015). Domestic uptake of green energy promoted by opt-out tariffs. *Nature Climate Change*, 5, 868-871.
- Frey, E. & Rogers, T. (2014). Persistence: How Treatment Effects Persist After Interventions Stop. *Policy Insights from the Behavioral and Brain Sciences*, 1, 172-179.
- Gneezy, A. (2017). Field Experimentation in Marketing Research, *Journal of Marketing Research*, 54, 140-143.
- Goswami, I. & Urminsky, O. (2016). When Should the Ask Be a Nudge? The Effect of Default Amounts on Charitable Donations. *Journal of Marketing Research*, 53, 829-846.
- Li, M. & Chapman, G.B. (2013). Nudge to health: Harnessing decision research to promote health behavior. *Social and Personality Psychology Compass*, 7, 187-198.
- Madrian, B.C. & Shea, D.F. (2001) The Power Of Suggestion: Inertia In 401(k) Participation And Savings Behavior. *Quarterly Journal of Economics*, 116, 1149-1187.
- Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in Psychology Research: How Often Do they Really Occur? *Perspectives on Psychological Science*, 7, 537-542.
- Szaszi, B., Palinkas, A., Szollosi, A., Palfi, B. & Aczél, B. (2018) A Systematic Scoping Review of Choice Architecture Interventions: towards a better understanding why and when nudges work. *Journal of Behavioral Decision Making*.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New Haven, CT: Yale University Press