Dataverzameling via internet: een update

Toon Taris*, Paul Schreurs** & Kees Jan Sepmeijer**

De snelle verbreiding van het gebruik van het internet heeft ervoor gezorgd dat het verzamelen van gegevens via dat medium sterk in de belangstelling van onderzoekers is komen te staan. In deze studie bespreken we de voor- en nadelen van deze manier om gegevens te verzamelen. Voordelen betreffen de snelheid en efficiency, alsmede de mogelijkheid om interviewer- en invoerfouten te elimineren. Nadelen zijn de soms lage response rates, mogelijke selectieve respons en mogelijke problemen betreffende de kwaliteit van de verkregen gegevens. Bij wijze van illustratie werden de resultaten van een studie onder ruim 45.000 medewerkers van Nederlandse thuiszorginstellingen gepresenteerd. Daaruit bleek dat hoogopgeleiden, mannen en jongeren relatief vaak via internet responderden. Confirmatieve factoranalyse liet zien dat de factorstructuur van de onderzochte werkgerelateerde concepten niet verschilde voor internet- en niet-internetgebruikers, maar dat de internetrespondenten gemiddeld gesproken duidelijk ongunstiger dachten over hun werkomstandigheden dan anderen. Ten slotte bespreken we de implicaties van deze bevindingen voor het onderzoek in de sociale, arbeids- en organisatiepsychologie.

Trefwoorden: dataverzameling, internet, respons, validiteit, representativiteit

1 Inleiding

In zijn *presidential address* aan de leden van de *American Association for Public Opinion Research* constateerde Don A. Dillman onlangs dat het aantal manieren waarop onderzoekers hun gegevens kunnen verzamelen de afgelopen veertig jaar sterk is toegenomen (Dillman, 2002). Waar enquêteurs omstreeks 1960 nog op pad gingen met vragenlijsten, interviewformulieren en kaarten met antwoordcategorieën dan wel vragenlijsten per post verzonden werden, kwam een decennium later het telefonisch enquêteren in zwang. Hoewel aanvankelijk betwijfeld werd of zo'n ongebruikelijke manier van data verzamelen kwalitatief even goede gegevens opleverde als het face-to-face interview, concludeerde Seymour Sudman – een andere autoriteit op het gebied van dataverzamelingstechnieken – in 1974 dat *'it was apparent that we could now conduct [...] surveys of 30 minutes or even longer over the telephone'* (Dillman, 2002, p. 475). Telefonisch enquêteren was *salonfähig* gewor-

* Radboud Universiteit Nijmegen, Sectie Arbeids- en Organisationspsychologie. **Instituut Werk en Stress, Bilthoven. Correspondentieadres: Dr. T.W. Taris, Radboud Universiteit Nijmegen, Sectie Arbeids- en Organisationspsychologie, Postbus 9104, 6500 HE Nijmegen. Telefoon 024-361 2639, fax 024-361 5937, e-mail t.taris@psych.ru.nl.

den, en sommigen voorspelden dat het face-to-face interview volledig zou worden verdrongen door de telefonische enquête.

Sindsdien heeft de techniek niet stilgestaan. Eén van de meest opvallende methodische vernieuwingen van het afgelopen decennium was wel de opkomst van het verzamelen van gegevens via inter- en intranet. Dataverzameling via deze media heeft grote aantrekkingskracht, omdat zo snel en goedkoop gegevens verzameld kunnen worden bij grote aantallen respondenten (Stanton & Rogelberg, 2003). In de literatuur is echter relatief weinig bekend over de kwaliteit van op dergelijke manier verkregen gegevens, bijvoorbeeld in termen van representativiteit van de steekproef, responspercentages en de equivalentie van via internet en pen-en-papier verkregen antwoorden (Dillman & Bowker, 2001; Thompson, Surface, Martin & Sanders, 2003). Kortom, dataverzameling via internet lijkt een op het eerste gezicht aantrekkelijke mogelijkheid om snel en zonder hoge kosten grote hoeveelheden gegevens te verzamelen, maar na enig nadenken zullen potentiële gebruikers echter constateren dat het gebruik van dit nieuwe instrument omgeven is met onzekerheden, mogelijke problemen en onbeantwoorde vragen.

In het navolgende willen we de lezers meer inzicht bieden in de kansen die dataverzameling via internet biedt, alsmede de valkuilen die daarbij van belang kunnen zijn. We gaan daarbij uit van de situatie waarin onderzoekers een min of meer afgebakende populatie willen benaderen (bijvoorbeeld de medewerkers van een bepaalde arbeidsorganisatie) en willen voldoen aan de in wetenschappelijk onderzoek gebruikelijke eisen van validiteit, representativiteit en betrouwbaarheid van de gegevens. In het eerste deel van deze bijdrage worden enkele aspecten van het verzamelen van gegevens via internet besproken. Daarbij gaat het om twee vragen: (1) wie doen er mee (representativiteit van het onderzoek)? en (2) wat is de kwaliteit van de verkregen gegevens? In het tweede deel presenteren we bij wijze van illustratie de resultaten van een onderzoek naar de vergelijkbaarheid van de antwoorden van thuiszorgmedewerkers die participeerden in een studie naar de kwaliteit van hun werkomstandigheden.

2 Dataverzameling via internet: twee vragen die gebruikers zich vooraf zouden moeten stellen

Dataverzameling via internet biedt veel voordelen. Hiervoor noemden we al dat deze manier om gegevens te verzamelen *relatief snel* leidt tot resultaten, vergeleken met meer traditionele instrumenten als de schriftelijke enquête en het face-to-face interview. In zijn eenvoudigste vorm houdt een internet-enquête niet meer in dan dat een mailtje naar de beoogde respondenten wordt gestuurd, met daarin het verzoek om in te loggen op het internetadres waar de vragenlijst te vinden is. Ten opzichte van de schriftelijke enquête hoeven er dus geen vragenlijsten meer gedrukt te worden, in enveloppen te worden gedaan of op de post gedaan te worden; er hoeven geen gegevens meer ingevoerd te worden; en er kan begonnen worden met de analyse van gegevens op het moment dat de eerste respondenten hun antwoorden hebben gegeven (wat al binnen enkele minuten het geval kan zijn). De tijdwinst ten opzichte van face-to-face interviews is nog groter, omdat het veldwerk dan snel enkele weken of zelfs maanden vergt.

Het ontbreken van de veldwerk- en invoerfase heeft ook gevolgen voor de *kosten* van het onderzoek. De aanloopkosten van een internet-enquête zijn relatief hoog. De vragenlijst moet immers geprogrammeerd en getest worden,

en daarvoor is vaak hulp van specialisten nodig. In het verdere verloop van het onderzoek worden die kosten echter gemakkelijk terugverdiend, vooral als het onderzoek onder een grote groep deelnemers wordt afgenomen: de marginale kosten van een extra respondent zijn laag. Ook als dezelfde vragenlijst vaker wordt afgenomen (bijvoorbeeld bij onderzoek naar de tevredenheid van klanten of medewerkers) kan de investering in een internetversie van de desbetreffende vragenlijst zeer de moeite waard zijn.

Een en ander gaat ook gepaard met *methodische voordelen*. Interviewers in face-to-face studies en telefonische enquêtes interpreteren de antwoorden van de respondenten bijvoorbeeld soms op eigen wijze, stellen eigen vragen, suggereren antwoorden, proberen hun eigen inzichten te toetsen, of worden door de respondent on aardig, aardig of zelfs seksueel aantrekkelijk gevonden (Dijkstra, 1983, voor een amusante staalkaart). Al dit soort factoren kan de kwaliteit van de verkregen gegevens negatief beïnvloeden; omdat er in een internetstudie geen interviewer nodig is om de gegevens te verzamelen, kan interviewerbias hier worden uitgeschakeld. Een traditionele schriftelijke enquête biedt in deze opzichten overigens dezelfde voordelen ten opzichte van face-to-face en telefonisch onderzoek.

Er zijn dus goede redenen om dataverzameling via internet te prefereren boven traditionele methoden. Dat betekent echter niet dat web-based enquêtes onder alle omstandigheden te verkiezen is boven meer traditionele benaderingen. Aangenomen mag worden dat onderzoekers primair *betrouwbare, valide en generaliseerbare kennis* willen genereren. Er is echter niet noodzakelijkerwijze een positieve relatie tussen enerzijds de snelheid waarmee een bepaalde benadering resultaten oplevert of de kosten daarvan, en anderzijds de kwaliteit van de daaruit resulterende gegevens. Een verstandige keuze voor een bepaalde wijze van dataverzameling betekent dat onderzoekers inzicht moeten hebben in de mogelijke problemen die daarbij een rol kunnen spelen. Met betrekking tot dataverzameling via internet spelen er twee belangrijke problemen, betreffende (1) de representativiteit van de onderzoeksgroep, en (2) de kwaliteit van de verkregen gegevens. Hierna worden deze problemen verder uitgediept.

2.1 Wie doen er mee?

In het ideale survey-onderzoek heeft ieder lid van de beoogde onderzoekspopulatie dezelfde kans om in de steekproef te worden opgenomen, de personen in de steekproef zijn random geselecteerd en geven tezamen een goede indruk van de beoogde populatie, en iedereen in de steekproef werkt mee aan het onderzoek (Dillman, 1991). In de praktijk wordt maar zelden aan deze voorwaarden voldaan. De respons op internetsurveys is lang niet altijd hoog genoeg om generaliseerbaarheid naar enige populatie te kunnen garanderen. Daarnaast maken factoren als non-random steekproeftrekking, problemen met internet junkmail (*spam* – wordt de vragenlijst door de geadresseerden herkend als een serieus onderzoek, of opgevat als ongewenste post?) en beperkte opslaglimieten (de mailbox van de respondent kan vol zijn, zodat post niet ontvangen kan worden) dataverzameling via internet tot een soms problematische aangelegenheid, zelfs als de beoogde populatie in principe toegang heeft tot het internet (Sills & Song, 2002).

Dekkingsfouten. Wanneer niet alle leden van de beoogde onderzoekspopulatie in potentie in de steekproef kunnen worden opgenomen, is sprake van een dekkingsfout. Het aantal mensen dat toegang heeft tot het internet groeit bijvoorbeeld snel, waardoor personen die tot de onderzoekspopulatie behoren maar die nog maar kort over een internetadres beschikken, vaak niet benaderd

zullen worden voor een onderzoek. Daarnaast verouderen internetadressen zélf zeer snel, wat betekent dat lang niet alle verstuurde boodschappen de geadresseerden zullen bereiken. Soms wordt niet minder 30% van de verstuurde vragen om medewerking aan een onderzoek teruggekaatst omdat de desbetreffende adressen niet meer actueel zijn (Swoboda, Mühlberger, Weitkunat & Schneeweiss, 1997).

Fouten bij de steekproeftrekking. Van een fout bij de steekproeftrekking ('sampling error') is sprake als de steekproef bestaat uit een subset van een bepaalde heterogene onderzoekspopulatie. Vanwege het non-random en niet-representatieve karakter van zo'n steekproef kunnen de resultaten van een onderzoek in deze steekproef niet gegeneraliseerd worden naar de gehele onderzoekspopulatie. De eerste auteur van dit artikel was enige jaren geleden bijvoorbeeld aanwezig bij de presentatie van de resultaten van een onderzoek naar relaties en seksualiteit. De oudere deelnemers in het onderzoek (80 jaar en ouder) bleken seksueel nog opvallend actief – een bevinding die een deel van de aanwezigen verwarde en een ander deel hoop voor de toekomst gaf, totdat bleek dat de desbetreffende gegevens via internet waren verzameld. Het lijkt aannemelijk dat een hoogbejaarde die in staat is een destijds nieuw medium als internet te gebruiken ook in andere opzichten afwijkt van de doorsnee *senior citizen*: het gebruik van internet heeft in dit geval wellicht geleid tot een fout bij de steekproeftrekking.

Non-respons. Non-respons heeft betrekking op het verschil tussen de verkregen en de oorspronkelijk beoogde waarnemingen (in casu de getrokken steekproef, bestaande uit respondenten en non-respondenten). Onderzoekers hopen gewoonlijk dat de non-respons random is, dat wil zeggen dat respondenten niet systematisch verschillen van diegenen die niet respondeerden. Echter, wanneer de respons minder dan 100% bedraagt, bestaat de mogelijkheid van *selectieve vertekening*: non-respondenten verschillen dan op bepaalde aspecten wél van respondenten. Eén vuistregel is dat de kans op grote vertekening als gevolg van selectieve (of non-random) non-respons afneemt naarmate de respons hoger is (Taris, 2000, voor kanttekeningen).

Alweer enige tijd geleden lieten Hox en De Leeuw (1994) zien dat methoden om gegevens te verzamelen verschillen in de respons die ze genereren. Op basis van een meta-analyse van 45 studies vonden zij een gemiddeld responspercentage van 70.3% voor face-to-face onderzoeken, 67.2% voor telefonische enquêtes, en 61.3% voor postenquêtes. Gewoonlijk is de respons op een internet-enquête (die vaak varieert van 20 tot 30%) dus duidelijk lager dan die in studies waarin de gegevens op andere wijze worden verkregen (Sills & Song, 2002, voor een overzicht).¹ Dillman, Phelps et al. (2001) noteerden in hun studie naar de respons op verschillende dataverzamelmethodeën bijvoorbeeld een respons van 13% voor diegenen die via internet moesten responderen, tegen 44% en 75% voor diegenen die respectievelijk via de telefoon dan wel via een pen-en-papier vragenlijst moesten reageren. Dergelijke verschillen in respons kunnen ertoe leiden dat de demografische kenmerken van internet-respondenten afwijken van die van anderen. Zo blijken personen die via internet responderen vaak jonger en hoger opgeleid dan anderen, en vaker van het mannelijk geslacht te zijn (Bandilla, Bosnjak & Altdorfer, 2003). Dat kan erop wijzen dat er sprake is van selectieve respons, waardoor de via internet verkregen resultaten mogelijk niet gegeneraliseerd kunnen worden naar de bedoelde populatie.

Relevantie voor de sociale en arbeids- en organisatiepsychologie. In hoeverre zijn bovenstaande overwegingen relevant voor het onderzoek zoals dat in de sociale, arbeids- en organisatiepsychologie wordt uitgevoerd? Veel van derge-

lijk onderzoek vindt plaats in de overzichtelijke context van een of meer organisaties. Verwacht mag worden dat deze een goed inzicht zullen hebben in de mate waarin hun personeel toegang heeft tot het inter- of intranet. Ook is aannemelijk dat zij beschikken over een up-to-date bestand van de e-mailadressen van hun werknemers. Een en ander impliceert dat dekkingsfouten en fouten bij de steekproeftrekking hier niet voor hoeven te komen. Problemen ten aanzien van selectieve non-respons kunnen worden opgespoord door organisatiegegevens te vergelijken met daarmee corresponderende gegevens van de steekproef. In potentie kan non-respons zelfs vrijwel volledig worden voorkomen, bijvoorbeeld door werknemers die nog niet hebben gerespondeerd eraan te herinneren dat zij nog een vragenlijst moeten invullen. Daarbij komen overigens wel problemen met de anonimiteit en vertrouwelijkheid van het onderzoek om de hoek kijken (Thompson et al., 2003) – Big Brother laat dan immers blijken het doen en laten van zijn medewerkers in de gaten te houden, wat door non-respondenten als zeer onaangenaam kan worden ervaren.

De internetsurvey kan binnen organisaties dus een handige manier zijn om gegevens te verzamelen, vooral wanneer alle medewerkers toegang hebben tot het inter- of intranet. Wanneer onderzoekers echter leden van een minder precies gedefinieerde populatie willen bereiken, zijn de hiervoor genoemde problemen levensgroot aanwezig. Respondenten kunnen bijvoorbeeld geworven worden via nieuwsgroepen, advertenties op aan een bepaald verschijnsel gewijde internetpagina's of door publiciteit in de media te zoeken. Een en ander leidt vaak tot gelegenheidssteekproeven die allesbehalve representatief zijn voor de beoogde populatie. Op dit soort steekproeven gebaseerde beschrijvende gegevens kunnen uiteraard niet gegeneraliseerd worden. Dat betekent echter niet dat het a priori zinloos is om dergelijke gegevens te verzamelen. Psychologen zijn vaak eerder geïnteresseerd in *verbanden* tussen variabelen dan in de gemiddelde scores van hun steekproef op die variabelen. Wanneer aangenomen kan worden dat bepaalde psychologische processen universeel geldig zijn, kan zelfs een slecht gedefinieerde en niet-representatieve gelegenheidssteekproef het inzicht in een bepaald verschijnsel vergroten.

2.2 Kwaliteit van via internet verzamelde gegevens

Wellicht de belangrijkste vraag die onderzoekers zich zullen stellen wanneer zij overwegen om hun data via internet te verzamelen is, of de aldus verkregen gegevens (en de daarop gebaseerde resultaten) hetzelfde zijn als wanneer zij gebruik zouden maken van pen en papier. Het antwoord op die vraag is: *vermoedelijk niet*. Hiervoor zagen we immers al dat er vaak demografische verschillen zijn tussen de mensen die via internet antwoorden en diegenen die een andere manier van dataverzameling prefereren; het lijkt aannemelijk dat zulks zal doorwerken in de antwoorden van beide groepen. Maar stel dat we de effecten van dit soort verschillen zouden wegfilteren (bijvoorbeeld door er statistisch voor te controleren, dan wel respondenten te matchen op bepaalde kenmerken), zijn er dan nog verschillen tussen de antwoorden van mensen die via internet, dan wel via een andere methode van dataverzameling worden benaderd? Dillman (2002) vraagt zich in deze context bijvoorbeeld af of respondenten het beantwoorden van internetvragenlijsten zien als het spelen van een computerspelletje:

'Repeatedly, I have observed cognitive interview respondents read only part of each question before going to response categories [...] The sense of not wanting to mark a wrong answer because of the difficulty of the correction process that I have often observed in interviews of paper questionnaire respondents, seems not to be part of the response process [...] I wonder if there is a game

mentality guiding responses to web surveys for figuring out how to answer quickly [...] Does this produce as a side effect an increase in the proportion of meaningless or inaccurate responses? [...] Is the sense of obligation for providing honest answers declining, while responding to surveys is becoming more of a game than a case of serious reflection and response?' (p. 489)

Een ander bezwaar van het gebruik van het internet betreft mogelijke problemen met de anonimiteit en de vertrouwelijkheid van de gegeven antwoorden. Om te voorkomen dat personen die niet tot de steekproef behoren de internetvragenlijst invullen (of dat personen die wél in de steekproef vallen de vragenlijst meer dan eens invullen) moeten participanten vaak een code intypen om toegang tot de vragenlijst te krijgen. Dit kan ervoor zorgen dat respondenten twifelen aan de mate waarin hun antwoorden vertrouwelijk worden behandeld. Thompson et al. (2003) rapporteren bijvoorbeeld dat 10% van hun respondenten zich zorgen maakt over deze aspecten van internet-enquêteren; dergelijke respondenten zouden gemotiveerd kunnen zijn om neutrale en sociaal wenselijke antwoorden te geven. Anderzijds kan beargumenteerd worden dat het ontbreken van een interviewer er bij een internet-enquête toe kan bijdragen dat respondenten juist *minder* sociaal wenselijk antwoorden dan anders het geval zou zijn.

Jammer genoeg is er weinig onderzoek naar deze kwesties gedaan. Wel kan hier deels worden teruggevallen op onderzoek uit het begin van de jaren '90 naar de vergelijkbaarheid van antwoorden op pen-en-papier tests en de gecomputeriseerde equivalenten daarvan. Met de opkomst van de microcomputer werden destijds veel traditionele psychologische tests opgenomen in allerlei gecomputeriseerde *test managers*; destijds maakten velen zich zorgen over de vergelijkbaarheid van de resultaten van beide manieren van afnemen. Psychometrisch onderzoek (dat vaak gebruikmaakte van experimentele proefopzetten, en waarin geavanceerde methoden als confirmatieve factoranalyse en item-responsmodellen werden toegepast) liet echter zien dat verschillen tussen de resultaten van beide manieren van afname vaak te verwaarlozen waren; factorstructuren waren bijvoorbeeld vrijwel identiek (Donovan, Drasgow & Probst, 2000; King & Miles, 1995, voor exemplarische studies). Dergelijk onderzoek naar de equivalentie van meetinstrumenten in de context van internet- versus pen-en-papier vragenlijsten ontbreekt echter tot nu toe. Enkele quasi-experimentele studies vergelijken de antwoorden van groepen respondenten die hetzij via internet, hetzij via pen-en-papier antwoorden gaven. Eén studie laat zien dat in 27 tot 73% van de gevallen (items) significante verschillen in gemiddelden tussen beide groepen worden geconstateerd (Bandilla et al., 2003; zie ook Dillman et al., 2001). Hoe deze bevinding geïnterpreteerd moet worden, is echter onduidelijk.

3 Illustratie: Een casestudy in de thuiszorg

De hiervoor beschreven overwegingen en resultaten suggereren dat het kan uitmaken of respondenten via internet of anderszins worden benaderd. Personen die via internet worden benaderd, blijken wat betreft hun demografische achtergrond en hun antwoorden te verschillen van respondenten die op andere wijze antwoorden; de via internet behaalde responspercentages zijn relatief laag; en er zijn redenen om aan de kwaliteit van de via internet verkregen gegevens te twifelen. Maar in hoeverre spelen deze problemen ook bij onderzoek in Nederlandse arbeidsorganisaties? Hierna presenteren we, bij wijze van illustratie van deze mogelijke problemen, de resultaten van een

onderzoek naar (a) verschillen tussen personen die via internet dan wel via pen-en-papier respondeerden; en (b) verschillen in de antwoorden die deze twee groepen geven. Het onderzoek werd uitgevoerd in de context van een studie naar de werkbeleving van Nederlandse thuiszorgmedewerkers, die konden kiezen of ze een vragenlijst via internet dan wel op de gebruikelijke wijze via pen-en-papier invulden.

3.1 Verschillen personen die via internet responderen van andere respondenten?

Mensen kunnen verschillen in de mate waarin zij toegang hebben tot het internet, alsook in de mate waarin zij geneigd zijn gebruik te maken van dit relatief nieuwe medium. Dat kan betekenen dat personen die ervoor kiezen een vragenlijst via het internet te beantwoorden, verschillen van diegenen die een meer traditionele aanpak prefereren. In het schaarse eerdere onderzoek naar het verband tussen kenmerken van respondenten en de voorkeur voor een bepaalde manier om te antwoorden werd geconstateerd dat mannen, welgestelden en jongeren vaak oververtegenwoordigd zijn in internetsurveys (Bandilla et al., 2003; Sills & Song, 2002), wat deels te wijten kan zijn aan het feit dat deze groepen vaker zouden kunnen beschikken over computers met internetaansluiting dan vrouwen, minder welgestelden en ouderen. Anderzijds bleek uit de studie van Thompson et al. (2003) dat mannelijke en vrouwelijke werknemers niet verschilden wat betreft hun voorkeur voor een bepaald medium. Het bewijs voor de stelling dat het internet vooral geschikt is om jonge, hoogopgeleide mannen te ondervragen is dus niet eenduidig – reden om in de huidige studie na te gaan of de kenmerken van diegenen die via internet respondeerden, verschilden van die van respondenten die pen, papier en post gebruikten.

Methode

Respondenten. In het kader van de arboconvenantenaanpak van werkdruk werd in 2002 een sectorbreed onderzoek gedaan naar de werkbeleving van werknemers in de thuiszorg (Taris et al., 2004). Alle werknemers van alle thuiszorginstellingen ontvingen op hun huisadres een gestructureerde vragenlijst waarin gevraagd werd naar hun werkbeleving (werkdruk, mogelijkheden en dergelijke), alsmede enkele uitkomstvariabelen (lichamelijke en psychische klachten, betrokkenheid en dergelijke). De begeleidende brief gaf aan dat zij hun antwoorden ook via internet konden geven; daartoe noemde de brief het internetadres waar de betreffende vragenlijst te vinden was. De respondenten bepaalden dus zelf hoe zij aan het onderzoek meededen. De respons bedroeg 43.2% ($N = 45.106$); slechts 1.3% responderde via het internet; 3.6% was man; gemiddelde leeftijd was 42.5 jaar, $SD = 10.2$; 64.5% van de respondenten had een opleiding afgerond op LBO-niveau of lager, 8.5% had een opleiding afgerond op MBO-niveau, en de resterende 27.8% had een opleiding op ten minste HBO-niveau.

Variabelen/statistische procedure. In deze eerste studie werden de achtergrondvariabelen sekse, leeftijd (5 categorieën), opleiding (6 categorieën, variërend van uitsluitend lagere school tot universitair niveau), etniciteit (autochtoon versus allochtoon) en functiegroep (9 categorieën) betrokken. Via logistische regressieanalyse werd nagegaan of achtergrondvariabelen voorspelden welk medium de deelnemers gebruikten om te responderen.

Resultaten

Tabel 1 laat zien dat de achtergrondvariabelen 7% van de variantie in het al dan niet responderen via internet verklaarden. In overeenstemming met eerder onderzoek vonden we dat mannen en jongeren vaker via internet responderden dan vrouwen en ouderen. De odds ratio voor sekse was 3.01 (mannen hebben drie keer zoveel kans om via internet te responderen dan vrouwen); voor leeftijd vonden we voor de hoogste leeftijdscategorie een odds ratio van .20, wat betekent dat deze leeftijdsgroep vijf keer minder vaak via internet respondeert dan de jongste leeftijdsgroep, in casu de referentiecategorie). Consistent met eerder onderzoek bleken de laagst geschoolde respondenten het minst via internet te responderen, al waren de verschillen met de hogere opleidingscategorieën vaak niet significant. Ten slotte maakten de respondenten in de onderscheiden functiegroepen in verschillende mate gebruik van de mogelijkheid om de vragenlijst via internet in te vullen; vooral personen die werkzaam waren in de algemene ondersteuning (waaronder secretariële en administratieve functies, waarin veel met computers wordt gewerkt) bleken via internet te responderen. Etniciteit was van geen belang.

Tabel 1 Resultaten van een multivariate logistische regressieanalyse op de kans om via internet te responderen (*chi-kwadraat* ($df = 18$, $N = 45.106$) = 391.5, $p < .001$, 98.6% correct geclassificeerd, $R^2 = .071$)

Variabelen	N	Exp(b)	95% betrouwbaarheids-interval
<i>Sekse</i>			
vrouw	39.564	1.00 ^a	
man	1.473	3.01***	2.28 - 3.97
<i>Leeftijd</i>			
< 21 jaar	593	1.00 ^a	
21-30 jaar	5.456	.80	.48 - 1.32
31-40 jaar	10.845	.57***	.34 - .93
41-50 jaar	14.090	.42***	.26 - .69
51 jaar en ouder	10.053	.20***	.12 - .35
<i>Opleiding</i>			
Lagere school	1.954	1.00 ^a	
Lager beroepsonderwijs	22.485	1.51	.81 - 2.80
Middelbaar beroepsonderwijs	3.221	6.80***	3.59 - 12.87
Hoger beroepsonderwijs	9.307	1.49	.78 - 2.83
Universitair onderwijs	932	1.70	.77 - 3.75
<i>Etniciteit</i>			
Autochtoon	39.807	1.00 ^a	
Allochtoon	1.230	1.14	.72 - 1.80
<i>Functiegroep</i>			
Hulp A en B	14.275	1.00 ^a	
Verzorgende C, D, E	8.313	.91	.70 - 1.19
Wijkverpleegkundige	6.780	.71*	.54 - .94
Leiding en staf	2.463	1.30	.90 - 1.87
Kraamzorg	3.361	.90	.63 - 1.28
Ouder- en kindzorg	2.522	.63	.38 - 1.04
Maatschappelijk werk	311	1.67	.76 - 3.66
Algemene ondersteuning	1.586	2.44***	1.74 - 3.42
Overig	1.426	1.54*	1.02 - 2.31

^a = Referentiecategorie.

* = $p < .05$, ** = $p < .01$, *** = $p < .001$.

Conclusie/discussie. In overeenstemming met eerder onderzoek bleek ook in deze studie dat vrouwen, laagopgeleiden en ouderen relatief weinig van internet gebruikmaakten om te responderen. In de huidige studie konden de respondenten zelf kiezen van welk medium zij gebruikmaakten om de vragenlijst in te vullen; diegenen die het internet niet wilden gebruiken konden hun vragenlijst ook per post verzenden. In andere internetstudies is zoiets niet altijd mogelijk, wat kan betekenen dat in dergelijke studies vrouwen, ouderen en laagopgeleiden ondervertegenwoordigd zijn.

Een praktisch punt betreft het zeer lage percentage respondenten dat per internet antwoordde (1.3%). De overgrote meerderheid van de deelnemers gebruikte dus pen en papier. Gezien de kenmerken van de onderzoeksgroep (laagopgeleid, voornamelijk vrouwen, veel 40-plussers, relatief veel functies waarin weinig of niet met de computer wordt gewerkt) kan men zich afvragen of een dergelijk gebrek aan enthousiasme voor het internet te voorspellen was geweest, en of het zinvol was om de deelnemers de mogelijkheid te bieden om via internet te responderen. Omdat in absolute zin toch nog vrij veel mensen via internet respondeerden (586 respondenten) zullen de kosten in dit concrete geval toch nog tegen de baten opwegen. Desalniettemin onderstreept dit resultaat onze stelling dat onderzoekers van tevoren goed moeten bedenken of het zinvol is om hun steekproef via internet te benaderen of dat andere, meer traditionele methoden passender zijn.

3.2 Verschillen de antwoorden van internetrespondenten van die van anderen?

Hiervoor constateerden we dat deelnemers die via internet respondeerden in bepaalde opzichten (seks, leeftijd, opleiding, type functie) afweken van andere deelnemers. Zulke verschillen zijn vooral betekenisvol als er ook verschillen zijn in de *antwoorden* die de leden van verschillende groepen geven. Immers, als beide groepen niet verschillen in hun antwoordgedrag, is er ook geen sprake van vertekening. Hiervoor gaven we al aan dat er enige redenen zijn om verschillen te verwachten tussen de antwoorden van internetrespondenten en andere respondenten. Internetrespondenten zouden minder precieze antwoorden geven (Dillman, 2002), wat zou kunnen blijken uit relatief grote varianties op de gevraagde concepten, alsook uit relatief lage factorladingen van de items die op deze concepten laden. Internetrespondenten zouden daarnaast bevreesd kunnen zijn voor hun privacy, wat ertoe zou moeten leiden dat zij vaker neutrale antwoorden geven dan anderen – er zou dus verschil moeten zijn tussen de gemiddelde scores van de groep internetrespondenten en die van de andere respondenten, terwijl in de eerste groep de varianties van de concepten relatief laag zouden moeten zijn (men zal immers vaak het neutrale, sociaal wenselijke antwoord geven).

Methode

Respondenten. Ten behoeve van dit deelonderzoek werden de gegevens van de internetrespondenten vergeleken met die van de overigen. Hiervóór bleek al dat seks, leeftijd, functiegroep en opleidingsniveau samenhangen met de kans om al dan niet via internet te responderen. Omdat we de invloed van deze factoren wilden minimaliseren, werd de analyse beperkt tot de vrouwelijke werknemers in de twee laagste (omvangrijkste en meest homogene) functiegroepen (Hulp A en B, en Verzorgenden C, D en E). Vervolgens werden twee groepen gecreëerd: zij die per internet respondeerden ($N = 245$) en zij die dat niet deden. Elke respondent in de eerste groep werd op leeftijd, functie en opleidingsniveau gematcht met een respondent in de tweede groep. De uiteindelijke onderzoeksgroep bestond uit 245 internetrespondenten en 245

andere respondenten; beide subgroepen waren identiek in termen van sekse, leeftijd, opleiding en functieniveau. Eventuele verschillen in de antwoorden van beide groepen kunnen dus niet aan deze achtergrondfactoren worden toegeschreven.

Variabelen. De groepen werden vergeleken op vijf vanuit arbeids- en gezondheidspychologisch perspectief interessante concepten. Twee daarvan (werkdruk en regelmogelijkheden) betreffen centrale kenmerken van de werksituatie (Karasek & Theorell, 1990) die in zeer veel onderzoek worden gemeten. Een derde kenmerk van de werksituatie, de mate waarin men wordt geconfronteerd met ongewenste intimiteiten, werd gekozen omdat het een 'moeilijk', privacygevoelig onderwerp betreft waarover men niet gemakkelijk met anderen spreekt; wellicht dat verschillen tussen manieren om gegevens te verzamelen juist op dergelijke concepten zichtbaar worden. De twee andere variabelen betroffen uitputting als indicator van negatief welbevinden en toewijding als indicator van positief welbevinden.

Werkdruk werd gemeten met vier door Karasek (1985) ontwikkelde items, waaronder 'Moet u erg snel werken?' (1 = 'nooit', 4 = 'altijd', $\alpha = .82$). *Regelmogelijkheden* werd gemeten met twee eveneens van Karasek (1985) afkomstige items, namelijk 'Kunt u zelf beslissen hoe u uw werk uitvoert' en 'Heeft u vrijheid bij het uitvoeren van uw werkzaamheden' (1 = 'nooit', 4 = 'altijd', $r = .51$).² De mate waarin respondenten geconfronteerd werden met ongewenste intimiteiten, werd gemeten met drie zelfgeconstrueerde items, respectievelijk 'Hoe vaak heeft u tijdens uw werk te maken met ongewenste intimiteiten door cliënten?', 'Hoe vaak wordt u tijdens uw werk geconfronteerd met grapjes of opmerkingen met een seksuele ondertoon, die u als ongewenst beschouwt?', en 'Hoe vaak wordt u tijdens uw werk aangeraakt of vastgepakt op een manier die u als ongewenst beschouwt?' (1 = 'nooit', 5 = 'altijd', $\alpha = .85$). *Emotionele uitputting* werd gemeten met de desbetreffende schaal uit de Utrechtse Burnout Schaal (Schaufeli & Van Dierendonck, 2000). Deze bestaat uit vijf items, waaronder 'Een hele dag werken vormt een zware belasting voor mij' (0 = 'nooit', 6 = 'altijd', $\alpha = .86$). *Toewijding* werd gemeten met vijf door Schaufeli en Bakker (2004) ontwikkelde items, waaronder 'Mijn werk inspireert mij' (0 = 'nooit', 6 = 'altijd', $\alpha = .90$).

Statistische procedure. Via confirmatieve factoranalyse werd nagegaan of de factorstructuren van de vijf hiervoor genoemde concepten verschilden voor de twee onderscheiden groepen. Daartoe werden vier modellen getoetst en vergeleken. Het eerste model (M_1) bevatte 19 gemeten variabelen (items) en vijf factoren die onderling samenhangen. Elke factor correspondeerde met één van de vijf hiervoor besproken concepten (uitputting, toewijding, werkdruk, regelmogelijkheden, ongewenste intimiteiten), en de gemeten variabelen laadden steeds op de factor die ze zouden moeten meten. Dit model werd voor beide onderzoeksgroepen geschat, waarbij de parameters van dit model (factorladingen, factorvarianties en covarianties, errorvarianties van de items) konden variëren. Dit model was het uitgangspunt voor onze analyses.

Het tweede model (M_2) was identiek aan M_1 , zij het dat nu de factorladingen voor beide groepen gelijk gehouden werden. Als dit model niet acceptabel zou blijken, verschilt de structuur van de vijf gemeten begrippen voor internetrespondenten en de overige respondenten, hun scores op de desbetreffende vijf concepten kunnen niet zonder meer met elkaar worden vergeleken. Het derde model (M_3) was gelijk aan M_2 , maar de factorvarianties en -covarianties werden voor beide groepen constant gehouden. Het laatste model (M_4) was identiek aan M_3 , zij het dat nu ook de errorvarianties gelijk waren voor beide groepen. Wanneer M_4 acceptabel zou blijken, zijn er geen verschillen tussen

internetrespondenten en andere respondenten wat betreft de *structuur en de variantie* van hun antwoorden; deze kunnen zonder meer met elkaar vergeleken worden. Dat betekent echter niet dat hun *antwoorden* aan elkaar gelijk zijn. Om na te gaan of de gemiddelden van beide groepen respondenten op de vijf schalen gelijk zijn, werd een multivariate variantieanalyse (MANOVA) uitgevoerd, waarna univariate follow-up-toetsen werden toegepast.

Resultaten

Tabel 2 presenteert de resultaten van onze factoranalyses. Model M_1 , dat veronderstelde dat de items van de vijf onderscheiden concepten in beide groepen op de bedoelde wijze op deze concepten laadden, bleek goed bij de gegevens te passen (RMSEA = .062; NNFI > .90; waarden van lager dan .08 (RMSEA) en hoger dan .90 (NNFI) indiceren een acceptabele fit, Byrne, 2001). Het voor beide groepen aan elkaar gelijkstellen van de factorladingen leidde tot een significante verslechtering van de fit van het model met 47.3 chi-kwadratpunten bij een winst van 14 vrijheidsgraden, $p < .001$. Desalniettemin bleef de fit van het model acceptabel (RMSEA = .064, NNFI > .90); hoewel significant zijn deze verschillen praktisch gesproken te verwaarlozen. Voor beide groepen wordt dus een vergelijkbare factorstructuur teruggevonden, wat betekent dat hun antwoorden goed vergeleken kunnen worden. Het betekent ook dat Dillmans (2002) vrees dat het gebruik van internet leidt tot een '*increase in the proportion of meaningless or inaccurate responses*' (p. 489) voor de huidige onderzoeksgroep vooralsnog niet wordt ondersteund. Dan zouden we immers voor de internetrespondenten duidelijk lagere factorladingen hebben moeten vinden dan voor de overige respondenten.

Tabel 2 *Resultaten van een serie confirmatieve factoranalyses waarin de structuur van een 5-factormodel wordt vergeleken voor diegenen die via internet responderden (N = 245) met een vergelijkbare groep niet-internetrespondenten (N = 245)*

Model		Aantal df	Chi- kwadraat	RMSEA	NNFI
M_1	5-factormodel, factoren gecorreleerd	284	583.1	.062	.92
M_2	Als M_1 , factorladingen echter gelijk	298	630.4	.064	.92
M_3	Als M_2 , factorvarianties/-covarianties echter gelijk	313	689.0	.066	.91
M_4	Als M_3 , item error varianties echter gelijk	332	1238.3	.092	.80

N.B. Verschil M_1 vs. M_2 : chi-kwadraat is 47.3 ($df = 14$); verschil M_2 vs. M_3 : chi-kwadraat is 58.6 ($df = 15$); verschil M_3 vs. M_4 : chi-kwadraat is 549.3 ($df = 19$); alle p 's < .001.

Het voor beide groepen gelijkstellen van de factorvarianties en -covarianties (M_3) leidde evenmin tot een in praktisch opzicht noemenswaardige verslechtering van de fit van het model. Toen echter ook de errorvariantie van de items aan elkaar gelijk werden gesteld (M_4), bleek de fit van het model zeer sterk te verslechteren en was niet meer acceptabel; in 10 van de 19 gevallen bleek de variantie voor de internetrespondenten significant groter te zijn dan voor de andere respondenten. Respondenten verschilden vooral op de items van de emotionele-uitputtingsschaal en de toewijdingsschaal.

Dit blijkt ook uit de in tabel 3 gepresenteerde resultaten van onze multivariate variantieanalyse. De tabel laat zien dat de varianties van werkdruk, regelmogelijkheden en het ervaren van ongewenste intimiteiten niet sterk variëren voor beide groepen. De verschillen zijn duidelijk groter voor emotionele uitputting en toewijding. Als we kijken naar de gemiddelden, dan blijken de

internetrespondenten op *alle* onderscheiden concepten significant ongunstiger te scoren dan de anderen.

Tabel 3 *Vergelijking van de gemiddelde scores van een groep internetrespondenten (N = 245) met een vergelijkbare groep niet-internetrespondenten (N = 245) op enkele werkgerelateerde variabelen*

Variabelen	Wijze van responderen				Univariate F (1, 488)
	Internet M	Internet SD	Pen-en-papier M	Pen-en-papier SD	
Werkdruk	2.20	.56	1.95	.57	23.9**
Regelmogelijkheden	2.71	.48	2.93	.52	23.6**
Ongewenste intimiteiten	1.26	.39	1.12	.33	18.8**
Emotionele uitputting	1.84	.81	1.32	.69	57.50**
Toewijding	4.66	1.34	4.94	1.18	5.6*

N.B. Multivariate $F(5, 484) = 16.0, p < .001$, * = $p < .05$, ** = $p < .001$.

Samenvatting/Discussie. De resultaten van onze analyses laten zien dat de factorstructuur van de vijf door ons bekeken concepten niet sterk verschilt voor internetrespondenten en andere respondenten. Wel blijkt de variantie van (de items van) twee van de vijf concepten (emotionele uitputting en toewijding) behoorlijk van elkaar te verschillen, zodanig dat deze groter is voor de internetrespondenten. Voorts rapporteren de internetrespondenten relatief ongunstige werkomstandigheden (hogere werkdruk, minder regelmogelijkheden, meer ongewenste intimiteiten) en werkuitkomsten (meer uitputting, minder toewijding).

Het feit dat voor de groep internetrespondenten relatief hoge varianties worden gevonden (met name voor uitputting en toewijding), betekent dat zij wat meer uiteenlopende antwoorden geven dan de andere respondenten. Dit zou kunnen betekenen dat zij inderdaad minder precieze antwoorden geven, dan wel zomaar wat aanklikken (Dillman, 2002); anderzijds vonden we dat de factorladingen voor beide groepen zeer sterk vergelijkbaar waren. Dit betekent dat onnauwkeurigheid bij het beantwoorden van de items minder voor de hand ligt als verklaring voor de relatief grote spreiding op deze items. Bovendien valt moeilijk in te zien waarom de internetrespondenten bij het beantwoorden van de items van uitputting en toewijding een gehaaste en onnauwkeurige stijl van invullen zouden hanteren, en niet voor de drie andere concepten. Het feit dat de variantie van de items en schalen relatief groot is voor de internetrespondenten, impliceert ook dat voor deze groep met gebrek aan anonimiteit en vertrouwelijkheid samenhangende sociale wenselijkheid geen belangrijke rol heeft gespeeld bij het beantwoorden van de items. Men zou dan immers verwachten dat de spreiding relatief laag zou zijn voor deze groep. Kortom, we zien geen aanwijzingen dat de kwaliteit van de via internet verzamelde gegevens lager is dan die van de via pen-en-papier verkregen gegevens.

Wel blijken de deelnemers die via internet respondeerden relatief ongunstige werkomstandigheden te rapporteren. De oorzaak hiervoor is onduidelijk. Er zou sprake kunnen zijn van een methode-effect, maar in eerder onderzoek naar verschillen tussen de antwoorden van internetrespondenten en anderen werd niet eerder geconstateerd dat diegenen die via internet respondeerden consequent negatiever (of positiever) waren dan de andere respondenten (Bandilla et al., 2003; Dillman et al., 2001). Ook zouden internetrespondenten hun werkomstandigheden en werkuitkomsten inderdaad als slechter kunnen beschou-

wen dan anderen. In dat geval is er echter geen duidelijke reden waarom deze respondenten ervoor kiezen om hun antwoorden via internet te geven.

4 Discussie

Doel van deze bijdrage was de lezer enig inzicht te geven in de kansen die het internet biedt om gegevens te verzamelen, alsmede de valkuilen die hier kunnen spelen. De voordelen van het internet zijn evident; het kan een relatief goedkope manier zijn om veel gegevens te verzamelen. Vanwege het ontbreken van interviewers is interviewerbias als vertekende factor uitgeschakeld; omdat de gegevens niet meer afzonderlijk hoeven te worden ingevoerd, is ook hier een bron van mogelijke fouten verdwenen. Tegenover deze voordelen staan echter nadelen en punten die extra aandacht behoeven. Het meest evidente probleem betreft de kwaliteit van de via internet verkregen gegevens. Daaraan werden twee aspecten onderscheiden, betreffende (1) de mate waarin de verkregen en de gewenste steekproef corresponderen: de respons op internetvragenlijsten is vaak relatief laag, en personen die via internet responderen wijken zowel qua achtergrondkenmerken als qua antwoorden af van respondenten die via meer traditionele methoden worden benaderd; en (2) de mate waarin de via internet verkregen antwoorden afwijken van via andere methoden verkregen gegevens: er zijn vermoedens dat personen die via internet responderen minder precies en zorgvuldig antwoorden dan anderen.

Onze illustratie van het onderzoek onder ruim 45.000 thuiszorgmedewerkers liet inderdaad zien dat diegenen die via internet respondeerden (slechts 1.3% van de totale onderzoeksgroep) zowel in termen van hun achtergrondkenmerken (leeftijd, sekse, opleidingsniveau en type functie) als wat betreft de door hen gegeven antwoorden afweken van diegenen die een papieren vragenlijst invulden. Bemoedigend was dat de factorstructuur van de door ons onderzochte concepten niet verschilde over de onderscheiden groepen, wat betekent dat de antwoorden van beide groepen respondenten zinvol met elkaar vergeleken kunnen worden. Onduidelijk is echter waaraan de geconstateerde verschillen in de antwoorden van beide groepen toegeschreven konden worden. Er was geen reden om te vermoeden dat internetrespondenten minder nauwkeurig antwoordden dan anderen; we zagen in de antwoorden van de internetrespondenten geen aanwijzingen dat deze het invullen van vragenlijsten als een computerspelletje beschouwden (Dillman, 2002).

Implicaties. Het wetenschappelijk onderzoek naar de kwaliteit van via internet verkregen gegevens staat nog in de kinderschoenen. Het ontbreken van experimentele en toetsende studies betekent dat er geen sprake is van een zich gestaag ontwikkelende hoeveelheid harde kennis; wel kunnen er, op basis van het beschikbare bewijs, allerlei vuistregels worden geformuleerd. Eén daarvan is dat internet vooral ingezet kan worden als de beoogde respondenten daadwerkelijk benaderd kunnen worden via internet. Het lijkt aannemelijk dat dit probleem in de tamelijk overzichtelijke context van onderzoek binnen organisaties relatief gemakkelijk ondervangen kan worden. Organisaties weten immers of hun medewerkers via internet benaderd kunnen worden en zullen gewoonlijk een actueel overzicht van hun adressen hebben.

Een andere vuistregel is dat het gebruik van internet in vergelijking met meer traditionele methoden om gegevens te verzamelen niet zal leiden tot verschuivingen in de factorstructuur (in casu de betekenis) van concepten. Dat betekent dat de gegevens van op verschillende manieren benaderde respondenten zinvol met elkaar vergeleken kunnen worden. Tegelijk is duidelijk dat

de antwoorden van via internet antwoordende respondenten kunnen verschillen van die van andere respondenten, en dat die verschillen vaak moeilijk geïnterpreteerd kunnen worden. Deze constatering leidt tot de aanbeveling om respondenten het liefst op één bepaalde wijze te benaderen, zodat methode-effecten als verklaring van verschillen tussen respondenten uitgeschakeld zijn. Ook is het verstandig om de gegevens van herhaalde metingen steeds op dezelfde wijze te verzamelen; zo blijven de gegevens over de tijd vergelijkbaar. Samenvattend, het verzamelen van gegevens via het internet lijkt voor sociaal, arbeids- en organisatiepsychologen een in potentie nuttige aanvulling op de gereedschapskist te zijn. In bepaalde situaties (vooral als onduidelijk is of men de beoogde respondenten via internet kan benaderen, of als er een groot risico is op selectieve respons) is het veiliger of zelfs noodzakelijk om de gegevens met de gebruikelijke pen-en-papiermethode te verzamelen. In andere situaties (met name wanneer men onderzoek verricht in organisaties) zullen onderzoekers echter de nadelen van het verzamelen van gegevens via internet kunnen vermijden, zodat zij ten volle kunnen profiteren van de evidente voordelen van deze benadering. Gegeven deze voordelen mag worden verwacht dat het verzamelen van gegevens via internet de komende jaren een grote vlucht zal nemen.³

Noten

- 1 Hierbij moet worden aangetekend dat de door Sills en Song (2002) bekeken studies vaak een niet-wetenschappelijk karakter hadden, wat de aard en uitvoering van en de respons op de studies kan hebben beïnvloed.
- 2 Een derde item van deze schaal, 'Kunt u deelnemen aan besluiten die uw werk raken', bleek niet met de andere twee samen te hangen, vermoedelijk omdat zulks voor de onderhavige zeer laag opgeleide onderzoeksgroep niet van toepassing is. Bij de hogere functieniveaus werden wél acceptabele factorladingen van dit item gevonden.
- 3 Meer informatie over internet-enquêtes is te vinden op een in het kader van het 5e Kaderprogramma FP5 door de Europese Unie gefinancierde internetpagina (www.websm.org)

Literatuur

- Bandilla, W., Bosnjak, M. & Altdorfer, P. (2003). Survey administration effects? A comparison of web-based and traditional written self-administered surveys using the ISSP environment module. *Social Science Computer Review*, 2, 235-243.
- Byrne, B.M. (2001). *Structural equation modeling with AMOS*. Mahwah, NJ: Erlbaum.
- Dijkstra, W. (1983). *Beïnvloeding van antwoorden in survey-interviews*. Delft: Bouwplatenwerkplaats (academisch proefschrift).
- Dillman, D.A. (1991). The design and administration of surveys. *Annual Review of Sociology*, 17, 225-249.
- Dillman, D.A. (2000). *Mail and internet surveys: The tailored design method*. New York: Wiley.
- Dillman, D.A. (2002). Navigating the rapids of change: Some observations on survey methodology in the early 21st century. *Public Opinion Quarterly*, 66, 473-494.
- Dillman, D.A. & Bowker, D.K. (2001). The web questionnaire challenge to survey methodologists. In U.D. Reips & M. Bosnjak (Eds.), *Dimensions of internet science* (pp. 159-178). Lengerich: Pabst Science Publishers.
- Dillman, D.A., Phelps, G., Tortora, R., Swift, K., Kohrell, K. & Berck, J. (2001). *Response rate and measurement differences in mixed mode surveys: Using mail, telephone, interactive voice response and the internet*. Ongepubliceerd paper, beschikbaar op <http://survey.sesrc.wsu.edu/dillman/papers.htm>.

- Donovan, M.A., Drasgow, F. & Probst, T.M. (2000). Does computerizing paper-and-pencil job attitude scales make a difference? New IRT analyses offer insight. *Journal of Applied Psychology*, 85, 305-313.
- Hox, J.J. & De Leeuw, E.D. (1994). A comparison of nonresponse in mail, telephone, and face-to-face interviews: Applying multilevel modeling to meta-analysis. *Quality & Quantity*, 28, 329-344.
- Karasek, R.A. (1985). *Job Content Instrument: Questionnaire and user's guide*. Los Angeles: Department of Industrial and Systems Engineering, University of Southern California.
- Karasek, R.A. & Theorell, T. (1990). *Healthy work: Stress, productivity, and the reconstruction of working life*. New York: Basic Books.
- King, W.C. & Miles, E.W. (1995). A quasi-experimental assessment of the effect of computerizing noncognitive paper-and-pencil measurements: A test of measurement equivalence. *Journal of Applied Psychology*, 80, 643-651.
- Schaufeli, W.B. & Bakker, A.B. (2004). Bevlogenheid: Een begrip gemeten. *Gedrag & Organisatie*, 17, 89-112.
- Schaufeli, W. & Van Dierendonck, D. (2000). *UBOS: Utrechtse Burnout Schaal*. Lisse: Swets Test Publishers.
- Sills, S.J. & Song, C. (2002). Innovations in survey research: An application of web-based surveys. *Social Science Computer Review*, 1, 22-30.
- Stanton, J.M. & Rogelberg, S.G. (2002). Beyond online surveys: Internet research opportunities for industrial-organizational psychology. In S.G. Rogelberg (Ed.), *Handbook of research methods in industrial and organizational psychology* (pp. 275-294). Malden, MA: Blackwell.
- Swoboda, W.J., Mühlberger, M., Weitkunat, R. & Schneeweiss, S. (1997). Internet surveys by direct mailing: An innovative way of collecting data. *Social Science Computer Review*, 15, 242-255.
- Taris, T.W. (2000). *A primer in longitudinal data analysis*. London: Sage Publications.
- Taris, T.W., Kompier, M.A.J., Geurts, S.A.E., Schreurs, P.J.G., Schaufeli, W.B., De Boer, E. & Sepmeijer, K.J. (2004). Stress management interventions in the Dutch domiciliary care sector: Findings from 81 organizations. *International Journal of Stress Management*, 10, 297-325.
- Thompson, L.F., Surface, E.A., Martin, D.L. & Sanders, M.G. (2003). From paper to pixels: Moving personnel surveys to the web. *Personnel Psychology*, 56, 197-227.

Summary

Web-based data collection: an update

Toon Taris, Paul Schreurs & Kees Jan Sepmeijer, *Gedrag & Organisatie*, Volume 18, June 2005, nr. 3, pp. 181-195.

The rapid increase in the number of internet users has given web-based surveys the potential to become a powerful tool in survey research. The present study discusses the pros and cons of this technique for gathering data. Pros include speed, cost effectiveness and exclusion of interviewer and data entry errors; cons include low response rates, selective response, and possible problems with data quality. As an illustration we present the results of a study among well over 45,000 Dutch home care employees, showing that men, higher educated and younger employees were more likely to respond through the internet than others. Confirmatory factor analysis revealed that whereas the factor structures of selected work-related concepts were similar for respondents who used the internet vs. other respondents, their average scores differed strongly. Implications of these findings on data collection practices in social, work and organizational psychology are discussed.

Key words: data collection, response, validity, representativeness