

User Information Satisfaction: een gebruikersperspectief op het succes van informatiesystemen

M. Gelderman¹

1 Introductie

In 1993 bedroeg de totale omzet van de Nederlandse computerservice- en IT-bureaus meer dan 8 miljard gulden (Centraal Bureau voor de Statistiek 1995, p. 33). Het is aannemelijk dat een groot deel van dit bedrag wordt uitgegeven aan de ontwikkeling van informatiesystemen (IS). Bovendien op dit bedrag komen de interne ontwikkelingskosten die een deel van het Nederlandse bedrijfsleven maakt. Als dergelijke bedragen worden uitgegeven, loont het de moeite om te kijken of deze schaarse middelen wel optimaal worden aangewend. Er bestaat een behoefte aan evaluatie van IS, opdat optimaal van de ingezette middelen geprofiteerd kan worden en gemaakte fouten een maximaal leereffect ressembleren. Reeds in 1973 constateerde het Committee on Management Information Systems van de American Accounting Association deze behoefte en gaf tegelijk aan dat adequate procedures en theorie ontbraken. Nog in 1990 stelden Abdel-Hamid en Madnick dat 'failure to learn from mistakes has been a major obstacle to improving software project management'. Sinds 1973 lijkt er niet veel veranderd te zijn. Er zijn echter wel positieve ontwikkelingen te melden.

Zo kan bovengenoemd probleem voor een deel worden ondervangen door het gebruik van user information satisfaction (UIS), zoals dat in dit

artikel wordt gepresenteerd. Vanzelfsprekend biedt ook UIS geen panacee; het is echter wel een interessante ontwikkeling die geleidelijk zijn weg van de academische wereld naar de praktijk vindt. Een recent empirisch onderzoek in Canada laat zien dat 26% van de onderzochte (grote) ondernemingen op de een of andere manier de gebruikerstevredenheid met EDP/MIS (electronic data processing/management information systems) meet (Conrath en Mignen 1990). Uit hetzelfde onderzoek blijkt dat men over het algemeen van mening is dat de baten van de evaluatie opwegen tegen de kosten. In paragraaf 2 wordt op het belang van evaluatie van IS ingegaan.

In de daarop volgende paragraaf wordt ingegaan op de voor- en nadelen verbonden aan de evaluatiecriteria 'mate van gebruik' en 'gebruikertevredenheid'. In paragraaf 4 wordt nauwkeuriger ingegaan op de ontwikkeling van een gestandaardiseerd instrument voor het meten van UIS.² Bij de bespreking wordt een chronologische indeling van de ontwikkeling van dit instrument in de (veelal Amerikaanse) literatuur aangehouden. In paragraaf 5 worden de huidige stand van zaken en een Nederlandse vertaling van dit instrument besproken. Ten slotte worden de samenvatting en conclusies van het artikel gepresenteerd.

Met dit artikel wordt een leemte in de Nederlandstalige literatuur omtrent kwaliteit van informatiesystemen ingevuld, zoals die kan worden afgeleid uit het door Rijsenbrij (1993) gepubliceerde overzicht van ruim 100 Nederlandstalige publicaties over dit onderwerp – een overzicht dat overigens zeker de aandacht van de geïnteresseerde lezer verdient. Uit dit overzicht blijkt dat kwaliteit van informatiesystemen in de

¹ M. Gelderman is verbonden aan de vakgroep kosten- en winstvraagstukken van de faculteit der economische wetenschappen en econometrie van de Vrije Universiteit te Amsterdam. Hij verricht promotie-onderzoek naar het succes van management support systemen.

Nederlandstalige literatuur met name is beschouwd vanuit de ontwikkelzijde, de verhouding opdrachtgever/ontwikkelaar en de optiek van de (controleerend) accountant; de waarde van een informatiesysteem voor de uiteindelijk gebruiker is onderbelicht gebleven. De gekozen benadering sluit aan bij de conclusie van Van Egten (1994) dat het objectief meten van de kwaliteit van informatie niet mogelijk is, maar dat beoordeling daarentegen aan het subject, i.c. de gebruiker gebonden dient te zijn.

2 Het belang van evaluatie

Evaluatie van IS zal onder meer noodzakelijk zijn om een adequaat oordeel te kunnen vormen omtrent het functioneren van de voor de ontwikkeling van IS verantwoordelijke afdelingen en personen. De evaluatie vormt op die manier een onderdeel van het normale proces van prestatiebeoordeling in een organisatie (Dickson en Wetherbe 1985). Het einddoel van de evaluatie is natuurlijk altijd een beter IS, c.q. een betere informatievoorziening aan het management. Ook kunnen de resultaten van een evaluatie gebruikt worden om het ontwikkelingsproces te verbeteren.

Tabel 1: De mate van het gebruik van verschillende succesdefinities in internationaal gepubliceerd empirisch onderzoek naar succesfactoren van management support systems in de periode 1975-1994

<i>Afhankelijke variabele</i>	<i>Aantal studies</i>
Mate van gebruik	17
User Information Satisfaction	27
Overige kenmerken	13

(Bron: Gelderman 1995a).

Evaluatie dient niet alleen om terug te kijken op het verleden. Een belangrijke functie is het 'continu' monitoren of een systeem nog wel voldoet, met andere woorden verschaft het IS nog steeds die informatie die noodzakelijk is voor het functioneren van de organisatie. In principe kan zelfs een stap verder worden gegaan en kan gekeken worden of er in de toekomst aanpassingen van het IS noodzakelijk zijn. Evaluatie signaleert respectievelijk voorkomt op deze wijze dat onaanvaardbare discrepanties tussen de informatievoorziening en de informatiebehoefte van de organisatie ontstaan. Als de evaluatie leidt tot de conclusie dat er problemen (te verwachten) zijn, biedt diezelfde evaluatie een aanknopingspunt

voor het oplossen van deze problemen. De resultaten geven aan wat de tekortkomingen zijn, zodat het systeem kan worden aangepast of een nieuw systeem kan worden ontwikkeld, met als resultaat een betere informatievoorziening.

De beschikbaarheid van goede evaluatiecriteria is niet alleen belangrijk voor de praktijk, maar ook voor onderzoek op het gebied van IS. Een belangrijk onderzoeksterrein is het vaststellen van succesfactoren voor ontwikkeling en implementatie van IS.³ Bij dergelijk onderzoek wordt het succes van IS gerelateerd aan de score op één of meer onafhankelijke variabelen. Vanzelfsprekend is het hierbij noodzakelijk dat 'succes' zo adequaat mogelijk wordt gemeten. Daarnaast is het van belang dat succes door verschillende onderzoekers op dezelfde wijze wordt gedefinieerd, zodat studies onderling vergelijkbaar zijn en eventuele verschillen in uitkomst aan de onafhankelijke variabelen of omgevingsfactoren kunnen worden toegeschreven. Tabel 1 geeft een overzicht van de succesmaatstaven zoals die zijn gebruikt in onderzoek op een specifiek deelgebied van IS-research, namelijk succesfactoren voor management support systems. Het merendeel der onderzoekers kiest voor het gebruik van een vorm van UIS als succescriterium. Gezien het feit dat de mate van gebruik met name in oudere studies is toegepast, is de suprematie van UIS in het huidige wetenschappelijk onderzoek duidelijk.

3 Evaluatiecriteria

In de vorige paragraaf is het belang van de evaluatie van IS geschetst. Deze paragraaf bespreekt twee criteria — mate van gebruik en user information satisfaction — die gebruikt kunnen worden om deze evaluatie concreet gestalte te geven.⁴ Daadwerkelijk evalueren is immers nodig om de hierboven genoemde doelen te realiseren. De ideale evaluatie verschaft tegen aanvaardbare kosten (1) een perfect oordeel over de verantwoordelijke personen en afdelingen, (2) perfecte aanwijzingen voor de verbetering van het ontwikkelingsproces en geeft (3) op exact het juiste moment precies aan welke veranderingen noodzakelijk zijn. De ideale evaluatie bestaat niet en zal waarschijnlijk ook nooit gevonden worden. In plaats van direct het 'succes' van een IS te meten moeten benaderingen voor dit succes gezocht worden: maatstaven die gemakkelijk vast te stellen

zijn en waarvan het aannemelijk is dat zij samenhangen met het werkelijke succes van het is. Dit werkelijke succes is een maximale bijdrage aan de doelrealisatie van de organisatie (DeBrabander en Thiers 1984; Doll en Torkzadeh 1988; Ives, Olson en Baroudi 1983; Kim 1990; Rivard en Kaiser 1989).

Indien er in de praktijk al geëvalueerd wordt, kiest men voor relatief eenvoudige middelen (Conrath en Mignen 1990). Een eerste facet is dan de technische evaluatie, waarbij downtime, aantal storingen, aantal fouten en dergelijke zaken worden gemeten. In dit artikel wordt niet op deze vorm van evaluatie ingegaan; de discussie beperkt zich tot de evaluatie van de *informatieverschaffende functie* van het is.⁵ Hiermee wordt aansluiting gezocht bij een ontwikkeling – die bij voorbeeld bij de MESDAG-onderzoeksgroep aan de Vrije Universiteit te signaleren valt – waarbij is vanuit de vraag- (dat is de gebruiker), in plaats van de aanbodzijde wordt bekeken.⁶ In concreto betekent dit dat zowel de technische werking van het is, als het gebruik van het is als transactieverwerkend of productiesysteem (bij voorbeeld bij banken of verzekeringsmaatschappijen) niet besproken zullen worden. Er wordt slechts gekeken naar het gebruik van het systeem om informatie te verschaffen. Of deze informatie direct, of via een intermediair wordt verschaft wordt hierbij niet van primair belang geacht.

3.1 *Mate van gebruik*

Eén van de meest voor de hand liggende, en derhalve reeds lang toegepaste maatstaven, is de mate van gebruik van het is. Hierbij kan zowel worden gedacht aan het aantal raadplegingen/transacties als de gebruikstijd van het systeem. Beide criteria gaan ervan uit dat de gebruiker een goed systeem wel en een slecht systeem niet zal gebruiken. Enerzijds wordt verondersteld dat de gebruiker de keuze heeft het systeem al dan niet te gebruiken (Barki en Huff 1990; Barki en Huff 1985; Baroudi, Olson en Ives 1986; Doll en Torkzadeh 1988; Ginzberg 1978; Ives, Olson en Baroudi 1983; Ives en Olson 1984; Lucas 1978; Olson en Ives 1981). Anderzijds wordt er vanuit gegaan dat gebruikers hun informatiebehoefte voorziet zullen gebruiken. Een essentiële vooronderstelling is dat het gewenst is dat een

systeem in de informatiebehoefte van de gebruiker voorziet, dat dit niet zo vanzelfsprekend is als in eerste instantie het geval lijkt, zal uit het vervolg van de discussie blijken.

Het toepassen van de mate van gebruik als evaluatiecriterium heeft als belangrijkste voordeel dat het eenvoudig en goedkoop is. Vaak kan de meting zelfs automatisch door de software worden uitgevoerd. Daarbij komt dat het criterium mate van gebruik voor alle betrokkenen begrijpbaar is. Nadelen zijn echter ook aanwezig. Allereerst dienen de hierboven genoemde vooronderstellingen te kloppen. Een eerste punt van discussie is dan of gebruikers hun eigen informatiebehoefte kennen. Indien dit het geval is, dienen gebruikers een juist en volledig model van hun beslissingssituatie te hebben (Ackoff 1967); dit is virtueel onmogelijk en onderzoek naar expert judgment geeft aan dat dit inderdaad problematisch is (MacIntosh 1985).

Ten tweede dient de gebruiker een systeem dat in de informatiebehoefte voorziet ook daadwerkelijk te gebruiken. Ook dit zal niet altijd het geval zijn, de gebruikers zullen een systeem slechts gebruiken indien dit systeem in hun *perceptie* in die informatiebehoefte voorziet. Indien gebruikers een verkeerd beeld hebben van de capaciteiten van het is zullen zij dit niet gebruiken. De oplossing zal dan echter veelal liggen in extra voorlichting en educatie en niet in verbetering van het systeem.

Ten derde impliceert het gegeven dat de mate van gebruik beïnvloed wordt door de match tussen de informatiebehoefte van de gebruiker en de verschaft informatie dat deze maatstaf afhankelijk is van de doelcongruentie tussen gebruiker en organisatie. Deze zal echter niet per definitie aanwezig zijn. Indien er bij voorbeeld beursgegevens aan het systeem worden toegevoegd zal de manager die privé aandelen bezit het systeem meer gaan gebruiken, de organisatie profiteert hier echter niet van. Een ander voorbeeld van de afwezigheid van doelcongruentie wordt gesignaleerd door Ives, Olson en Baroudi (1983). Zij merken op dat een is om oneigenlijke redenen gebruikt kan worden. De gebruiker 'vertrouwt' dan op een is terwijl het bekend is dat dit systeem niet betrouwbaar is. Een reden om dit te doen kan gelegen zijn in het afwentelen van verantwoordelijkheid. Hiernaast bestaat het probleem dat gebruik van het is niet altijd vrijwillig is (Lucas

1978). Ook wanneer er geen expliciete verplichting bestaat om een is te gebruiken kan men zich afvragen of er sprake is van vrijwillig gebruik. Zo kan de organisatiecultuur gebruik van het systeem afdwingen. Ook is het mogelijk dat de informatie uit het systeem op geen enkele andere wijze beschikbaar is (Ginzberg 1978).

Van een evaluatiecriterium mag voorts verwacht worden dat een hogere score op het criterium altijd gepaard gaat met betere performance. Bij toepassing van mate van gebruik is dit niet per definitie het geval. Een drietal voorbeelden mag dit verduidelijken:

- 1 Indien gebruikstijden worden gemeten, leidt een tragere respons van het systeem tot meer gebruik. Het opvragen van dezelfde informatie kost dan immers meer tijd. Het systeem wordt nu positiever beoordeeld, terwijl de prestaties zijn gedaald.
- 2 Indien aantallen raadplegingen (transacties) worden gemeten leidt een inefficiënte opbouw van het systeem (bij voorbeeld informatie die in één rapport zou kunnen staan wordt verdeeld over meerdere rapporten) tot een *beter*e evaluatie. Ook dit is niet gewenst.
- 3 In de meting van de mate van gebruik zit 'ruis'. Zo is het denkbaar dat een systeem vlak na introductie meer gebruikt zal worden dan later. In eerste instantie experimenteren gebruikers met het systeem en/of weten nog niet hoe ze de gewenste informatie kunnen opvragen, later wordt het systeem efficiënter gebruikt. Het gebruik zal *afnemen*, terwijl de organisatie per saldo *beter* af is.

Bij het gebruik van een criterium ten behoeve van oordeelsvorming is verder de vergelijkbaarheid met derden of met een ideale score van belang. De mate van gebruik leent zich hier niet goed voor. Een 'ideale mate van gebruik' bestaat niet en onderlinge vergelijkbaarheid van gebruikscijfers is slechts in zeer beperkte mate mogelijk.

Voorts is het onderscheidend vermogen van het gebruikscriterium beperkt. Het is aannemelijk dat gebruikers ophouden een systeem te gebruiken wanneer een grens wordt overschreden. Wanneer dit zo is, onderscheidt het criterium 'mate van gebruik' alleen de zeer slechte systemen van de minder slechte systemen (Barki en Huff 1985). Resultaten van een recente meta-analyse wijzen uit dat de toepassing van de mate van gebruik als succes criterium over de jaren heen steeds tot

geringere resultaten heeft geleid (Gelderman 1995a). Indien inderdaad steeds minder systemen zo slecht zijn dat zij beneden de grenswaarde vallen, bevestigt dit het hiervoor gestelde.

Ten slotte geeft het gebruik van de mate van gebruik als succes criterium bij problemen slechts aan 'dat er iets fout zit' en niet wat er fout zit, c.q. waar de oplossing gezocht dient te worden. Overigens zal bij de bespreking van uis – dat over het algemeen beter scoort dan mate van gebruik – blijken dat hier sprake is van een trade-off: het aanwijzen van specifieke oplossingsmogelijkheden gaat ten koste van de algemene toepasbaarheid van een succes criterium.

Samenvattend is de mate van gebruik van een is geen ideaal evaluatiecriterium. Dit impliceert niet dat mate van gebruik niet vastgesteld mag worden, er zal echter ten minste nog een aantal andere aspecten in de evaluatie betrokken moeten worden.

3.2 Gebruikerstevredenheid

Een belangrijke vooruitgang wordt bereikt door te evalueren op basis van tevredenheid van de gebruikers⁷ met het systeem. Met de introductie van het concept gebruikerstevredenheid wordt aan een aantal tekortkomingen van het gebruikscriterium tegemoet gekomen. Zo kan een grotere tevredenheid met een is niet worden bereikt met kunstmatige zaken als een slechtere responstijd en het inefficiënt inrichten van het systeem. Ook is het criterium gebruikerstevredenheid toe te passen in omstandigheden waar sprake is van expliciet of impliciet verplicht gebruik van het is. Bovendien dekt de gebruikerstevredenheid grotere verschillen in het succes van systemen, dan de mate van gebruik. Een aantal punten verdient echter ook bij de toepassing van gebruikerstevredenheid als evaluatiecriterium de aandacht.

Zo bekritiseert Melone (1990) het uis-criterium omdat het veronderstelt dat een tevreden gebruiker het is ook daadwerkelijk zal gebruiken. Dit lijkt mij echter geen al te gewaagde veronderstelling. Wel problematisch is dat ook uis doelcongruentie tussen organisatie en gebruiker veronderstelt. Evenals bij de mate van gebruik het geval is, is het mogelijk om op oneigenlijke wijze een hogere gebruikerstevredenheid te bewerkstelligen indien de doelen van gebruiker en organisatie niet

parallel lopen. Indien beursdata worden verschaft zal de manager die privé-aandelen aanhoudt het systeem niet alleen meer gaan gebruiken, maar er ook tevredener mee zijn. Verder vraagt toepassing van gebruikerstevredenheid als evaluatiecriterium meer inspanning van de organisatie, dan het meten van de mate van gebruik van een systeem: UIS kan niet automatisch gemeten worden. Ook zal bij toepassing enige uitleg richting gebruikers noodzakelijk zijn.

In eerste instantie blijft bovendien een tweetal onduidelijkheden bestaan. Allereerst is het de vraag of en op welke wijze het mogelijk is om gebruikerstevredenheid te vergelijken met derden of met een ideaal. Ten tweede is het niet duidelijk of de toepassing van gebruikerstevredenheid als evaluatiecriterium aangeeft in welke richting de oplossing van een eventueel probleem gezocht dient te worden.

Voor het probleem van de doelcongruentie is geen directe oplossing voorhanden. Het gebruik van een *gestandaardiseerde* meetwijze van gebruikerstevredenheid zal echter de kosten van de evaluatie verlagen, maakt benchmarking mogelijk en geeft *globaal* de richting aan waarin de oplossing van problemen gezocht dient te worden. De hier bedoelde 'gestandaardiseerde meetwijze' kenmerkt zich er door dat de gebruikerstevredenheid wordt vastgesteld met behulp van een vragenlijst waarin over iedere samenstellende factor van de gebruikerstevredenheid – voorbeelden van dergelijke factoren zijn nauwkeurigheid van de informatie en presentatie van de informatie – een aantal, deels overlappende, vragen wordt gesteld. De scores op de individuele vragen worden vermenigvuldigd met een wegingsfactor en vervolgens opgeteld om te komen tot een score op de factor.⁸ De scores op de factoren worden weer gewogen en opgeteld om tot een totale score voor de gebruikerstevredenheid te komen.⁹ Om benchmarking en gebruik van vooraf vastgestelde wegingsfactoren mogelijk te maken dient natuurlijk wel *altijd* met een *identieke* vragenlijst te worden gewerkt. Aanpassing van de vragen zou bovendien de validatie van een instrument teniet doen. Ook voor onderzoeksdoeleinden is het gewenst met een gestandaardiseerde vragenlijst te werken. Op deze manier kunnen onderzoeksresultaten vergeleken en geïntegreerd worden. Deze gestandaardiseerde vragenlijst wordt in het vervolg van dit artikel het User

Information Satisfaction Instrument, oftewel UIS-instrument genoemd. Zoals gebruikelijk is bij standaarden, blijkt er meer dan een versie van het UIS-instrument te zijn.

4 Ontwikkeling van het UIS-instrument

De periode voor 1983 valt te beschouwen als de prehistorie van het UIS-tijdperk.¹⁰ Een aantal personen, die later op dit gebied zouden publiceren, was wel met onderzoek naar UIS bezig, maar de resultaten hiervan zijn met name terug te vinden in ongepubliceerde PhD-thesen, working papers en congrespresentaties. Een uitzondering vormt Swanson (1974), die het begrip 'mis-appreciation' introduceert en dit ook op een formele wijze – met behulp van een meetinstrument (vragenlijst) – operationaliseert. Verder zijn er enige publicaties (bij voorbeeld Alavi en Henderson 1981; Ginzberg 1981; McIntyre 1982; Powers en Dickson 1973; Vasarhelyi 1977) te vinden waarin een relatie gelegd wordt tussen op *informele* wijze vastgestelde gebruikerstevredenheid met een is en een aantal succesfactoren.

Bailey en Pearson

In 1983 verschijnt in *Management Science* het artikel 'Development of a tool for measuring and analyzing computer user satisfaction' van Bailey en Pearson (1983). In dit artikel (dat een weergave is van het promotieonderzoek van Pearson) wordt de ontwikkeling van een vragenlijst voor het meten van UIS getoond. Het gepresenteerde instrument bevat 39 verschillende factoren, die ieder gemeten worden met behulp van een viertal deelvragen.¹¹ De keuze van deze factoren is gebaseerd op theorie, bestaand onderzoek, interviews met 'data processing professionals' en interviews met managers. Laatstgenoemde managers hebben ook de vragenlijst ingevuld teneinde betrouwbaarheid en voor zover mogelijk validiteit van het instrument vast te stellen; de resultaten van dit onderzoek blijken bevredigend. Wel dient te worden opgemerkt – en de auteurs doen dit ook – dat het gebruik van dezelfde groep managers voor het ontwikkelen en testen van het instrument tot geflatteerde waarden kan leiden.

Ives, Olson en Baroudi

Vijf maanden later verschijnt in *Communications of the ACM* een artikel van de hand van Ives, Olson en Baroudi (1983), dat voortbouwt op het

werk van Bailey en Pearson. Het begrip User Information Satisfaction wordt hier als term geïntroduceerd en gedefinieerd als 'the extent to which users believe the information system available to them meets their information requirements' (Ives, Olson en Baroudi 1983, p. 785). Ives et al. herhalen het onderzoek van Bailey en Pearson naar betrouwbaarheid en validiteit en breiden dit enigszins uit. Ook hun resultaten zijn bevredigend. Tevens ontwikkelen zij een verkorte versie met 24 vragen – dat is twaalf factoren (schalen) gemeten met behulp van ieder twee items – van het oorspronkelijk 156 vragen tellende instrument.

Problemen

Een probleem met het door Bailey en Pearson ontwikkelde en door Ives et al. aangepaste instrument is dat het gebaseerd is op theorieën over het succes van IS. Er wordt niet alleen gemeten of de gebruikers tevreden zijn met het hen ter beschikking staande systeem: de uiteindelijke UIS-score is tevens gebaseerd op een aantal factoren waarvan men op theoretische gronden of op basis van eerder empirisch onderzoek veronderstelt dat zij tevredenheid *beïnvloeden*. Het instrument meet onder andere de mate van top management involvement, error recovery, de wijze waarop EDP-services aan gebruikers worden doorbelast en de ondersteuning door de leverancier. Dit leidt tot een tweetal problemen. Ten eerste wordt verondersteld dat de achterliggende theorie juist is. Dit levert zowel problemen op voor de praktische toepassing, als voor de toepassing in onderzoek. In gevallen waar er bij voorbeeld geen sprake is van top management involvement resulteert dit per definitie in een lagere UIS-score. Indien top management involvement geen invloed heeft op het succes van het systeem is dit echter niet terecht. Dit kan bovendien tot de cynische situatie leiden dat ontwikkelaars negatief worden beoordeeld wegens een gebrek aan top management involvement indien zij ondanks een weigering van het management om te participeren toch een redelijk systeem hebben kunnen bouwen. Analoge redeneringen gelden voor de andere op theorie gebaseerde variabelen. Logischerwijs is het instrument door hetgeen hiervoor is gesteld, niet of minder geschikt voor onderzoek. Het instrument kan niet gebruikt worden om vast te stellen of er een relatie bestaat tussen top management involvement en het succes van een systeem. Door opname van deze

factor in het instrument is dit immers per definitie het geval (Doll en Torkzadeh 1988).

Ten tweede is de gebruiker niet altijd in staat om alle vragen te beantwoorden. Sommige vragen zijn niet in iedere situatie van toepassing. Als er geen leverancier is, is de kwaliteit van zijn service van weinig belang.¹² Andere vragen, bij voorbeeld die over error recovery, zullen voor vele gebruikers te technisch van aard zijn (Doll en Torkzadeh 1988).

Tegenover genoemde nadelen staat ook een belangrijk voordeel: indien de theorie achter het instrument juist en volledig is, legt gebruik van het instrument de vinger onmiddellijk op de zere plek. De resultaten van de evaluatie geven dan bij voorbeeld aan dat de lage gebruikerstevredenheid voor een belangrijk deel wordt veroorzaakt door te weinig top management involvement.

Doll en Torkzadeh

In 1988 introduceren Doll en Torkzadeh een nieuw instrument voor het meten van 'end-user computing satisfaction' dat impliciet aan bovengenoemde bezwaren – de impliciete veronderstelling omtrent de juistheid van de achterliggende theorie en het niet altijd beantwoordbaar zijn van een aantal vragen – tegemoet komt en expliciet twee andere veranderingen introduceert. Allereerst meet het nieuwe instrument UIS met een specifieke applicatie en niet, zoals het Bailey en Pearson instrument, algemene UIS – een stap die noodzakelijk is door technische ontwikkelingen die maken dat de gebruiker end-user is geworden en zijn uitvoer niet meer bij de EDP-afdeling ophaalt. Verder betreft het nieuwe instrument de gebruikersvriendelijkheid van de applicatie in de beoordeling. Ook deze uitbreiding van het instrument wordt gedreven door de tendens richting end-user computing. Door het weglaten van de vragen over gebruikersvriendelijkheid is het instrument overigens ook bruikbaar om UIS te meten van een indirecte gebruiker van het systeem.¹³

Het nieuwe UIS-instrument is ontwikkeld door op basis van een pilot-study uit een relatief lange lijst met vragen de items met de hoogste betrouwbaarheid en de geringste redundantie ten opzichte van andere vragen te selecteren. Uiteindelijk bleven na de definitieve studie nog twaalf vragen over die de factoren content, accuracy, format, ease-of-use en timeliness afdekken. Tests op betrouwbaarheid en validiteit leiden tot tevre-

denstellende resultaten.¹⁴ Doll en Torkzadeh rapporteren eveneens percentiel-scores op hun instrument, die toepassing van de vragenlijst voor benchmarking mogelijk maken. Tevens wijst het instrument de richting van het probleem aan. De algemene toepasbaarheid van het instrument is echter ten koste gegaan van de specificiteit waarmee het probleem wordt aangeduid. Er wordt slechts aangegeven dat de problemen te maken hebben met één (of meer) van de genoemde vijf factoren (terwijl het Bailey en Pearson instrument 39 meer specifieke oorzaken kan aanwijzen).

5 Huidige stand van zaken en Nederlandse versie van het UIS-instrument

Ondertussen is het door Doll en Torkzadeh ontwikkelde instrument opnieuw getest in een door beide oorspronkelijke auteurs tezamen met Xia verrichte studie (Doll, Xia en Torkzadeh 1994). In dit geval is gebruik gemaakt van confirmatieve factoranalyse — LISREL-modellering — om de factorstructuur van het instrument en de betrouwbaarheid van de individuele items en de factoren vast te stellen. De resultaten van deze analyse zijn niet geheel eenduidig, maar in het algemeen bevredigend.

Op basis van het Doll en Torkzadeh instrument is tevens een Nederlandstalig UIS-instrument ontwikkeld. Allereerst is een vertaling gemaakt van het volledige instrument. Vervolgens is aan de éne helft van een groep studenten tijdens een bedrijfssimulatie de Nederlandse versie en aan de andere helft het oorspronkelijke Engelstalige instrument voorgelegd. De vragen met betrekking tot gebruikersvriendelijkheid zijn weggelaten, daar de informatie op papier werd verschaft. Tussen de antwoorden op de Nederlandse en de Engelse versie werden geen significante verschillen geconstateerd. De resultaten van een betrouwbaarheids- en factoranalyse komen overeen met de gegevens zoals die door Doll en Torkzadeh zijn gepubliceerd (Gelderman 1995b).¹⁵ Naar aanleiding van dit pilot-onderzoek is een aangepast Nederlandstalig instrument ontwikkeld, dat is weergegeven in tabel 2. Eén vraag – ‘do you find the output relevant’ – laadde zowel in het onderzoek van Doll en Torkzadeh, als in het pilotproject zeer sterk op meerdere factoren en is derhalve uit het instrument verwijderd. Aan de dimensie tijdigheid van informatie zijn twee vragen toegevoegd. Hiermee wordt de betrouwbaarheid van

deze dimensie verhoogd en wordt het instrument beter bruikbaar voor statistische analyses.¹⁶

Tabel 2: De vragen van het UIS-instrument

Inhoud van de informatie uit het systeem

Verschaft het systeem de informatie die u nodig heeft?

Komt de inhoud van de informatie overeen met uw behoeften?

Zijn de door het systeem versochte rapportages bijna precies wat u nodig heeft?

Verschaft het systeem voldoende informatie?

Nauwkeurigheid van de informatie

Is het systeem nauwkeurig?

Bent u tevreden met de nauwkeurigheid van het systeem?

Is naar uw gevoel het systeem betrouwbaar?

Vindt u dat u op het systeem kunt vertrouwen?

Presentatie van de informatie

Denkt u dat de uitvoer op een bruikbare wijze wordt gepresenteerd?

Is de informatie duidelijk?

Bent u tevreden met de layout van de uitvoer?

Is de uitvoer gemakkelijk te begrijpen?

Tijdigheid van de informatie

Krijgt u de informatie die u nodig heeft op tijd?

Verschaft het systeem informatie die up-to-date is?

Worden de gegevens in het systeem vaak genoeg bijgewerkt/ververst?

Worden de gegevens in het systeem snel genoeg bijgewerkt/ververst?

Gebruikersvriendelijkheid van het systeem

Is het systeem gebruikersvriendelijk?

Is het systeem gemakkelijk te gebruiken?

Is het systeem efficiënt?

De antwoorden op de vragen bestaan uit een vijfpunts-schaal met aan de beide uitersten ‘bijna nooit’ en ‘bijna altijd’. De in rozein weergegeven vragen zijn vertalingen van de oorspronkelijke (Engelse) vragen van Doll en Torkzadeh (Doll en Torkzadeh 1988). De cursief weergegeven vragen zijn door de auteur van dit artikel toegevoegd.

Ter verdere validatie van het hier gepresenteerde instrument wordt een grootschalige enquête opgezet onder Nederlandse managers. Tevens bestaan er plannen voor het uitvoeren van een aantal casestudies waarin het hier gepresenteerde instrument wordt gebruikt om in een praktijksituatie het is te beoordelen. In deze studies wordt niet alleen het UIS-instrument gebruikt, maar worden tevens extra vragen aan de respondenten gesteld om verklaringen voor het succes van het systeem vast te stellen. Hierbij kan aan twee soorten verklaringen worden gedacht. Allereerst kan een verklaring voor het succes in de specifieke situatie

van de casestudie worden gezocht. Een uitkomst kan dan zijn dat een in een filiaal-organisatie geïntroduceerd systeem alleen in de grotere filialen succesvol is. Een tweede soort verklaringen dat gezocht wordt, is meer algemeen van aard. Een uitkomst kan dan zijn dat de opleiding of de cognitieve stijl van de gebruiker van belang is voor de succeskans van het systeem. Het onderscheid tussen beide soorten verklaringen zal overigens niet altijd gemakkelijk te maken zijn. Beide verklaringen hebben gemeen dat zij helpen bij het vinden van de oplossing voor geconstateerde problemen. Zij geven aan op welke problemen de organisatie zijn aandacht moet richten.

Naast deze algemene validatie is onderzoek nodig naar de wijze waarop het gebruik van het UIS-instrument de gebruikersperceptie van het systeem beïnvloedt (zie ook Melone 1990). Met andere woorden: hoe verandert de wijze waarop gebruikers tegen een IS aankijken doordat zij een UIS-vragenlijst hebben ingevuld? Zeker indien men regelmatig de gebruikerstevredenheid met een systeem wenst te meten is inzicht in de invloed van de meting op de gebruiker van belang. Onderzoek zal moeten uitwijzen wat een optimale meetfrequentie is en of het gewenst is telkens dezelfde groep gebruikers voor de UIS-meting te gebruiken.

De gegevens uit het enquête-onderzoek en de casestudies zullen tevens gebruikt worden om gegevens te verzamelen die noodzakelijk zijn om het UIS-instrument te gebruiken voor benchmarking. Hiervoor zijn echter nog verdere gegevens nodig. Wie gebruik wenst te maken van het hier gepresenteerde instrument wordt derhalve verzocht contact op te nemen met de auteur van dit artikel.

6 Samenvatting en conclusies

In dit artikel is ingegaan op het belang van evaluatie van IS voor de praktijk. Evaluatie is van belang om een adequaat oordeel te kunnen geven over de verantwoordelijke personen, om te voorkomen dat slechte systemen in gebruik blijven en om input te verschaffen voor de betere ontwikkeling van betere systemen. Uit praktijkonderzoek blijkt dat ondernemingen die hun IS evalueren met behulp van gebruikerstevredenheid dit ook daadwerkelijk als positief ervaren. Ook wetenschappers kunnen niet meer om UIS heen: het is het meest gebruikte succes criterium in empirisch

onderzoek naar management support systemen.

Bij evaluatie van IS geniet het gebruik van een gevalideerd UIS-instrument als succes criterium de voorkeur. Evaluatie kan hiermee geschieden tegen aanvaardbare kosten, benchmarking is mogelijk en de richting van de oplossing van eventuele problemen wordt globaal aangewezen. Het enige reële alternatief is de mate van gebruik van het IS, maar dit criterium kent een aantal tekortkomingen. Zo kan het niet worden toegepast indien gebruik niet vrijwillig is, is meer gebruik niet altijd beter, is het slechts in beperkte mate mogelijk een ideale mate van gebruik vast te stellen dan wel gebruikscijfers onderling te vergelijken en lijkt het erop dat het gebruikscriterium alleen de zeer slechte van de overige systemen onderscheidt.

Van de beschikbare instrumenten is de hier gepresenteerde Nederlandse aanpassing van het door Doll en Torkzadeh instrument het bruikbaarst. Allereerst geniet dit instrument de voorkeur vanwege de algemene toepasbaarheid. De gestelde vragen zijn op ieder willekeurig systeem van toepassing en zijn ook door technisch niet onderlegde gebruikers te beantwoorden. Verder is het instrument gericht op gebruikerstevredenheid met een bepaald systeem en niet met de informatievoorziening in het algemeen. De beperkte omvang van het hier gepresenteerde instrument maakt praktische toepassing haalbaar; de voor het invullen benodigde tijd is beperkt. Hier staat tegenover dat de richting van de oplossing van de problemen slechts globaal wordt aangeduid en dat de vertaling van het instrument nog niet volledig gevalideerd is.

L I T E R A T U U R

- Abdel-Hamid, T. K. en S. E. Madnick, (1990, Fall). The elusive silver lining: how we fail to learn from software development failures. *Sloan Management Review*.
- Ackoff, R. L., (1967, december). Management misinformation systems. *Management Science* 14 (4), B:147-B:156.
- Alavi, M. en J. C. Henderson, (1981, november). An evolutionary approach for implementing a decision support system. *Management Science* 27 (11), pp. 1309-1323.
- American Accounting Association, (1972-1973). *Report of the committee on management information systems*. American Accounting Association.
- Bailey, J. E. en S. W. Pearson, (1983, mei). Development of a

- tool for analyzing and measuring computer user satisfaction. *Management Science* 29 (5).
- Barki, H. en S. Huff, (1990, mei). Implementing decision support systems: correlates of user satisfaction and system usage. *INFOR* (Canada) 28 (2), pp. 89-101.
- Barki, H. en S. L. Huff, (1985, december). Change, attitude to change, and decision support system success. *Information & Management* 9 (5), pp. 261-268.
- Baroudi, J. J., M. H. Olson en B. Ives, (1986, maart). An empirical study of the impact of user involvement on system usage and information satisfaction. *Communications of the ACM* 29 (3), pp. 232-238.
- Bollen, K. A., (1989). *Structural equations with latent variables*. Wiley series in probability and mathematical statistics. Applied probability and statistics section. John Wiley & Sons.
- Centraal Bureau voor de Statistiek, (1995, februari). *Maandstatistiek van de Binnenlandse Handel en Dienstverlening*.
- Conrath, D. W. en O. P. Mignen, (1990). What is being done to measure user satisfaction with EDP/MIS. *Information & Management* 19, pp. 7-19.
- Cox, J. F., R. W. Zmud en S. J. Clark, (1981). Auditing an MRP system. *Academy of Management Journal* 24 (2), pp. 386-402.
- DeBrabander, B. en G. Thiers, (1984, februari). Successful information system development in relation to situational factors which affect effective communications between MIS-users and EDP-specialists. *Management Science* 30 (2), pp. 137-155.
- Delen, G. en D. Rijsenbrij, (1990, januari). Kwaliteitsattributen van automatiseringsprojecten en informatiesystemen. *Informatie* 32 (1), pp. 46-55.
- Dickson, G. W. en J. C. Wetherbe, (1985). *The management of information systems*. McGraw-Hill.
- Doll, W. J. en G. Torkzadeh, (1988, juni). The measurement of end-user computing satisfaction. *MIS Quarterly* 12, pp. 259-274.
- Doll, W. J., W. Xia en G. Torkzadeh, (1994, december). A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS Quarterly*, pp. 453-461.
- Egten, C. A. V., (1994). *Kwaliteit van bestuurlijke informatie in een organisatie*. Ph. D. thesis, Vrije Universiteit Amsterdam.
- Gallagher, C. A., (1974, maart). Perceptions of the value of a management information system. *Academy of Management Journal* 17 (1), pp. 46-55.
- Gelderman, M., (1995a, augustus). *Factors affecting the success of management support systems: analysis and meta-analysis*. Paper presented at the IS/MAS forum of the Annual Conference of the American Accounting Association, Orlando, Florida.
- Gelderman, M., (1995b). *Translation and validation of the Doll and Torkzadeh user information satisfaction instrument*. Unpublished Working Paper, Vrije Universiteit Amsterdam.
- Ginzberg, M. J., (1978, augustus). Finding an adequate measure of OR/MS effectiveness. *Interfaces* 8 (4), pp. 59-62.
- Ginzberg, M. J., (1981, april). Early diagnosis of MIS implementation failure: promising results and unanswered questions. *Management Science* 27 (4), pp. 459-478.
- Ives, B. en M. H. Olson, (1984, mei). User involvement and MIS success: a review of research. *Management Science* 30 (5), pp. 586-603.
- Ives, B., M. H. Olson en J. J. Baroudi, (1983, oktober). The measurement of user information satisfaction. *Communications of the ACM* 26 (10), pp. 785-793.
- Jong, C. de, K. Huizingh, P. Oude Ophuis en B. Wierenga, (1994). *Kritische succesfactoren voor marketing decision support systemen*. Eburon, Delft.
- Jöreskog, K. G. en D. Sörbom, (1989). *LISREL7: a guide to the program and applications* (2nd ed.). SPSS Inc.
- Kim, K. K., (1990). Task characteristics, decentralization, and the success of hospital information systems. *Information & Management* 19, pp. 83-93.
- Lucas, H. C., (1978, Winter). The evolution of an information system: from keyman to every person. *Sloan Management Review*, pp. 39-52.
- MacIntosh, N. B., (1985). *The Social Software of Accounting and Information Systems*. John Wiley & Sons.
- McIntyre, S. H., (1982, januari). An experimental study of the impact of judgment based marketing models. *Management Science* 28 (1), pp.17-33.
- Melone, N. P., (1990, januari). A theoretical assessment of the user-satisfaction construct in information systems research. *Management Science* 36 (1), pp. 76-91.
- Nolan, R. L. en D. C. Croson, (1995). *Creative destruction: a six-stage process for transforming the organization*. Boston, Mass.: Harvard Business School Press.
- Olson, M. H. en B. Ives, (1981). User involvement in systems design: an empirical test of alternative approach. *Information & Management* 4 (4), pp. 183-195.
- Powers, R. F. en G. W. Dickson, (1973, Spring). Misproject management: Myths, opinions, and reality. *California Management Review* XV (3), pp. 147-156.
- Rijsenbrij, D., (1993, januari). Kwaliteit in Informatie. *Informatie* 35 (1), pp. 20-38.
- Rivard, E. en K. Kaiser, (1989, January 15). The benefit of quality IS. *Datamation* 35, pp. 53-58.
- Swanson, E. B., (1974, oktober). Management information systems: appreciation and involvement. *Management Science* 21 (2), pp.178-188.
- Vasarhelyi, M., (1977, Spring). Man-machine planning systems: A cognitive style examination of interactive decision making. *Journal of Accounting Research* 15 (1), pp. 138-153.
- Ven, A. H. V. d. en D. L. Ferry, (1980). *Measuring and*

NOTEN

1 Ik wil Tom Groot (Vrije Universiteit Amsterdam), Rien Hamers (Rabobank Nederland), Edu Spoor (Vrije Universiteit Amsterdam), Elbert de With (Vrije Universiteit Amsterdam) en Esther IJskes (Vrije Universiteit Amsterdam) en beide referenten bedanken voor commentaar op een eerdere versie van dit artikel.

2 Een gestandaardiseerd instrument is een vragenlijst die op een telkens identieke wijze wordt gebruikt. Op deze wijze loont het de moeite om in de ontwikkeling van goede vragen te investeren. Bovendien wordt het mogelijk om de resultaten van het gebruik van een dergelijke vragenlijst onderling te vergelijken. Gestandaardiseerde vragenlijsten zijn afkomstig uit de psychometrie, maar worden steeds meer op andere terreinen toegepast. Voorbeelden van toepassing in de bedrijfseconomie zijn het bekende werk (175 verwijzingen in de Social Science Citation Index van het moment van publicatie tot augustus 1995) van Van de Ven en Ferry (1980), die zeven jaar besteedden aan de ontwikkeling van een instrument voor het vaststellen van kenmerken van organisaties (de titel van het boek waarin dit project en de resultaten besproken worden luidt typerend genoeg *Measuring and assessing organizations*), ook bekend is het SERVQUAL-onderzoek dat zich richt op het meten van de kwaliteit van dienstverlening, terwijl Nolan en Croson (1995) verwijzen naar een instrument voor het meten van klanttevredenheid met een specifiek product.

3 Globaal gesproken zijn er twee richtingen te onderscheiden in empirisch onderzoek naar het succes van is: process en factor research. De eerste onderzoeksrichting richt zich op de ontwikkeling van het succes en de tweede probeert op enig moment in de tijd het succes te verklaren.

4 Tabel 1 bevat ook het hoofdje 'overige kenmerken'; het betreft hier onderzoek dat bij voorbeeld de winst (mede) als afhankelijke variabele gebruikt.

5 De gekozen benadering valt derhalve onder te brengen in de dynamische en de informatiedimensie uit het framework van Delen en Rijsenbrij (1990). Zie hierover ook Jong, Huizingh, Oude Ophuis en Wierenga (1994); met name hoofdstuk 6.

6 We kunnen dit ook zien als een uiting van een meer algemene trend. Nolan en Croson hebben het bij voorbeeld in hun boek *Creative Destruction* over 'a fundamental shift from an inward-focused "make-and-sell" orientation to an outward looking "sense-and-respond" orientation' (Nolan en Croson 1995, p. 30).

7 Dit sluit aan bij de eerder genoemde tendens om is meer vanuit de vraag- dan vanuit de aanbodzijde te bekijken. Bij

meting van de mate van gebruik wordt het impliciet als gegeven beschouwd dat er een systeem is, en wordt gegeven dit systeem maximaal gebruik nagestreefd ('make-and-sell' orientation die zeer wel gebaseerd kan zijn op het traditionele management science/operations research denken waarin mathematisch wordt 'aangetoond' dat toepassing van een techniek 'gewenst' is), terwijl bij meting van gebruikerstevredenheid waarde wordt toegekend aan het oordeel van de gebruiker (de afnemer) over het is.

8 Het is gebruikelijk om in dit geval de individuele vragen items te noemen en in plaats van het woord factor het woord schaal te gebruiken. De wegingsfactoren worden lading genoemd. Wanneer een instrument T -equivalent of parallel is, zijn alle ladingen gelijk en kunnen derhalve ongewogen worden opgeteld (Jöreskog en Sörbom 1989).

9 In feite wordt de op de individuele vragen toegepaste procedure herhaald. Er wordt wel gesproken van tweede-ordefactoren.

10 Dit blijkt ook uit het onderzoek van Conrath en Mignen, waar het merendeel van de organisaties die gebruikerstevredenheid meten, pas na 1983 begonnen is met deze metingen (Conrath en Mignen 1990, Table 5). Overigens is het met enige goede wil mogelijk een artikel van Gallagher (1974) in de *Academy of Management Journal* te beschouwen als een eerste serieuze poging tot het definiëren van *UIS*. Genoemde auteur streeft er echter uiteindelijk naar om een bedrag in geld vast te stellen dat de waarde van het systeem weergeeft.

11 De onderscheiden factoren zijn top management involvement, organizational competition with the EDP unit, priorities determination, charge-back method of payment for services, relationship with the EDP staff, communication with the EDP staff, technical competence of the EDP staff, attitude of the EDP staff, schedule of products and services, time required for new development, processing of change requests, vendor support, response/turnaround time, means of input/output with EDP center, convenience of access, accuracy, timeliness, precision, reliability, currency, completeness, format of output, language, volume of output, relevancy, error recovery, security of data, documentation, expectations, understanding of systems, perceived utility, confidence in the systems, feeling of participation, feeling of control, degree of training, job effects, organizational position of the EDP function, flexibility of systems en integration of systems.

12 Deze factor is dan ook door Ives et al. geëlimineerd.

13 Op deze wijze valt echter niet te achterhalen of een te lage gebruikersvriendelijkheid van een systeem ertoe heeft geleid dat de eindgebruiker het systeem niet zelf (direct) gebruikt.

14 Een factoranalyse bevestigt de veronderstelde structuur en de Cronbach's alpha's zijn respectievelijk 0.89, 0.91, 0.78, 0.85, 0.82 en 0.92)

15 De betrouwbaarheden (Cronbach's alpha van de

content-, accuracy-, format- en timeliness-schaal zijn goed (respectievelijk 0.88, 0.88, 0.86, 0.74 en voor de schaal als geheel 0.90). Alle factorladingen zijn significant; alle t -waarden liggen boven de 5. De goodness-of-fit index van het factormodel (vergelijkbaar met R^2 in een multiële regressie) ligt boven de 0.8, maar een χ^2 -test wijst uit dat de fit tussen het model en de waarnemingen onvoldoende is. Hierbij dient overigens wel opgemerkt te worden dat de eisen die in een confirmatieve factoranalyse aan de kwaliteit van het model worden gesteld veel hoger zijn dan de eisen die in een traditionele

vaststelling van de betrouwbaarheid worden gehanteerd. Reliabilites van de individuele items zijn respectievelijk 0.63, 0.65, 0.74, 0.64; 0.77, 0.91, 0.66, 0.40; 0.61, 0.57, 0.72, 0.53; 0.57, 0.45.

16 Mathematisch kan worden aangetoond dat wanneer ten minste drie indicatoren per factor worden gebruikt en gegeven een aantal randvoorwaarden de oplossing van het LISREL-model uniek is (Bollen 1989). Op deze manier wordt het mogelijk om multivariate modellen over het succes van is te testen.