

## Begroten van software als groepsproces : een alternatieve benadering

**Citation for published version (APA):**

Howard, M. S., Kusters, R. J., Heemstra, F. J., & Genuchten, van, M. J. I. M. (1992). Begroten van software als groepsproces : een alternatieve benadering. *Bedrijfskunde : Tijdschrift voor Modern Management*, 64(2), 143-153.

**Document status and date:**

Gepubliceerd: 01/01/1992

**Document Version:**

Uitgevers PDF, ook bekend als Version of Record

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

Ir. M. Howard, dr. R.J. Kusters, dr. ir. F.J. Heemstra en dr. ir. M. van Genuchten\*

## Begroten van software als groepsproces: een alternatieve benadering

### 1. Inleiding

Software is economisch gezien een belangrijk produkt geworden. De verkoop van software lag in 1989 rond de 65 miljard dollar (Shaw 1990). De kosten van 'in-house' software-ontwikkeling in de Verenigde Staten werden in dat jaar tussen de 150 en 200 miljard dollar geschat. Ondanks deze stormachtige ontwikkeling is een groot aantal aan software-ontwikkeling gerelateerde problemen niet opgelost. Een daarvan is dat software vaak later klaar is dan gepland. Reeds enige jaren lopend onderzoek op de TU Eindhoven is gericht op het vinden van benaderingen die verbeteringen kunnen aanbrengen in het begrotingsproces van software. Een van deze benaderingen is gebaseerd op het systematisch hanteren van groepsdiscussies.

In dit artikel zullen we de motivatie voor het volgen van een dergelijke benadering aangeven en enige gegevens presenteren die deze benadering ondersteunen. Deze gegevens zijn het resultaat van een serie experimenten die eveneens in dit artikel zullen worden beschreven.

In paragraaf 2 wordt dieper ingegaan op de

argumenten voor het hanteren van een groepsgewijze benadering voor het begroten van softwareprojecten. Verder worden de bestaande begrotingsmethoden geëvalueerd vanuit een groepsperspectief. In paragraaf 3 wordt een experiment beschreven waarmee deze argumenten worden ondersteund. Ten slotte worden in paragraaf 4 consequenties getrokken voor verder onderzoek.

### 2. Het begroten van softwareprojecten vanuit een nieuw perspectief

#### *Motivatie*

We zijn begonnen met een studie naar eerder onderzoek dat gericht is op het vergelijken van de kwaliteit van de beslissingen van individuen met die van groepen in verschillende omstandigheden (Shaw, 1981; Wallach e.a., 1962). De conclusie daarvan was dat het verschil afhangt van een aantal factoren, zoals het type beslissing, het type beslissingsproces en de kennis en ervaring van de betrokkenen. Het is dus niet zo dat een groepsbenadering in alle situaties 'beter' beslissingen oplevert. In deze paragraaf zullen we proberen aan te tonen dat het gebruik van groepen in de specifieke situatie van het begroten van software-ontwikkeling wel voordelen heeft. Bij de bespreking van deze voordelen zullen we uitgaan van drie belangrijke problemen die men vaak tegenkomt bij het begroten.

Het eerste probleem betreft de complexiteit van het begroten. De begroter moet rekening houden met de invloed van allerlei kos-

\* Ir. M. Howard, dr. R.J. Kusters en dr. ir. F.J. Heemstra zijn verbonden aan de vakgroep Bestuurlijke Informatiesystemen en Automatisering van de Faculteit Technische Bedrijfskunde van de Technische Universiteit Eindhoven respectievelijk als assistente in opleiding, universitair docent en universitair hoofddocent. Dr. ir. M. van Genuchten was voorheen werkzaam als assistent in opleiding bij dezelfde vakgroep en thans als adviseur bij Lighthouse Consultancy, een onderdeel van Philips.

tenbeïnvloedende factoren. Voorbeelden van deze factoren zijn de kwaliteit van het ontwikkelteam, de complexiteit en de vereiste betrouwbaarheid van de software en het effect van allerlei programmeerhulpmiddelen. Men kan weliswaar een groot aantal van deze factoren identificeren, maar de meerderheid ervan is lastig te definiëren, laat staan te kwantificeren. Voor een individu is het haast onmogelijk om bij het opstellen van een begroting de enorme hoeveelheid kwalitatieve informatie te verwerken en rekening te houden met alle relevante factoren en met hun onderlinge relaties. We vatten dit probleem samen onder de noemer 'informatie-overload'.

Als de literatuur vanuit deze invalshoek wordt bekeken, dan valt één resultaat onmiddellijk op. In taken die oordeelsvermogen en taxatie vereisen, worden betere resultaten geboekt door groepen dan door individuen (Moorhead en Griffin, 1989). Dit wordt veroorzaakt doordat aan elk van de leden van een groep een deel van het werk kan worden toegewezen. Bij deze taakverdeling kan men gebruik maken van de expertise die in de groep aanwezig is. Ieder groepslid is het best in staat op zijn eigen terrein de informatie te verwerken en in samengevatte vorm in de groepsdiscussie in te brengen. Het tweede probleem heeft te maken met de onzekerheid die een begroter ervaart. Er is weliswaar een ruime hoeveelheid informatie beschikbaar, maar deze informatie is vaak redundant of onbetrouwbaar en dus niet direct bruikbaar voor het maken van een begroting. Dit gevoel van onzekerheid bij een individu wordt versterkt wanneer hij wordt geconfronteerd met de povere resultaten van zijn eerdere pogingen op dit terrein. Het is daarom geen verrassing dat de begroter vaak advies zoekt bij anderen. Hij zal andere mensen raadplegen om een gevoel te krijgen voor de wijze waarop die het probleem zouden aanpakken of ter geruststelling zoals men op school kan zien bij een reactie

als: 'Ik keek alleen maar op zijn blaadje om te zien of ik het goed had!'

De literatuur bevestigt dat in complexe situaties waarbij veel factoren uit verschillende domeinen een rol spelen, een groepsgewijze aanpak vaak voordelen oplevert. Omdat veel factoren een rol spelen bij het begroten kan een persoon deze nooit integreren in een enkel mentaal model. Vaak zal men een soort 'tunnelvisie' ontwikkelen, waarbij de beschikbare informatie uiterst selectief wordt gefilterd om deze te kunnen hanteren (Sheridan, 1981). Het gevolg is dat informatie vaak van uit een enkel perspectief wordt bekeken. Er is geen garantie dat dit perspectief de belangrijkste informatie bevat. In een groep zal een dergelijk selectieve benadering van informatie niet zo gauw optreden. Vooroordelen worden zo voorkomen en er wordt niet alleen bevestigende informatie meegenomen (Einhorn en Hogarth, 1978; Schustack en Sternberg, 1981; Mynatt e.a., 1977; Wason en Johnson-Laird, 1972).

Een derde probleem betreft de geringe betrokkenheid van de betrokken partijen. Het succes of falen van een softwareproject wordt voor een groot deel beïnvloed door de mate van betrokkenheid van de deelnemers. Dit geldt natuurlijk in meer of mindere mate voor elk type taak. In versterkte mate geldt dit bij een taak als software-ontwikkeling omdat hier sprake is van een abstract en ongreepbaar eindresultaat. Een schatting die de ontwikkelaars niet reëel vinden zal hun motivatie negatief beïnvloeden. Men voelt zich niet medeverantwoordelijk voor een begroting die 'van bovenaf' wordt opgelegd. Dit kan resulteren in tijd- en budgetoverschrijdingen. Een ander punt is dat bij een project verschillende belangen een rol kunnen spelen. In een groepsdiscussie kunnen deze politieke en persoonlijke belangen, de zogenaamde 'hidden agenda's', expliciet aangevoerd worden, wat de motivatie positief kan beïnvloeden. Het gevolg hiervan is dat een

begroting eerder als een te halen doelstelling gezien wordt.

De literatuur bevestigt inderdaad dat begroten in groepsvorm een hogere motivatie en meer interesse in de uit te voeren taak tot gevolg zal hebben. Dit leidt tot een hogere produktiviteit en een positievere instelling als er problemen tijdens het ontwikkelproces optreden (Huber, 1980; Snizek en Henry, 1990).

In de praktijk zijn er mensen aan te wijzen die als begrotingsexpert worden beschouwd en er in slagen redelijk betrouwbare begrotingen te maken. Dit zou onze argumentatie voor groepsbegrotingen kunnen weerspreken. Deze deskundigen zijn echter in staat alleen in grote lijnen aan te geven welke stappen bij het opstellen van de begroting worden gevolgd, maar niet welke kennis bij deze stappen wordt gebruikt (Vicinanza, 1990; Howard, 1991). Hun deskundigheid is gebaseerd op intuïtie en is als zodanig moeilijk overdraagbaar. Ook wordt het motivatieprobleem van de ontwikkelaars niet opgelost.

De geciteerde literatuur geeft aan dat een groepsbenadering bij begroten voordelen heeft. Deze indruk wordt bevestigd door een enquête die uitgevoerd is onder de bezoekers van een congres<sup>1</sup> over begroten. De respondenten (29 van de circa 45 aanwezigen) gaven aan dat:

- ze doorgaans een begroting in een groep maken of het advies van anderen vragen. (ja: 23; nee: 3; geen antwoord: 4);
- ze van mening zijn dat een groep over het algemeen een betrouwbaardere schatting zal opleveren dan een individuele begroting (ja: 19; nee: 3; geen antwoord: 9)

Naast deze positieve indruk komt uit de literatuur ook een aantal probleempunten naar voren. Een voorbeeld is het zogenaamde 'risky shift'-phenomeen (Stoner, 1968 en Cartwright, 1973): de tendens van een groep zich te conformeren aan het gedrag van een

*opinion leader*. Bij het eventuele herontwerp van de begrotingstaak zal rekening moeten worden gehouden met deze en andere factoren. Voordat men overgaat tot groepsbesluitvorming en de nodige organisatorische maatregelen hiervoor gaat treffen, zal uiteraard het nut daarvan geverifieerd moeten worden. Het experiment dat in de volgende paragraaf wordt beschreven, is een eerste aanzet in de richting van een dergelijk verificatie

#### *Bespreking van de bestaande begrotingsmethoden*

Voor het experiment wordt besproken, zal eerst op de bestaande begrotingsmethoden worden ingegaan. Boehm (1981) onderscheidt een aantal methoden voor het begroten van een softwareproject. De belangrijkste zijn:

- de expertbegroting;
- een begroting op basis van analogie;
- een begroting met behulp van parametrische modellen

Allereerst valt op dat de groepsbegroting niet in deze lijst voorkomt. Verder blijkt dat geen van deze methoden expliciet ingaat op problemen met onzekerheid, informatie-*overload* en motivatie. Geen van de methoden gaat in op het aspect 'betrokkenheid van een groep' of moedigt een dergelijke betrokkenheid zelfs maar aan. Het wordt aan de individuele begroter overgelaten of en op welke wijze anderen bij het maken van de begroting worden betrokken.

Een uitzondering is het gebruik van parametrische modellen. Deze aanpak kan gezien worden als een voorzichtige poging de nadelige effecten van informatie-*overload* te verminderen. Een dergelijk model stelt relevante informatie in een geconcentreerde vorm ter beschikking. Het is echter niet aangetoond (Kemerer, 1987, Kusters e a., 1990) dat dergelijke modellen in de praktijk functioneren.

### 3. Het experiment

#### *De aanpak*

Met behulp van het experiment willen we onderzoeken wat de effecten van het begroten in een groep zijn. Hierboven is vermeld dat een parametrisch model de nadelige effecten van informatie-*overload* kan verminderen. Vandaar dat een dergelijk model ook in het experiment is opgenomen. Aan de deelnemers van het experiment is een schatting van de inspanning gevraagd voor het ontwikkelen van een informatiesysteem. Daarbij is een casebeschrijving gebruikt.

#### *De case*

De casebeschrijving is afgeleid van een definitiestudierapport. De studie is een aantal jaren geleden werkelijk uitgevoerd in een multinationale onderneming. De studie betreft de ontwikkeling van een systeem ter ondersteuning van het beheer van grote aantallen financiële waardepapieren (aandelen, obligaties, etc.). Kern van het systeem is een online transactieverwerkend systeem, gericht op de verwerking van mutaties in de posities die in zo'n vierduizend verschillende fondsen worden ingenomen. Verdere eisen waren een koppeling met het financiële informatiesysteem en het voldoen aan vrij strikte EDP-voorwaarden. Het systeem moest een bestaand, gedeeltelijk geautomatiseerd systeem vervangen.

De casebeschrijving bestond uit vijf pagina's tekst en gaf informatie over de systeemontwikkelomgeving en de organisatie waarbinnen het systeem moest functioneren. Daarnaast werd een opsomming gegeven van functionele, niet-functionele en organisatorische eisen. We gebruikten een iets aangepaste versie van het originele rapport. Op basis van het originele definitiestudierapport is het werkelijke systeem ontwikkeld. Al snel na de start van het project traden er problemen op. Het bleek noodzakelijk de planning diverse malen aan te passen, met

als uiteindelijk gevolg een grote overschrijding van de origineel geplande inspanning (tabel 1).

*Tabel 1* Geschatte en gerealiseerde inspanning van het originele project

	Inspanning (in mensmaanden)	
	Plan	Realisatie
Functioneel ontwerp	32	36
Technisch ontwerp	32	52
Programmeren en testen	36	58
Acceptatietest	12	86
Conversie en implementatie	0	3
<b>Totaal</b>	<b>112</b>	<b>235</b>

Na afloop van het project vond een evaluatie plaats, gericht op het vinden van de oorzaken van deze overschrijdingen. Door gebruikers en ontwikkelaars werden vier factoren geïdentificeerd die een versturende invloed hadden op de uitvoering van het project. Deze risicofactoren waren:

- Gebrek aan tijd en belangstelling van de toekomstige gebruikers. De materiekennis die essentieel was voor het ontwerp van het systeem was alleen bekend bij een gebruiker. Dit was een afdelingshoofd die niet in staat was voldoende aandacht aan het project te besteden.
- Er waren verschillende afdelingen betrokken bij het project. Dit resulteerde in verschillende, vaak conflicterende gebruikerseisen. Slechts een van deze afdelingen was afhankelijk van het systeem. De andere afdelingen hadden geen belang bij het efficiënt oplossen van optreden problemen.
- Onbekendheid met de ontwikkelomgeving. De ontwikkelorganisatie gebruikte voor de eerste keer een vierde-generatietaal voor de ontwikkeling van een omvangrijk systeem.
- Gebrek aan materiekennis bij de ontwikkelaars veroorzaakte een onderschatting

van de omvang en complexiteit van het systeem

Bij de evaluatie bleek dat de gebruikers het project ondanks de tijd- en kostenoverschrijdingen een succes vonden. Het systeem voldeed uitstekend aan de verwachtingen van de gebruikers en is momenteel operationeel.

#### *De deelnemers*

Het experiment is een aantal malen uitgevoerd als onderdeel van een cursus post-academisch onderwijs 'Begroten en beheersen van software-ontwikkeling'. De deelnemers van deze cursussen (24 voor experiment 1 en 22 voor experiment 2) werd gevraagd de inspanning voor het project te begroten. Ze hadden een gemiddelde werkervaring van acht jaar. Vijfendertig van hen hadden ervaring met het begroten van software-ontwikkelprojecten. Allen waren voldoende bekend met het proces van software-ontwikkeling. Gegeven dat ze allen vrijwillig voor de cursus hadden gekozen, mag worden verondersteld dat in principe de bereidheid bestond om serieus aan het experiment mee te werken. Dit werd later bevestigd aan de hand van de cursusevaluatie.

#### *De procedure*

In het experiment werd aan de deelnemers vijf schattingen voor het project gevraagd. Het experiment duurde een halve dag. De procedure verschilde bij de twee experimenten enigszins. Tijdens het eerste experiment werden de volgende stappen gevolgd:

1 *Bestudering van het casemateriaal* Elke deelnemer beschikte over het definitiestudierapport. Na verloop van tijd kreeg men de gelegenheid vragen te stellen aan een van de onderzoekers. Deze was bekend met het project en speelde de rol van gebruiker. De vragen betroffen voornamelijk de organisatorische *setting* van het systeem en de taken

die het systeem moest gaan ondersteunen. Deze ronde was plenair, zodat alle deelnemers dezelfde informatie kregen.

2 *De eerste schatting*. De deelnemers werd gevraagd een schatting voor inspanning en doorlooptijd te maken. Deze schatting was gebaseerd op hun eigen ervaring en op het casemateriaal. Ook werd hen gevraagd de voornaamste risicofactoren te identificeren en de mate van vertrouwen in hun eigen schatting weer te geven op een vijfpuntschaal.

3. *De modelschatting* Vervolgens kreeg elke deelnemer de beschikking over een geautomatiseerde versie van een parametrisch model (Estimacs). Iedere deelnemer werd verzocht de vragen die het tool stelde te beantwoorden waarna het tool een begroting produceerde.

4 *De tweede schatting* Deze stap was hetzelfde voor beide experimenten. Aan iedereen werd een tweede individuele schatting gevraagd. Deze was naast de eigen ervaring en de casebeschrijving verder gebaseerd op de informatie die in de vorige stap werd verkregen. De vragen betreffende risicofactoren en mate van vertrouwen werden ook hier herhaald.

5 *De groepsschatting* De deelnemers werden verdeeld in zes groepen van ieder vier personen. Groepen van twee en drie werden vermeden. Bij een groep van drie is de kans op een twee tegen een deadlock te groot, terwijl in een groep van twee personen de kans op onderling ondersteunend gedrag te groot is (Davis, 1964; Hare, 1981). Elke groep maakte een gezamenlijke schatting van inspanning en doorlooptijd en identificeerde gezamenlijk de belangrijkste risicofactoren. Ook hier werd om de mate van vertrouwen in het eindresultaat van de groep gevraagd.

6. *De derde schatting* Ten slotte werd aan alle deelnemers gevraagd een laatste individuele schatting te geven op basis van alle informatie die gedurende het experiment

verkregen was. Ook hier werd weer gevraagd om de risicofactoren en om een indicatie van de mate van vertrouwen.

De resultaten van elke schatting werden direct in een vragenlijst genoteerd. Het eerste experiment verliep als zojuist geschetst. Tijdens het tweede experiment werden de stappen 3 en 5 verwisseld. Eerst werd de groepsschatting gemaakt en pas in tweede instantie werd gebruik gemaakt van een schattings-tool. Ook moet worden opgemerkt dat, gegeven het totaal aantal deelnemers (22), werd afgeweken van de groepsomvang van vier. Er werden twee groepen van drie deelnemers elk gevormd.

De opzet van het experiment stelde ons in staat de invloed van de twee onafhankelijke variabelen 'groepseffect' en 'tool-effect' te meten. De variabelen werden één voor één toegevoegd en elk effect werd apart geregistreerd. De mogelijkheid dat de volgorde waarin de variabelen werden toegevoegd invloed heeft, werd opgevangen door het omdraaien van deze volgorde bij de twee experimenten. Deze opzet houdt echter geen rekening met het mogelijk effect van het herhalen van de schattingen. Dit zou men kunnen opvangen door middel van het hanteren van controlegroepen. Gezien de benodigde grote aantallen deelnemers is hier niet voor gekozen.

#### *Het meten van het effect van groepen en tools*

De meest voor de hand liggende maat van een dergelijk effect is de schatting zelf. In dit geval kon hiervan niet gebruik worden gemaakt. Alle deelnemers kwamen immers uit verschillende omgevingen. Dit kan op zich al voor een verschil in de begroting zorgen. De deelnemers zijn niet 'gecalibreerd' of 'gestandaardiseerd' op één enkele omgeving. Ook het gebruik van de variantie als maat voor een mogelijk effect is twijfelachtig, om-

dat het gebruik hiervan uitgaat van de veronderstelling dat er een gemeenschappelijk referentiepunt is ten opzichte waarvan de spreiding wordt bekeken. Een dergelijk referentiepunt is in dit geval niet waarschijnlijk, vandaar dat voor andere maten is gekozen. Een van deze maten was de mate van vertrouwen die werd uitgesproken in de schattingen, en dan met name de veranderingen die optraden na de tweede en de derde schatting. Een andere maat was of deelnemers bij hun oorspronkelijke begroting bleven, of daar veranderingen in aanbrachten. Een toename in vertrouwen kan ook worden veroorzaakt door een stijgende vertrouwdheid met het probleem, omdat het tijdens het experiment een aantal malen bekeken is. Vooral als de begroting tijdens het experiment niet verandert kan dit het geval zijn. Tijdens het experiment wordt dan met name die informatie verwerkt die de oorspronkelijke mening bevestigt. Als echter een stijging in vertrouwen samenvalt met een verandering in de begroting, dan is het redelijk te veronderstellen dat andere factoren, zoals het effect van een groepsdiscussie of het effect van het hanteren van een *tool*, een rol spelen.

De uitvoering van het oorspronkelijke project werd verstoord door een aantal factoren. De deelnemers moesten deze risicofactoren tijdens het maken van de begroting identificeren. Daarom is als derde maat in deze studie het aantal geïdentificeerde relevante risicofactoren besproken. Deze maat werd gezien als een indicatie van de effectiviteit waarmee de beschikbare informatie werd verwerkt.

#### *Resultaten*

De deelnemers hebben vijf maal een begroting van de inspanning gemaakt. Een samenvatting van deze resultaten is in tabel 2 te vinden. Al eerder is aangegeven dat wij ui-

terst voorzichtig zijn geweest om op basis van deze resultaten conclusies te trekken. Wel geven de cijfers een indicatie van de spreiding van de begrotingen. Het aantal begrotingen in de eerste kolom vereist toelichting. Het aantal deelnemers was 24. Eén deelnemer gaf geen modelschatting en zes personen gaven geen derde begroting; zij volgden de instructies niet correct op. Bij het tweede experiment traden deze problemen niet op.

De deelnemers gaven ook de mate van vertrouwen aan die zij stelden in hun begrotingen. Ze maakten daarbij gebruik van een schaal die varieerde van 1 (geen vertrouwen) tot 5 (absoluut vertrouwen). Deze resultaten staan in tabel 3 vermeld. Er ontbreken twee waarden omdat bij het eerste experiment deze vraag bij de modelschatting niet werd gesteld.

Het aantal deelnemers dat gaandeweg het

experiment de begroting wijzigde, geeft een indicatie van het effect van de groep en het effect van het *tool*. Deze resultaten worden in tabel 4 weergegeven. Bij de presentatie van deze gegevens is onderscheid gemaakt tussen het totaal aantal wijzigingen en het deel ervan wat in de richting van de groep (model) wijzigde.

Het aantal geïdentificeerde relevante risicofactoren is in tabel 5 weergegeven. Het aantal door individuen tijdens de eerste begroting benoemde risicofactoren wordt gepresenteerd. Daarnaast wordt het resultaat van de groepen gepresenteerd. Bij het bepalen van dit resultaat is de informatie van de afzonderlijke leden van de groep 'gepoold'. Hierdoor is gebruik gemaakt van de informatie die een groep in principe ter beschikking had, niet van de keuzen die de groep uiteindelijk gemaakt heeft.

Tabel 2. Resultaten van het experiment

Schatting	Experiment 1		Experiment 2	
	Inspanning (in mensmaanden)		Inspanning (in mensmaanden)	
	Gemiddelde	$\sigma$	Gemiddelde	$\sigma$
Eerste (N = 24)	65	77	48	43
Model (N = 23)	74	42	56	30
Tweede (N = 24)	55	40	62	39
Groep (N = 6)	57	30	82	62
Derde (N = 18)	43	19	67	43

Tabel 3. De mate van vertrouwen in de begrotingen

Schatting	Experiment 1		Experiment 2	
	Mate van vertrouwen		Mate van vertrouwen	
	Gemiddelde	$\sigma$	Gemiddelde	$\sigma$
Eerste (N = 24)	2,0	0,8	1,9	0,8
Model (N = 23)	—	—	2,2	0,8
Tweede (N = 24)	2,1	0,8	2,4	0,8
Groep (N = 6)	2,3	0,9	2,0	0,9
Derde (N = 18)	2,6	0,8	2,6	0,7



Tabel 4 Het aantal deelnemers dat tijdens het experiment de begroting wijzigde

	Experiment 1		Experiment 2		Totaal	
	aantal	percentage	aantal	percentage	aantal	percentage
Na model:						
- aantal wijzigingen	14	58	7	31	21	45
- aantal wijzigingen in de richting van het model	13	54	7	31	20	43
Na groepsessie:						
- aantal wijzigingen	18	100	18	81	36	90
- aantal wijzigingen in de richting van de groep	17	94	16	72	33	82

*Analyse van de resultaten*

De analyse bestaat uit drie delen:

- de kwaliteit van het casemateriaal;
- het effect van het model;
- het effect van de groepsessie

*De kwaliteit van het casemateriaal*

De meerderheid van de deelnemers (35 van de 46) gaf aan dat de beschikbare informatie onvoldoende was om er een begroting op te baseren. Een aantal deelnemers gaf aan dat zij nooit een project op basis van dergelijk specificaties zouden starten. Zij zouden eerst een nader vooronderzoek op 'fixed-price'-basis uitvoeren. Extra informatie werd gewenst, namelijk:

- een datamodel;
- betere functionele eisen;
- meer informatie over de ontwikkelomgeving;
- meer informatie over de toekomstige gebruikers;
- meer informatie over de wijze waarop het systeem gebruikt zal gaan worden

Het originele project is begroot en gestart op basis van een definitiestudierapport waar in geen datamodel beschikbaar was; alle beschikbare informatie over functionele eisen is in het casemateriaal weergegeven. Op deze punten beschikten de deelnemers over dezelfde informatie als de uitvoerders van het originele project. Het gebrek aan informatie betreffende de ontwikkel- en gebruikersomgevingen is een direct gevolg van de

Tabel 5 Het aantal geïdentificeerde relevante risicofactoren

Aantal geïdentificeerde risico-factoren	Percentage deelnemers na eerste schatting		Percentage groepen (beschikbare informatie)	
	experiment 1 (N = 24)	experiment 2 (N = 22)	experiment 1 (N = 6)	experiment 2 (N = 6)
0	13	4	0	0
1	20	28	0	0
2	42	50	0	0
3	17	18	17	33
4	8	0	83	67

opzet van het experiment. In werkelijkheid beschikten de uitvoerders van het oorspronkelijk project over meer informatie op deze punten, aangezien zij zelf uit die omgeving afkomstig waren.

Een vergelijking tussen de schattingen die tijdens het experiment zijn gemaakt en het werkelijke resultaat laat grote verschillen zien. Slechts acht deelnemers maken in eerste instantie een begroting van meer dan 100 mensmaanden. De onbekendheid met de materie van de verschillende deelnemers en de complexiteit van de applicatie zijn waarschijnlijk de oorzaak van de slechte kwaliteit van de schattingen. De onderschatting van het originele project (een schatting van 112 en een realisatie van 235 mensmaanden) geeft aan dat ook voor de professionals die de oorspronkelijke begroting opstelden dit geen eenvoudige taak was.

#### *De invloed van het model*

Het effect van een begrotingsmodel werd bekeken door een vergelijking van de individuele begrotingen voor en na het modelgebruik. Al eerder is gesteld dat niet veel waarde kan worden toegekend aan het gemiddelde van de begrotingen vanwege de verschillen in achtergrond tussen de deelnemers. Ook het begrotingsmodel kon niet worden gecalibreerd (Ceulenaere e.a., 1987). Daarom is hier alleen naar wijzigingen in de begroting gekeken en naar de richting van een dergelijke wijziging. Uit de resultaten volgt dat 21 van de 46 deelnemers een wijziging aanbrachten na gebruik van het model. Deze wijzigingen waren op één na in de richting die het model aangaf. De resultaten waren vergelijkbaar voor beide experimenten. Dit geeft aan dat de deelnemers inderdaad door het model werden beïnvloed. Geen van de deelnemers nam de schatting van het model over. Klaarblijkelijk werd de modelschatting als een tweede mening en niet als richtinggevend ervaren. De mate van vertrouwen in de begroting werd door het gebruik

van het model nauwelijks beïnvloed. Tijdens beide experimenten was weliswaar een stijging te zien, maar deze was te klein om een significant effect te suggereren.

#### *De invloed van de groepsdiscussie*

Het verschil tussen de individuele begroting voor en na de groepsdiscussie is weer als indicatie van het groepseffect genomen. Er was een duidelijk effect waar te nemen. Negentig procent van de deelnemers veranderde de begroting na afloop van de groepsdiscussie. Op drie na waren al deze veranderingen in de richting van de groepsbegroting. De resultaten waren voor beide experimenten min of meer hetzelfde. Een andere indicatie van het groepseffect werd verkregen door te kijken naar de mate van vertrouwen. Het is niet verrassend dat deze gedurende de tijd stijgt. Het is echter wel interessant dat in beide experimenten een stijging in de toename kon worden waargenomen na de groepsdiscussie terwijl de volgorde van activiteiten in de twee experimenten verschillend was. De stijging in vertrouwen hing samen met een verandering in de begroting en is dus niet alleen toe te schrijven aan een verder ingraven in een van te voren ingenomen positie. Men kan aannemen dat de groepsdiscussie en niet het verstrijken van de tijd verantwoordelijk is voor het waargenomen effect.

Verder is gekeken naar het aantal relevante risicofactoren dat door de individuen en door de groep is gevonden. Als de beschikbare informatie van de leden van een groep samen genomen werd, waren alle groepen, op drie na, in staat de vier meest relevante risicofactoren te benoemen. Bij de overige drie groepen was dit altijd nog drie van de vier factoren. Hier tegenover staat dat de individuen beduidend minder presteerden met een gemiddelde van twee geïdentificeerde factoren. Dit is een sterke indicatie dat groepen in staat zijn meer relevante informatie te genereren.

De groepen beschikten over voldoende informatie maar slaagden er niet in deze op voldoende wijze te interpreteren. Gedurende het tweede experiment werd aan de groepen gevraagd de belangrijkste risicofactoren te identificeren. De resultaten waren:

- een groep identificeerde één factor;
- vier groepen identificeerden twee factoren;
- een groep identificeerde drie factoren

Alhoewel de informatie in principe binnen de groep beschikbaar was, werd er geen goed gebruik van gemaakt. We concluderen hieruit dat het niet voldoende is om een aantal mensen samen te zetten. Wil men de voordelen van een groepsgewijze benadering krijgen, dan zal de groep een methode moeten hanteren die hen in staat stelt conflicterende meningen in kaart te brengen en daar op adequate wijze gebruik van te maken.

#### 4. Implicaties voor verder onderzoek

In dit artikel is een aantal argumenten, deels theoretisch en deels gebaseerd op praktische ervaringen, gepresenteerd. Een eerste conclusie is dat een groepsgewijze benadering bij het begroten van software-ontwikkeling veelbelovend lijkt. Theoretische argumenten geven aan dat zowel het verzamelen van relevante informatie als het verwerken ervan in een groep betere resultaten oplevert. Ook is het mogelijk een betere motivatie bij de betrokkenen te bewerkstelligen. De experimenten bevestigen de theorie dat groepen beter in staat zijn relevante informatie te verzamelen en, gegeven de hogere mate van vertrouwen die uitgesproken werd, een grotere betrokkenheid van de projectuitvoerders te realiseren.

Anderzijds is een aantal problemen geconstateerd. Het is in de literatuur al bekend dat de uitvoering van een taak binnen een groep bepaalde nadelen kan hebben, zoals conformeren aan een *opinion leader*. Ook tij-

dens de experimenten bleek dat het weliswaar goed mogelijk is de relevante informatie te verzamelen binnen een groep, maar dat het veel moeilijker is deze informatie ook op een adequate wijze te hanteren. Bij het ontwerp van de groepstaak moet daarmee rekening gehouden worden.

Het gebruik van groepen is een potentiële bijdrage. Wil men echter optimaal gebruik maken van deze voordelen, dan zal zowel de wijze waarop de groep samenwerkt als de ondersteuning die aan de groep wordt gegeven van belang zijn. Kijkend naar de wijze van samenwerking, zien we dat er enige vorm van opleiding of organisatie moet plaatsvinden. Het op willekeurige wijze bij elkaar zetten van mensen op zich levert onvoldoende voordeel op. Verzamelen en verwerken van informatie binnen een groep moet op zorgvuldige wijze plaatsvinden. Als de groepsinteractie op enige wijze verstoord wordt, is de kans groot dat de resultaten afwijken van wat mogelijk is. Ook dit komt overeen met wat in paragraaf 2 al op basis van de literatuur gesignaleerd was.

Ook de beschikbaarheid van ondersteunende hulpmiddelen kan van invloed zijn op het eindresultaat van een groepsproces. Men kan hierbij denken aan groepsbeslissingsondersteunende systemen (GDSS). Dergelijke systemen ondersteunen een systematische voortgang van het interactieproces. Ook geven ze aan iedereen in de groep de mogelijkheid om anoniem meningen en inzichten te presenteren. Tenslotte geven deze *tools* de mogelijkheid tot het parallel verwerken van informatie, iets wat in een groep tot veel tijdverspilling leidt (Dennis e a., 1988). Naast deze algemene ondersteunende hulpmiddelen kan men denken aan specifiek op dit probleem toegespitste *tools*. Bijvoorbeeld *tools* die informatie uit het verleden presenteren, 'what-if' analyses doorvoeren of een of meerdere begrotingsmodellen presenteren. Verder onderzoek op dit terrein zal zich richten op de inrichting van het groepsproces en

op de wijze waarop dit proces ondersteund kan worden. In eerste instantie zullen we ons daarbij richten op een deeltaak binnen het begroten, namelijk de risicoanalyse, en op een deel van de mogelijk ondersteunende hulpmiddelen, namelijk de generieke GDSS-tools. Er is een aantal experimenten gepland die binnen dit afgebakende terrein moeten leiden tot een opzet van een groepsbenadering.

**Noot**

- 1 ESCOM Meeting, Noordwijkerhout Nederland, juni 1991

**Literatuur**

- Boehm, B.W., *Software engineering economics*, Prentice Hall, 1981.
- Cartwright, D., 'The case of research on the risky shift', *American Psychologist*, march 1973, pp 222-231.
- Ceulenaere, A.M.E., M.J.I.M. van Genuchten en F.J. Heemstra, 'Calibrating a software cost model', *Information and Software Technology*, dec. 1987.
- Davis, J.H., *Group performance*, Addison-Wesley, Reading, Mass., 1964.
- Dennis, A.R., J.F. George, L.M. Jessup, J.F. Nunamaker en D.R. Vogel, 'Information technology to support electronic meetings', *MIS Quarterly*, vol 12, dec 1988, pp. 591-624.
- Einhorn, H.J., en R.M. Hogarth, 'Confidence in Judgment: Persistence of the illusion of validity', *Psychological Review*, vol. 85, 1978, pp 395-416.
- Hare, A.P., 'Group size', *American Behavioral Scientist*, vol 24, 1981, pp. 695-708
- Huber, G.P., *Managerial Decision Making*, Foreman, Glenview, Ill., 1980, pp. 90-115.
- Kemerer, C.F., 'An empirical validation of software cost estimation models', *Communications of the ACM*, vol 30, nr 5, 1987
- Kusters, R.J., F.J. Heemstra, en M.J.I.M. van Genuchten, 'Are software cost estimation models accurate?', *Information Technology and Software*, vol 32, nr 3, 1990.
- Moorhead en Griffin, *Organizational Behaviour*, Houghton Mifflin Cy., Boston, 1989
- Mynatt, C.R., M.E. Doherty en R.D. Tweney, 'Confirmation bias in a simulated environment: An experimental study of scientific inference', *Quarterly Journal of Experimental Psychology*, vol. 29, 1977, pp 85-95
- Schustack, M.W., en R.J. Sternberg, 'Evaluation of evidence in causal inference', *Journal of Experimental Psychology: General*, vol. 110, 1981, pp. 101-120.
- Shaw, M., *Group Dynamics: The Psychology of Small Group Behaviour*, 3rd ed New York, 1981, pp 57-68.
- Shaw, M., 'Prospects for an engineering discipline of software', *IEEE Software*, nov. 1990
- Sheridan, T., 'Understanding human error of aiding human diagnostic behaviour in nuclear power plants', In: J. Rasmussen en W. Rouse (red.), *Human Detection and diagnosis of system failures*. New York, Plenum Press, 1981
- Sniezek, J.A., en R.A. Henry, 'Revision, weighting and commitment in consensus group judgment', *Organizational Behaviour and Human Decision Processes*, vol 45, 1990, pp 66-84.
- Stoner, J.A.F., 'Risky and cautious shift in group decisions: the influence of widely held values', *Journal of Experimental Social Psychology*, vol. 4, 1986, pp. 442-459.
- Vicinanza, S., M.J. Prietula en T. Mukhopadhyay, 'Case-based reasoning in software cost estimation', In: J.I. DeGross, M. Alavi en H. Opeiland (red.), *Proceedings of 11th international conference on information systems*, ACM, 1990
- Wallach, M.A., N. Kogan en D.J. Bem, 'Group Influence on Individual Risk Taking', *Journal of Abnormal and Social Psychology*, pp. 75-86, 1962.
- Wason, P.C., en P.N. Johnson-Laird, *Psychology of reasoning: Structure and content*, Batsford, London, 1972