

## Hoe onderzoek je het denken van reizigers

### Inleiding

De filosoof Descartes had het niet zo met meningen, zelfs niet met meningen van de wetenschap. Gemakshalve ging hij ervan uit dat alles wat wie dan ook zei niet klopte. De psycholoog Piet Vroon had het ook niet zo met meningen. De mens zegt niet wat hij denkt en denkt niet wat hij zegt. Maar Descartes en Vroon zijn dood en het aantal telefonische, web-, journaal- en straat-enquêtes is groter dan ooit.

Keypoint vroeg 1121 reizigers wat beter is: de vertrektijd van een bus of een afteltijd op een dynamisch paneel. Het antwoord is: 'Vertrektijd is beter'. Dat is de overheersende mening. Keypoint stak zijn nek uit en publiceerde het verhaal in Verkeerskunde 04/2008 (Keuning en Roding, DRIS: lezen en wegwezen, Verkeerskunde 04/2008, p. 26-30). Het is leerzaam te zien hoe men tot deze foute conclusie komt.

#### *Figuur 1. De borden die Keypoint onderzocht*

Begrijpt men bord 1 (aan de hand van de twee testvragen 7 en 8)?

Lijn	Bestemming	14:01	Vertrek	Info
2	Dubbeldam		14:02	
11	Papendrecht centrum		14:04	
9	Gezondheidspark		14:05	+1

Begrijpt men bord 3 (aan de hand van de testvraag 11)?

Lijn	Bestemming	14:01	Vertrek
2	Dubbeldam		1 min
11	Papendrecht centrum		3 min
9	Gezondheidspark		5 min

Vertrektijd

Afteltijd

## 1. Inzicht met literatuur

In de literatuur kan men zien of het wiel dat je zoekt al uitgevonden is. Ebbink (Ebbink, B.A.J. (2007). De reiziger centraal!, een onderzoek naar een nieuwe methodiek om het openbaar vervoer op de wensen van de (potentiële) reiziger aan te laten sluiten. Amsterdam: Universiteit van Amsterdam. ( [www.railpedia.nl/download/attachments/2097229/Bart+Ebbink-jul07.pdf?version=1](http://www.railpedia.nl/download/attachments/2097229/Bart+Ebbink-jul07.pdf?version=1) ) heeft bijvoorbeeld gevraagd welke informatie reizigers op een station willen. Resultaat: de dynamische borden moeten nieuwsflitsen, weer- en hooikoortsberichten geven. Je zou dus kunnen concluderen dat je helemaal geen vertrek- of afteltijd moet tonen. Een andere conclusie kan zijn dat je dit soort zaken beter niet aan reizigers kunt vragen. Verder is onlangs in de Nederlandse vakliteratuur (Heuts, P., 'De psychologie achter informatieborden' [OV-Magazine 30 augustus 2007](http://OV-Magazine 30 augustus 2007)) afteltijd duidelijk als de beste oplossing geponeerd. In het artikel zwijgt Keypoint over deze contrasterende standpunten.

## 2. Inzicht met theorie



Het kan zijn dat de literatuur geen antwoord geeft. Je kunt dan de handelingen (Zie bijvoorbeeld [www.humanefficiency.nl/gui/handeling\\_psychologie.htm](http://www.humanefficiency.nl/gui/handeling_psychologie.htm) en [www.humanefficiency.nl/designers\\_understanding.htm](http://www.humanefficiency.nl/designers_understanding.htm) ) die de reizigers met ogen, geheugen

en verstand uitvoeren analyseren. Uit het onderzoek van Keypoint blijkt dat 12% aftellen niet begrijpt.

Iedereen kent het aftellen van de sport, tv-spelletjes en vertrekkende raketten. Zouden reizigers het aftellen echt niet begrijpen? Kinderen kunnen eerder aftellen dan klokkijken. Kan meer dan 12% van de busreizigers dan niet klokkijken? Maar juist als ze niet kunnen klokkijken zou ik aftellen. Dat is gemakkelijker en er is dan meer kans dat ze begrijpen wanneer de bus vertrekt. Stel dat de reiziger het niet begrijpt, hij ziet het toch voor zijn ogen gebeuren. Niet één keer maar telkens als er een nul verschijnt vertrekt een bus. Stel dat de reiziger het dan nog niet snapt, zet er dan desnoods bij: 'Bus vertrekt over 4 minuten'. Op het bord dat Keypoint in het onderzoek gebruikte werd alles uitgelegd, behalve de afteltijd. Opmerkelijk is dat de onderzoekers de ondervraagden het juiste antwoord in de mond leggen met de vraag of zij aftellen beter vinden omdat dat minder denkhandelingen geeft. De ondervraagden koppen de bal in en 58% zegt ja. Toch is de conclusie van de onderzoekers dat vertrektijd de voorkeur heeft.

We analyseren de handelingen daarom maar even verder. Figuur 2 laat zien dat het aantal handelingen bij vertrektijd veel groter is dan bij afteltijd. Er bestaat veel literatuur over het aantal fouten dat mensen bij dit soort opgaven maakt. Per cijfer maakt men grofweg 1% waarneem- en onthoudfouten (Vroemen, M. (1988). **Bestemmingen op treinkaartplaatsen automaten**, Eenknopsbediening of numeriek codesysteem. Utrecht: NV Nederlandse Spoorwegen, Pz 2.4. /CAE, stage-verslag). Dit komt overeen met de resultaten van Keypoint zelf: 15% voor de totale vertrektijdberekening. Ook blijkt uit het onderzoek van Keypoint dat er minder handelingen; c.q. minder fouten zijn, namelijk 12%. Daar komt nog bij dat sommetjes in een enquête gemakkelijker zijn dan sommetjes in een openbaarvervoering met slechtere zicht-, taak- en denkomstandigheden. Nog meer aanwijzingen dat aftellen waarschijnlijk beter is.

*Figuur 2. Psychologische handelingen bij vertrektijd en afteltijd*

Vertrektijd	Afteltijd
	
vertrektijd zoeken en lezen (vier cijfers)	afteltijd zoeken en lezen (maximaal 2 cijfers)
de actuele tijd zoeken en lezen (vier cijfers).	
de betrouwbaarheid van de actuele tijd bepalen (loopt die klok gelijk?)	
In het kortetermijngeheugen berekenen hoeveel tijd er nog over is	
vertraginginformatie zoeken en lezen	
in het kortetermijngeheugen de berekende resttijd corrigeren	resttijd zien
afstand tot bus schatten, loopsnelheid berekenen	afstand tot bus schatten, loopsnelheid berekenen
besluiten: bus of koffie halen	besluiten: bus of koffie halen

### 3. Inzicht met onderzoek

Theorie is maar theorie en daar houden we tegenwoordig niet van. Dus gaan we onderzoeken.

#### 3.1 Variabelen of borden onderzoeken

Meestal maken enquêteurs het zich gemakkelijk en vragen gewoon wat men ervan vindt. Keypoint heeft de moeite genomen twee borden te ontwerpen en heeft reizigers deze borden laten gebruiken. In het onderzoek worden deze borden vergeleken. Bij cognitief-psychologisch onderzoek kijk je naar het *effect van één variabele*. Als je vertrektijd of afteltijd wilt vergelijken, moet je dat zo doen dat de situaties die de proefpersonen aangeboden krijgen voor beide variabelen gelijk zijn. Keypoint kijkt naar het *effect van borden*. Deze borden verschillen op meer variabelen. Je weet dan niet door welk verschil de proefpersonen beter presteren. Het zou het wel eens kunnen dat de proefpersonen de voorkeur gaven aan vertrektijd omdat daarbij ook de vertraging zichtbaar was. Op het aftelbord was *niet te zien* of er vertraging was en of deze in de afteltijd verwerkt was. Het aftelbord toont ook de huidige tijd. Deze is niet nodig bij aftellen. Hoe minder getallen een bord heeft hoe beter een bord beoordeeld wordt. Het is niet eerlijk afteltijd op te zadelen met een nadeel van vertrektijd.

Goed aan het Keypoint onderzoek is dat men niet alleen meningen vraagt maar de reizigers ook met de borden laat werken. Welke vraag zou je nu moeten stellen om te weten hoe reizigers presteren bij afteltijd. De vraag die Keypoint stelt is: "Hoe laat vertrekt lijn 9 richting Gezondheidspark *werkelijk* vanaf de bushalte?" De performance op aftellen wordt gemeten via de vertrektijd. Dat betekent dat de nadelen van vertrektijd ineens ook gelden voor aftellen, bijvoorbeeld de kans dat er meer reken- en leesfouten gemaakt worden. De juiste vraagstelling is iets in de geest van: "Hier heeft u een foto van een busstation. U moet snel naar het Gezondheidspark. Wat doet u:

- a) ik ren naar de bus
- b) ik loop naar de bus
- c) ik ga eerst koffie drinken."

#### 3.2 Effecten van volgorde

Het gedrag van mensen verandert onder andere door ervaringen die zij opdoen. Als je wilt weten welk ontwerp het beste is dan moet je dus elk ontwerp even dikwijls als eerste aanbieden. Bij onderzoek komt daar bij dat de proefpersonen bij de tweede vraag meer op hun gemak zullen zijn en de procedure beter kennen. De effecten van volgorde moet je dus gelijk verdelen over beide condities.

Wanneer het gaat om onderzoek naar reisinformatie dan komt daar nog bij dat de eerste aanbieding de meest reële is. Reizigers gaan niet met twee bussen gelijktijdig op stap; ze kiezen maar één keer. Na de eerste vraag is je proefpersoon eigenlijk al niet meer representatief. De onderzoekers van Keypoint hebben de reizigers eerst vier dienstregelingsommetjes laten maken waarin gevraagd werd naar de *vertrektijd* en pas daarna maar één aftelsommetje. Vind je het gek dat reizigers vertrektijd dan belangrijk vinden?

Beheersing van volgorde is bovendien nodig om effecten van volgorde uit te sluiten. Nieuwe interfaces krijgen meestal niet het voordeel van de twijfel. Men sabelt ze liever neer met: "Ja maar de mensen zijn het oude gewend en nu moeten ze wat nieuws leren." Wanneer men de volgorde beheerst dan krijgen de proefpersonen de nieuwe interface even dikwijls als eerste, als de oude interface. Dikwijls blijkt dan dat de prestaties van gebruikers *onbekend met de nieuwe* interface beter zijn dan de prestaties van gebruikers *bekend met de oude* interface. Ondanks dat de oude interface dikwijls een extra niet beheersbaar voordeel heeft dat de proefpersonen de oude interface kennen. De meeste reizigers kennen vertrektijd omdat NS en luchthavens dit al decennia lang tonen. Als je het effect van volgorde controleert dan kun je de behoudenden neersabelen met de conclusie dat onervaren gebruikers met de nieuwe interface beter scoren dan ervaren gebruikers met de oude interface.

## 4. Inzicht met enquête

Onderzoek doen is erg lastig, dat blijkt. Bovendien overtuigt onderzoek vaak niet. Onderzoek gebruik je om psychologen bezig te houden zodat zij geen lastige vragen stellen. Psychologen trappen daar gemakkelijk in. Zij onderzoeken graag, daar zijn ze voor opgeleid. Zij weten bijvoorbeeld goed hoe je moet enquêteren.

### 4.1 De vormgeving van de vraagstelling

Als je de presentatie van informatie onderzoekt dan moet je de presentatie van je onderzoeksmateriaal zelf niet vergeten. De vormgeving van de vraagstelling heeft effect op het antwoord. Een mooie vriendelijke juffrouw krijgt andere resultaten dan een norse lelijke heer.

In Nederland scoren we meestal van links/slecht naar rechts/goed (bijvoorbeeld schoolcijfers). Het omdraaien van deze richting vergroot de kans op invulfouten. Keypoint draait die volgorde om zoals figuur 4 laat zien. Boven de vakjes staan verder de cijfers 1 tm 4. De invuller moet dus onthouden dat links belangrijk is en rechts onbelangrijk. Belasting van het kortetermijngeheugen vergroot de kans op invulfouten. Je moet belangrijk/onbelangrijk er gewoon bijzetten.

### 4.2 Niet in de mond leggen

Waarschijnlijk staat het in geen enkel methodologieboek, maar toch is het zo: "Je moet ondervraagden geen antwoorden in de mond leggen." Figuur 3 geeft een voorbeeld.

*Figuur 3. De 'Hoe vindt u vraag'*

#### 13. Hoe vindt u de manier waarop deze businformatie wordt weergegeven?

- gemakkelijk, ik hoef zelf geen denkstappen meer te maken
- lastig, want de officiële dienstregeling van de bus staat er nu niet op
- geen mening

Wie wil er nu liever niet te veel denkstappen maken als hij alleen maar naar huis wil? Alleen het eerste antwoord kan juist zijn. Toch blijkt 42% liever eerst wat onnodige denkstappen te maken voor zij in de bus stappen. Wat is hier aan de hand? Het onderzoek was erg gericht op de 'officiële dienstregeling'. De ondervraagden moesten zeven vertrektijden berekenen. Ook het taalgebruik in de enquête is erg technisch ('officiële dienstregeling' in plaats van 'geplande vertrektijd', 'businformatie' in plaats van 'vertrekmoment', 'elektronisch bushaldebord' in plaats van 'bord', 'werkelijke vertrektijd' in plaats van 'vertraagde vertrektijd'). De onderzoekers hebben de ondervraagden opgeleid tot busdienstregelaars en dan is de 'officiële dienstregeling' natuurlijk belangrijk.

Door maar drie alternatieven te geven is de vraag gesloten. Ondervraagden mogen geen argumenten hebben die de onderzoekers zelf niet van te voren konden bedenken. Dat is erg riskant, zeker als je daar zelf een expliciete theorie over hebt. Dat is ook erg riskant als je onderzoek doet naar denkstappen van mensen. Van alle beesten is de mens het best in het verzinnen van een grote variëteit aan slechte en goede oplossingen. Mijn ervaring is dat je meer dan tweemaal zoveel oplossingen en overwegingen krijgt dan je als onderzoeker vooraf bedacht hebt. Als onderzoeker kan ik er negen noemen die de onderzoekers niet voorlegden (Verhoef, L.W.M. (2008). Vertrektijd is passé, leve de afteltijd, OV-magazine, 03-jul, pag. 26-27) Het is niet zo handig een gesloten vraag te stellen. De eenvoudige verwerking van een gesloten multiple choice vraag is hier kennelijk belangrijker geweest dan een juiste conclusie.

### 4.3 Rangordenen

Om te weten te komen wat mensen belangrijk vinden kun je ze rijtjes laten maken. Deze methode is ook gebruikt door Keypoint. Zie figuur 4.

*Figuur 3. Wat vindt u belangrijk*

**19. Kunt u, na alle voorgaande vragen, aangeven wat u graag op een elektronisch bushaldebord zou willen zien?**

**Met de cijfers 1 tm 4 kunt u aangeven hoe belangrijk u de informatie vindt: 1 staat voor heel erg belangrijk en het loopt af tot 4, wat staat voor heel onbelangrijk.**

	1	2	3	4
- Lijnnummer	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- Eindbestemming van de bus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- Huidige tijd	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- Officiële vertrektijd van de bus (volgens de dienstregeling)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- Werkelijke vertrektijd van de bus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- Aantal minuten tot vertrek	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- Indicatie dat er geen recente informatie beschikbaar is over wanneer de bus komt	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- Opmerkingen in geval van calamiteiten	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

De resultaten zijn:

Belangrijk, score: 1,1: Werkelijke vertrektijd van de bus

Minder belangrijk, score: 1,4: Officiële vertrektijd van de bus

Nog minder belangrijk score, 2,5: Aantal minuten tot vertrek

De conclusie die Keypoint uit deze resultaten trekt is onderzoekstechnisch gezien uitermate interessant: "Opvallend hierbij was dat de 'werkelijke vertrektijd van de bus' vele malen hoger scoorde dan het 'aantal minuten tot vertrek'." Deze conclusie is dus in strijd met alle theoretische overwegingen. Hoe kan dat?

- Ten eerste scoorde vertrektijd niet 'vele malen' hoger maar 2,3 maal hoger. Deze slip of the tongue onthult vermoedelijk de impliciete theorie van de onderzoekers: vertrektijd is beter.
- Ten tweede scoort vertrektijd niet 'maal hoger' maar alleen maar 'hoger'. Er is namelijk sprake van een ordinale schaal en niet van een ratio-schaal.
- Ten derde hebben we hier ook weer een volgorde effect. Natuurlijk is lijnnummer belangrijk (in het oude systeem waarin de ondervraagden getraind zijn), natuurlijk is eindbestemming belangrijk (in het oude systeem), etc. Als je al vijf maal gezegd hebt dat iets heel belangrijk is dan denk je: "Alles zo belangrijk, dat kan niet" en je gaat wat lager scoren.
- De onderzoekers concluderen zelf dat 12% aftellen niet begrijpt. De onderzoekers laten in hun vraag de respondenten het belang van informatie rangordenen. Dat is een vorm van aftellen. Aftellen van informatiebelang is veel lastiger dan aftellen van tijd. Volgens de onderzoekers zelf kan minstens 12% van de ondervraagden de vraag die zij stelden dus niet begrijpen.

Methodologisch gezien kloppen de vraagstelling en conclusie dus niet. Ook inhoudelijk zijn er vraagtekens bij te plaatsen. Het belang van lijnnummer en eindbestemming is niet afhankelijk van de mening van de reizigers maar van het systeem dat je kiest. Busmaatschappijen hebben zeer lang geleden gekozen voor lijnnummering. NS heeft een uitstekend systeem van eindbestemmingen en gebruikt geen lijnnummers. Moet NS volgens Keypoint nu lijnnummers gaan invoeren? Het systeem van lijnnummers staat onder druk en meer bestemmingsgeoriënteerde systemen krijgen steeds meer de voorkeur. Moet deze goede trend nu de kop ingedrukt worden?

Dezelfde conclusies kan men trekken uit de andere voorkeuren van reizigers. De onderzoekers melden het rangnummer 6 voor verstoringinformatie. Stel deze vraag maar eens tijdens een verstoring. Met rangnummers dwingt men reizigers tot 'of-denken' terwijl er een genuanceerder 'en-denken' geldt.

#### **4.4 Gebruikers en ontwerpen**

De mening van mensen is van belang wanneer het gaat om opinies over fundamentele levensvragen. Die mogen anderen niet voor je kiezen. Maar een dokter die vraagt: "Welke ziekte had u zelf in gedachten" heeft veel tevreden en veel overleden patiënten.

Vaak verdedigt men de enquête-aanpak met: "Ja maar dit was alleen maar om te weten wat reizigers ervan vinden." Die verdediging gaat hier niet op: Het doel van het onderzoek is te weten te komen welke presentatie van businformatie de reiziger het best begrijpt. Heb je die gevonden dan is de vraag voor de gebruiker onbegrijpelijk. De vraag: "En hoe fietst het?" heeft geen zin als de fiets maar één trapper, geen zadel en geen rem heeft. Maar de vraag heeft ook geen zin wanneer de fiets goed is. Mijn ervaring is dat je dan als antwoord krijgt: "Man zeur niet, wat is het probleem, laat me toch lekker fietsen."

### **5. Inzicht met statistiek**

Proefpersonen selecteer je niet zomaar. Wanneer je wilt voorspellen hoe Nederland kiest dan moet je een representatieve steekproef nemen. Je moet niet alleen jonge mannelijke Limburgers vragen wat zij gaan stemmen.

De onderzoeker van Keypoint hebben Brabanders en Groningers onderzocht. Daarmee is het onderzoek volgens de onderzoekers geografisch representatief. Het blijkt dat geografie niets uitmaakt, concluderen de onderzoekers. Hadden de onderzoekers echt verwacht dat Brabanders anders klokkijken en aftellen dan Groningers? Als men zo graag 'geografisch representatief' wilde zijn waarom zijn de onderzoekers niet in Londen gaan kijken welke busreizigers nauwkeuriger op het balkon van een al vertrekkende bus springen: a) busreizigers die rennend dienstregelingsommetjes maken, of b) busreizigers die zien dat zij in tien seconden vijfentwintig meter moeten afleggen. Ook Houston is interessant. Daar telt men af terwijl je er nooit reizigers in een reeds vertrekkende raket ziet springen.

Een representatieve steekproef kies je wanneer je uitspraken wilt doen over een gehele populatie. Elke subgroep moet een evenredig deel van de proefpersonen leveren. Bij designergericht onderzoek is dat niet handig. Je neemt juist een selecte steekproef om fouten met het ontwerp beter waar te kunnen nemen. Je brengt je proefpersonen bijvoorbeeld onder druk zodat zij slechter presteren of je neemt meer ouderen. Keypoint wil ook naar selecte steekproeven toe door via internet te gaan enquêteren. Dat is natuurlijk wel erg gemakkelijk. Maar je krijgt wel een reisinformatiesysteem dat bedacht is door Cor van der Laak-achtige (Cor van der Laak was een typetje van cabaretier Kees van Kooten. Van der Laak had overal uitgesproken meningen over) whizzkids die niets anders te doen hebben dan op 'mijn mening' knoppen te drukken.

Misschien toch maar een psycholoog inhuren. Al was het alleen om te horen dat je met wat statistiek niet 1121 Cor van der Laaks hoeft aan te horen maar dat zo'n 200 al aan de ruime kant is om betrouwbare uitspraken te kunnen doen.

### **6. Inzicht met psychologie**

Het inhuren van een psycholoog heeft wel zijn prijs. Een praktijk die geen rekening houdt met de functiepsychologie - hoe mensen bewegen, waarnemen, taal gebruiken, leren en denken - kan rekenen op een harde confrontatie. Dat blijkt uit dit artikel maar ook uit het werk van functiepsychologen als Piet Vroon, Willem-Albert Wagenaar en Robert Cialdini. Gelukkig is de spiegel die de functieleer ons voorhoudt vaak een lachspiegel en gaat de les verloren in de lach. Als we uitgelachen zijn is de les vergeten en kunnen we gewoon doorgaan met onze eigen niet psychologisch verantwoorde brouwsels.

*Leonard Verhoef, psycholoog*