

Abstract

Forecasting Recessions: A Survey

*This article reviews several aspects about forecasting recessions. A first issue concerns the definition and dating of recessions. Secondly, the quantitative estimation and evaluation of recession forecasts is examined. Thirdly, popular recession indicators are discussed, viz. growth predictions obtained from large macroeconomic models or surveys, leading indicators and financial market information. The last section concludes and offers perspectives for further research.*



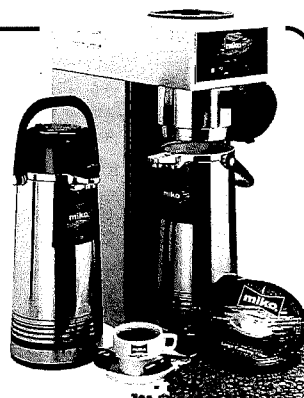
Office Coffee Service,  
een dienstverlening waar wij thuis in zijn!  
Meer dan 9.000 klanten zijn daar  
het bewijs van!

Onze succesformule:

Dit toestel volledig kosteloos bij U,  
inclusief alle nazicht en onderhoud.  
U betaalt **enkel** de koffie en de  
bijproducten!

Wenst U meer inlichtingen of  
een apparaat op proef, aarzel dan niet  
ons **vrijblijvend** te contacteren:

**M.C.S.** N.V. Hermesstraat 22  
B-1930 Zaventem



Tel. (02) 714.10.90  
Fax (02) 714.10.99

**200 jaar meesterschap in koffie**

Filip Lievens \*

Etienne Van Keer \*\*

## Assessment centers in België: de resultaten van een onderzoek naar de voorspellende kracht en de billijkheid

Trefwoorden: personeelsselectie; assessment centers; validiteit

*In België kennen assessment centers een steile groei. Ondanks deze populariteit is de waarde van assessment centers als selectie- en ontwikkelingsinstrument in België nog niet getoetst. Daarom onderzoekt deze studie de predictieve validiteit en de billijkheid van een assessment center van een grote bank. De steekproef bestaat uit 252 middle level managers, waaronder 201 mannen en 51 vrouwen (gemiddelde leeftijd = 38 jaar). De resultaten tonen dat dit assessment center aan de gestelde verwachtingen beantwoordt. Het instrument is voldoende valide om predicties toe te laten naar succes in hogere managementfuncties. Ook met betrekking tot billijkheid scoort dit assessment center uitstekend. Zo worden nagenoeg geen verschillen tussen beoordelingen van mannen en vrouwen gevonden. Ten slotte zijn er ook bijna geen verschillen in de beoordelingen van Walen en Vlamingen. Mogelijke verklaringen voor deze goede resultaten qua voorspellende validiteit en billijkheid liggen in de kwalitatief goede opbouw van het assessment center.*

\* Universiteit Gent, Vakgroep Personeelsbeleid en Arbeids- en Organisationspsychologie. Postdoctoraal onderzoeker van het Fonds voor Wetenschappelijk Onderzoek - Vlaanderen (F.W.O.).

\*\* SHL Belgium

## Inleiding

Assessment centers (hierna AC's) kennen een hoge vlucht. Uit een recente enquête blijkt dat 25% van de grote (met meer dan 200 werknemers) bedrijven in Vlaanderen regelmatig de AC-methode toepast in het kader van selectie en/of ontwikkeling (Lievens en Goemaere, 1999, blz. 56). Hoewel er vele definities over AC's de ronde doen, zou men een AC kunnen omschrijven als *"a standardized evaluation of behavior based on multiple inputs. Multiple trained observers and techniques are used. Judgments about behaviors are made, in major part, from specifically developed assessment simulations. These judgments are pooled in a meeting among the assessors or by a statistical integration process"* (Task Force on Assessment Center Guidelines, 1989, blz. 460). Deze definitie toont dat, naast assessoren (d.i. beoordelaars) en gedragskenmerken (d.i. observeerbare kenmerken waarop de kandidaten gescoord worden, zoals planning en organisatie, besluitvaardigheid, initiatief, stressbestendigheid, mondelinge communicatie en creativiteit), situatietests een derde typisch kenmerk van AC's vormen. Situatietests weerspiegelen belangrijke activiteiten en situaties uit de functie die kandidaten ambiëren. Aan deelnemers wordt bijvoorbeeld gevraagd om een complex en in een bedrijfscontext ingebed probleem te analyseren ("case-analyse") en nadien de resultaten hiervan mondeling te presenteren ("presentatie"). In andere situatietests ordenen en beantwoorden de deelnemers de ingekomen post van een zieke collega ("postbak", "in-basket" of "in-tray") of bereiken zij in groep een consensus over het toekomstige budget van een gefingeerde afdeling ("groepsdiscussie").

Het ontstaan van AC's gaat terug tot de Tweede Wereldoorlog. De methode is overgewaaid uit respectievelijk het Duitse, Engelse en Amerikaanse leger, maar American Telephone and Telegraph (AT&T) was de eerste grote onderneming die AC's op ruime schaal toepaste. Pas het laatste decennium hebben AC's zich als een belangrijke techniek binnen human resource management (HRM) in België geprofileerd. Toepassingen van AC's situeren zich ook niet enkel meer in de industriële sector maar eveneens in non-profitorganisaties.

De steile groei van AC's is onder meer een gevolg van de commerciële gedrevenheid waarmee AC's "verkocht" worden aan organisaties. Een argument dat hiervoor dikwijls gebruikt wordt, is dat AC's het mogelijk maken om valide voorspellingen te doen over de sterkten en zwakten van (toekomstig) personeel. Een ander verkoopargument is dat AC's

billijke instrumenten zijn. Dit betekent dat zowel mannen als vrouwen in een AC en in de AC-oefeningen gelijke kansen zouden hebben om te tonen wat ze kunnen. Voorzover wij weten, zijn er in België nog geen studies over AC's gepubliceerd om deze twee veronderstellingen te toetsen (Lievens en Vander Meeren, 1998, blz. 25). Om die lacune op te vangen, brengen wij in dit artikel verslag uit van een onderzoek naar de voorspellende validiteit en de billijkheid van het AC van de Bank Brussel Lambert (BBL). Bij wijze van achtergrond geven wij in eerste instantie een overzicht van het voorgaand internationaal AC-onderzoek in verband met deze twee invalshoeken.

### 1. Literatuurbevindingen over de voorspellende validiteit van assessment centers

Voorspellende of predictieve validiteit betreft de mate waarin een techniek, in dit geval de AC-methode, juiste voorspellingen mogelijk maakt over toekomstig werkgedrag. Het recentste grootschalige onderzoek naar de voorspellende kracht van AC's werd uitgevoerd door Gaugler, Rosenthal, Thornton en Bentson (1987, blz. 493). Hun meta-analyse vatte op een statistische manier de resultaten van 50 onafhankelijke studies samen. Hieruit bleek dat AC's een gemiddelde voorspellende validiteit van 0,37 behaalden. Ondanks het feit dat AC's over het algemeen als valide worden gezien, was de grote variatie in de gevonden validiteitscoëfficiënten frappant (Gaugler et al., 1987, blz. 493). Of om het eerder in statistische termen uit te drukken, de standaarddeviatie was relatief groot en bedroeg 0,12. Hieruit volgt dat men 90% zekerheid heeft dat een AC een predictieve validiteit tussen 0,21 en 0,53 bezit. Gaugler et al. (1987, blz. 493) identificeerden in hun onderzoek ook een aantal factoren die de validiteit gevoelig verhoogden. Zo bleken AC's een hoger voorspellend vermogen te bezitten als verschillende assessment-technieken werden gebruikt, als er naast managers ook psychologen als assessor optraden, en als er gebruik werd gemaakt van beoordelingen door mede-assessees. Ook was de hoogte van de validiteit afhankelijk van de aard van het criterium: voor prestatiebeoordelingen bedroeg de validiteit 0,36, voor potentieelinschattingen was dat 0,53.

Sinds deze meta-analyse zijn er een dozijn buitenlandse studies verricht naar de predictieve waarde van AC's (bijvoorbeeld Chan, 1996, blz. 167; Fleenor, 1996, blz. 319; Goffin, Rothstein en Johnston, 1996, blz. 746; Tzi-

ner, Ronen en Hacohen, 1993, blz. 225). Het grootste gedeelte van deze recente onderzoeken versterkt verder voorgaande bevindingen:

- a. de validiteit van AC's cirkelt tussen 0,30 en 0,40.
- b. AC's zijn meer valide voorspellers voor potentieel dan voor prestatie.

Opvallend was ook dat één studie aantoonde dat er een verband bestond tussen de goede prestatie van managers in een AC en de omzet van de onderneming (Russell en Domm, 1996, blz. 25).

De resultaten van onderzoek in Nederland liggen in dezelfde lijn. Zo rapporteerden Van Leest, Maas, Nienhuis en Wiersma (1991, blz. 474) een correlatie van 0,48 tussen de O.A.R. en prestatie in de functie. In twee kleinere studies schommelden de validiteitscoëfficiënten tussen 0,34 en 0,44 (Albers en Altink, 1993, blz. 246-248). In een recente grootschalige studie ( $N = 355$ ) verkregen Jansen en Stoop (1994, blz. 193) een correlatie van 0,55 tussen de O.A.R. en huidig salaris.

Samenvattend: over het algemeen illustreert voorgaand internationaal onderzoek de goede voorspellende waarde van AC's. Dit betekent dat uitmunten in de microcontext van een AC een goede indicator is voor het slagen in latere managementfuncties. Zoals gezegd, is er echter nog geen onderzoek naar de voorspellende validiteit van AC's in België gepubliceerd.

## 2. Literatuurbevindingen over de billijkheid van assessment centers

Billijkheid ("fairness") van een selectieprocedure verwijst naar de mate waarin een selectie-instrument leden van een specifieke groep (bijvoorbeeld mannen of vrouwen) meer bevoordeelt dan anderen (Schmitt en Chan, 1998, blz. 205). In het AC-domein hebben voorgaande internationale studies hoofdzakelijk geanalyseerd of er significant verschillende resultaten bekomen worden op grond van etniciteit en geslacht (Baron en Janman, 1995, blz. 101). Wat de invloed van etniciteit betreft, zijn de resultaten van de studies verdeeld. Sommige onderzoeken vinden geen significante verschillen in gemiddelde scores tussen etnische groepen (Kriek, Hurst en Charoux, 1994, blz. 21; Huck en Bray, 1976, blz. 13; Jaffee, Cohen en Cherry, 1972, blz. 22; Russell, 1975, blz. 3). Andere studies rapporteren dat blanken opvallend beter scoren dan andere etnische

groepen (Schmitt en Hill, 1977, blz. 261; Ramos, 1992, blz. 173). Een recent onderzoek (Goldstein, Yusko, Braverman, Smith en Chung, 1998) biedt een plausibele verklaring voor deze onduidelijke onderzoeksbevindingen. Goldstein en collega's (1998) vonden dat de mate waarin etniciteit AC-prestaties beïnvloedt een functie is van de cognitieve component in een simulatieoefening. Met andere woorden, als een AC vooral bestaat uit oefeningen met een sterke cognitieve component (zoals de postbak en de case-analyse), kan men eerder verwachten dat etniciteit de resultaten vertekent.

Wat de invloed van het geslacht van de kandidaten in AC's betreft, laten de meeste voorgaande studies zien dat mannen en vrouwen even goed presteren in AC's (Iles en Robertson, 1989, blz. 257; Moses, 1973, blz. 569; Moses en Boehm, 1975, blz. 527; Ritchie en Moses, 1983, blz. 227; Weijerman en Born, 1995, blz. 284). Niettemin stelden Walsh, Weinberg en Fairfield (1987, blz. 305) vast dat vrouwen een hogere score kregen als ze beoordeeld werden door mannen. In een andere buitenlandse studie vond Shore (1992, blz. 499) eveneens dat vrouwelijke kandidaten een streep voor hadden, indien ze beoordeeld werden op prestatiegerelateerde gedragskenmerken.

Samenvattend: het internationaal wetenschappelijk onderzoek spreekt meestal in het voordeel van de billijkheid van AC's. De AC-methode heeft haar populariteit in de Verenigde Staten trouwens mede aan dit non-discriminatieaspect te danken. Toegepast op de situatie in België zou het interessant zijn te onderzoeken of er significante verschillen bestaan in AC-prestaties tussen mannen/vrouwen en Vlamingen/Walen.

## 3. Doelstelling en onderzoeksvragen van de huidige studie

Voorzover wij weten, is dit onderzoek de eerste gepubliceerde studie naar de voorspellende validiteit en de billijkheid van een AC in België, namelijk het BBL AC voor directieleden. Op deze manier draagt deze studie bij tot de verwetenschappelijking van HRM-instrumenten (in dit geval AC's) in België (Vallet, Demarest en Henderickx, 1995, blz. 91). Vier onderzoeksvragen stonden centraal:

1. In welke mate houdt het eindadvies dat op het AC volgt verband met het behaalde functieniveau, het behaalde salaris, de jaarlijkse beoordeling door de hiërarchie en de beoordeling van aanwezige managementcompetenties door de hiërarchie?

2. In welke mate houden de eindscores op de AC-gedragsskenmerken verband met deze diverse criteria?
3. In welke mate is er een significant verschil tussen mannen en vrouwen in het AC?
4. In welke mate er een significant verschil tussen Vlamingen en Walen in het AC?

#### 4. Methode

##### A. Algemeen

Het AC voor directieleden van de BBL wordt reeds sinds 1990 georganiseerd. Elk AC duurde twee en een halve dag. Per residentieel center werden telkens zes deelnemers geëvalueerd. Het AC werd ontwikkeld volgens de *Richtlijnen en Ethische Overwegingen bij het Toepassen van de Assessment Center Methode*, die in 1992 werden opgesteld onder auspiciën van de Vlaamse Vereniging voor Organisatie-, Consumenten- en Arbeidspsychologen (VOCAP). Bij de ontwikkeling en uitwerking van het AC opteerde de BBL voor een formule op maat, bijgestaan door SHL Belgium. De AC's die het onderwerp waren van deze studie, hadden plaats tussen 1990 en 1995.

##### B. Doelstelling van het assessment center

De doelstelling van het AC voor directieleden van de BBL bestond uit potentieelbeoordeling met het oog op een brede waaier van functies. Men wou potentiële opvolgers identificeren voor functies op de verschillende niveaus van het management en meer in het bijzonder voor het toekomstig senior management. Hierbij waren twee kanttekeningen relevant. Allereerst waren de resultaten op het AC niet noodzakelijk verbonden met een eventuele promotiebeslissing. Zo speelden ook andere "traditionele" elementen (zoals bijvoorbeeld gegevens uit het persoonlijk dossier) een rol bij het nemen van promotiebeslissingen. Ten tweede bood het AC ook ruimere perspectieven van loopbaanbegeleiding in de volle betekenis van het woord. Zo reikte het AC de nodige informatie aan om iedere deelnemer te kunnen begeleiden in een optimaal ontplooiën van het potentieel waarover hij/zij beschikt. Dit gold zowel op zijn/haar eigen niveau (indien doorgroeimogelijkheden beperkt waren) als in het perspectief van een progressief opgaande carrière. Ook was het verrijken

van het inzicht van de deelnemer in eigen functioneren en eventuele vormingsbehoeften een belangrijke doelstelling van ieder AC.

##### C. Steekproef

De steekproef telde 252 deelnemers (d.i. middle level managers), waaronder 201 mannen en 51 vrouwen. Hiervan zijn 128 personen Nederlandstalig en 124 personen Franstalig. Van alle deelnemers hadden er 249 de Belgische nationaliteit.

De gemiddelde leeftijd van de totale groep bedroeg 38 jaar ( $SD = 6$  jaar). De jongste deelnemer was 28 jaar en de oudste was 53 jaar. Tot slot varieerde de dienstanciënniteit van 1 jaar tot 38 jaar (gemiddelde anciënniteit = 16 jaar,  $SD = 8$  jaar).

##### D. Assessoren

Acht ervaren assessoren (vier mannen en vier vrouwen, gemiddelde leeftijd = 38 jaar,  $SD = 10$  jaar) van SHL Belgium fungeerden als assessoren. Vier van de assessoren waren Vlamingen, de overige assessoren waren Walen. Alle assessoren waren psychologen en externe consultants. Gemiddeld waren ze reeds 11 jaar actief in de selectiepraktijk ( $SD = 9$  jaar). Af en toe traden ook stagiairs op als schaduwasessor.

Alle betrokken consultants hadden kort na hun indiensttreding bij SHL Belgium een grondige training ondergaan. Deze assessorentraining duurde drie tot vijf dagen. In deze training werden de volgende topics behandeld: het observeren en systematisch beoordelen bij AC's, het integreren van de verschillende beoordelingen in een rapport, het geven van feedback aan kandidaten, het opzetten en implementeren van een assessment/development center, enz. Verder kwamen ook de definitie van de gedragsskenmerken en de opzet van de AC-oefeningen aan bod.

##### E. Gedragsskenmerken en oefeningen

Vooraleer de AC-benadering werd geïmplementeerd, werd een functieanalyse uitgevoerd. Hiertoe beoordeelden 73 directieleden de relevantie van de gedragsskenmerken. Dit resulteerde in dertien strategisch relevante competenties. Deze competenties werden ondergebracht in vier

clusters: leiderschap, communicatie/ commerciële zin, ondernemingsgerichtheid en intelligentie/persoonlijkheid. Appendix A geeft een overzicht van deze competenties en hun gedragsverankerde definities.

Wat de AC-oefeningen betrof, werd gekozen voor een ruime benadering, waarin zowel meer klassieke tests (bijvoorbeeld cognitieve vaardigheidstests en een creativiteitstest) als vragenlijsten (bijvoorbeeld een persoonlijkheidsvragenlijst) en situatietests een plaats hadden. Het AC bestond uit de volgende situatietests: twee in-trays, een presentatieoefening, een rollenspel, twee groepsoefeningen (met telkens zes deelnemers). Een aantal van deze oefeningen situeerden zich op een zuiver praktisch-operationeel niveau, terwijl andere oefeningen op een meer complex strategisch niveau gesitueerd waren. In deze studie spitsen wij ons toe op de situatietests als typisch kenmerk van AC's. Appendix B geeft een overzicht van de gebruikte situationele oefeningen in het AC.

#### F. Procedure

De BBL maakte het gebruik van de AC-methodiek bekend aan het eigen personeel via diverse informatiekkanalen (bijvoorbeeld het personeelsblad, een circulaire aan relevante personen, etc.). Om deel te kunnen nemen aan het AC diende "de hiërarchie" medewerkers voor te dragen. De aanleiding hiertoe kon meervoudig zijn. Enerzijds kon de chef vermoeden dat de betrokken medewerker een hoger functieniveau aankon. Anderzijds kon de behoefte bestaan om een betere kijk te krijgen op de groeimogelijkheden van de persoon.

Net vóór het eigenlijke AC vond een directe bespreking tussen de deelnemer en zijn directie plaats. In dit pre-assessment-gesprek kwamen de volgende topics aan bod: de doelstellingen, de procedure, het post-AC-gebeuren en de implicaties van het AC. Ook eventuele vragen van de deelnemer kregen er een antwoord.

In het eigenlijke AC evalueerden de assessoren de prestaties van elke deelnemer op de verschillende gedragskenmerken die door de oefeningen werden gemeten. Bij het observeren maakten de assessoren gebruik van observatiechecklists. Nadien beoordeelden zij de kandidaat globaal op het gedragskenmerk in de oefening. Zoals gebruikelijk roteerden de assessoren over de verschillende oefeningen. De assessor diende deze beoordeling vervolgens kort te verantwoorden. Op de laatste dag kregen

de deelnemers de gelegenheid om zich over de gehele procedure uit te spreken. Daarna volgde een halve dag voor de integratiediscussie tussen assessoren. In deze discussie namen de assessoren de tijd om per kandidaat de eindbeoordelingen op de gedragskenmerken verder te nuanceren en uit te diepen. Bij hun eindberaad maakten de assessoren zowel gebruik van de resultaten op de simulatieoefeningen als van de scores op de klassieke tests.

Op basis van het AC werden vier types advies geformuleerd: (1) negatief, (2) "ja-maar", (3) positief voor een tactisch-operationele functie of (4) positief voor een strategische functie. Zoals gezegd, werden latere promoties niet uitsluitend op grond van de AC-resultaten toegekend, maar speelden ook meer traditionele elementen (zoals bijvoorbeeld gegevens uit het persoonlijk dossier) een rol. Bovendien reikte het AC ook de nodige informatie aan om de deelnemers te kunnen begeleiden in het optimaal ontplooien van hun potentieel, zowel op hun eigen niveau als in het kader van doorgroeimogelijkheden. Grosso modo bestond het post-AC-gebeuren uit twee etappes: een post-AC-gesprek met een consultant/assessor en een post-AC-gesprek met de directie.

#### G. Criteria

De BBL stelde drie criteria ter beschikking, namelijk het behaalde functieniveau van de kandidaten, het brutosalaris van de kandidaten, en de laatste appreciatie door de hiërarchie die beschikbaar was voor de kandidaten. Als vierde criterium fungeerde de zogenaamde *Inventory of Management Competencies* (IMC). De IMC is een gestructureerde vragenlijst, waarop de manager (d.i. de chef) medewerkers op 16 generieke managementcompetenties beoordeelt. Voorbeelden van deze managementcompetenties zijn: commercieel, technisch, analytisch, innovatief, flexibel, etc. Hierbij komt het erop aan te vermelden in welke mate een specifieke gedraging karakteristiek is voor de beoordeelde. In dit onderzoek gebruiken wij het IMC-totaal (d.i. de som per kandidaat op de 16 managementcompetenties, via een partiële correlatie statistisch gecorrigeerd voor mildheid bij beoordelaars).

Deze criteria werden eind 1995 (maand juni) verzameld. De correlaties tussen deze criteria varieerden van -0,36 (tussen brutosalaris en IMC-totaal) en 0,50 (tussen functieniveau en appreciatie door de hiërarchie).

Tabel 1. Correlatie tussen de appreciatie van het assessment center en objectieve criteria (met en zonder statistische correctie voor anciënniteit).

	Appreciatie	Functie-niveau	Salaris	IMC-totaal
Appreciatie assessment center	0,16**	0,28***	-0,33***	0,48***
Appreciatie assessment center (statistische correctie voor anciënniteit)	0,20**	0,35***	0,15*	0,33***

Noot N varieert van 132 tot 260. IMC = Inventory of Management Competencies.  
\*  $p < 0,05$ ; \*\*  $p < 0,01$ ; \*\*\*  $p < 0,001$ .

## 5. Resultaten

### A. Predictieve validiteit van eindadvies en AC-eindscores

In het kader van de eerste onderzoeksvraag gingen wij na in welke mate het eindadvies dat volgde op het AC een predictieve waarde had voor de vier criteria. Uit tabel 1 blijkt dat het verband tussen het IMC-totaal en de prestatie in het AC opvallend is. Verder correleerde het eindadvies dat volgde op het AC significant met zowel de appreciatie door de hiërarchie als het functieniveau. Het is merkwaardig dat de correlatie met salaris significant negatief uitvalt. Deze laatste correlatie wordt in sterke mate beïnvloed door de "anciënniteit", die medebepalend is voor het salaris. Anciënniteit wordt echter in het AC niet in overweging genomen. Daarom voerden wij een nieuwe analyse uit op basis van een partiële correlatie, waarbij de gegevens statistisch gecorrigeerd werden voor anciënniteit. Deze analyse elimineert de variantie ten gevolge van anciënniteit, zodanig dat alle kandidaten behandeld worden alsof ze dezelfde anciënniteit hebben. Na statistische correctie voor anciënniteit veranderen de correlaties duidelijk. Tabel 1 toont dat niet alleen het eindadvies nog positiever correleert met de appreciatie door de hiërarchie en het behaalde functieniveau, maar ook dat de correlatie met salaris significant positief wordt.

De voorspellende waarde van het eindadvies kan ook nog vanuit een andere invalshoek onderzocht worden. Zo kunnen we bekijken hoeveel percent van de individuen die volgens het eindrapport van het AC in aanmerking kwamen om door te groeien tot een directiefunctie, in de daaropvolgende jaren daadwerkelijk zijn gepromoveerd tot een directie-

functie. Zoals reeds gesteld, steunde de promotiebeslissing slechts gedeeltelijk op de resultaten (eindrapport) van de kandidaat op het AC. Ook werd de promotiebeslissing genomen door andere personen dan de consultants uit het AC. Tabel 2 toont bijzonder positieve resultaten. Enerzijds resulteerde in 86,7% van de gevallen een uitstekend eindadvies ook daadwerkelijk in een promotie tot de directie. Anderzijds promoveerde slechts 15% van de individuen die op grond van het AC een duidelijk negatief advies ontvingen, toch tot een directiefunctie. Als we de aantallen in deze categorieën via een  $\chi^2$ -test vergelijken met een selectie op grond van toeval, dan zijn de resultaten duidelijk significant, met name  $\chi^2(3) = 50,49$  ( $p < 0,0000$ ).

Tabel 2. Carrière-evolutie van kaderleden na assessment center.

Type eindadvies	Aantal personen met dit eindadvies na assessment center	Aantal personen daadwerkelijk gepromoveerd tot directiefunctie	Hit ratio
Positief advies (strategisch niveau)	30	26	86,7%
Positief advies (tactisch niveau)	46	30	65,2%
"Ja-maar"-advies	60	25	41,7%
Negatief advies	60	9	15%

De tweede onderzoeksvraag hield verband met de mate waarin de eindscores op de AC-gedragsskenmerken (over de oefeningen heen) verband hielden met de vier criteria. Zoals reeds gezegd, kwamen deze eindscores tot stand door combinatie (via discussie) van de resultaten op de diverse AC-opdrachten. De resultaten van deze analyse worden weergegeven in tabel 3. Ook hier hebben we een analyse uitgevoerd met een statistische correctie voor anciënniteit. De AC-resultaten tonen nogal wat correlaties met de appreciatie door de hiërarchie, waarbij vooral inzet (bijvoorbeeld verandering bevorderen, doorzetter-initiator), doeltreffende communicatie en strategische planning bepalend zijn voor het succes zoals het door het management beoordeeld wordt. De correlaties met het verworven functieniveau zijn beperkter, maar ook hier blijken inzet en inzicht (in de betekenis van verandering bevorderen) doorslaggevend, terwijl anderzijds ook een loyale opstelling tegenover de organisatie (waardehiërarchie) belangrijk is. De correlatie met salaris is niet aanwezig. Klaarblijkelijk vertoont het salaris nog onvoldoende variantie na de statistische correctie voor anciënniteit om significante correlaties

mogelijk te maken. Wat de correlatie met het IMC-totaal betreft, wijzen de hoge significante correlaties uit tabel 3 op een zeer goede predictieve waarde van bijna alle AC-resultaten voor het IMC-totaal. De resultaten tonen dat enkel communicatiebereidheid, het participatieve aspect van leidinggeven en de beoordeling van waardehiërarchie niet significant correleren met de beoordeling in de praktijk.

Tabel 3. Voorspellende waarde van eindscores op de assessment-center-gedragsskenmerken.

	Appreciatie (N = 174)	Functie- niveau (N = 174)	Salaris (N = 172)	IMC-totaal (N = 84)
Communicatiebereidheid	0,12	0,09	-0,05	0,18
Participatief leiden	0,13	0,04	-0,06	0,11
Actie-crisis-interventie	0,11	0,03	-0,16	0,22*
Integriteit-stressbestendigheid	0,02	0,09	-0,10	0,35**
Doeltreffende communicatie	0,16*	0,14	-0,08	0,35**
Open commerciële omgang	0,00	0,08	-0,12	0,27*
Verandering bevorderen	0,17*	0,18*	0,00	0,39***
Tactische planning	0,16*	0,10	-0,10	0,34**
Strategische planning	0,15*	0,15*	0,01	0,38***
Intelligente aanpak	0,12	0,05	-0,09	0,31**
Doorzetter-initiator	0,17*	0,14	0,01	0,39***
Emotionele stabiliteit	-0,04	0,05	-0,09	0,29**
Waardehiërarchie	0,11	0,21**	0,02	0,16

Noot De correlaties met functieniveau en salaris werden statistisch gecorrigeerd voor anciënniteit.

\*  $p < 0,05$ ; \*\*  $p < 0,01$ ; \*\*\*  $p < 0,001$ .

## B. Billijkheid van assessment center qua geslacht en taal

In het kader van de derde onderzoeksvraag onderzochten wij via *t*-tests of mannen/vrouwen significant beter presteerden in het AC. Dit deden wij zowel op het niveau van de eindscores op de gedragskenmerken als op het niveau van het eindadvies. Op grond van de *t*-tests vonden wij nagenoeg geen significante verschillen tussen mannen en vrouwen op de eindbeoordelingen van de gedragskenmerken. De enige uitzondering was de eindscore op het gedragskenmerk doeltreffende communicatie,  $t(183) = -2,13$ ,  $p < 0,05$ . Vrouwen scoorden beduidend beter op dit gedragskenmerk. Deze positieve indicaties van de billijkheid van het BBL-

AC werden ook bevestigd door het eindadvies van het AC uit te splitsen naar geslacht. Zoals uit tabel 4 blijkt, kregen ongeveer dezelfde percentages mannen (38,3%) en vrouwen (41,2%) een positief advies.

Tabel 4. Percentage mannen en vrouwen, opgesplitst naar type eindadvies.

	Mannen		Vrouwen	
	%	Cumulatief %	%	Cumulatief %
Positief advies (strategisch niveau)	14,8	14,8	11,8	11,8
Positief advies (tactisch niveau)	23,5	38,3	29,4	41,2
"Ja-maar"-advies	29,6	67,9	39,2	80,4
Negatief advies	32,1	100	19,6	100

Om de vierde onderzoeksvraag te beantwoorden, analyseerden wij of de scores van Vlamingen en Walen beduidend van elkaar verschilden. Op grond van de uitgevoerde *t*-tests bestonden er opnieuw nagenoeg geen significante verschillen tussen Vlamingen en Walen op de eindbeoordelingen van de gedragskenmerken. De enige uitzondering was de eindscore op het gedragskenmerk intelligente aanpak,  $t(182) = 2,01$ ,  $p < 0,05$ . Vlamingen scoorden beduidend hoger op dit gedragskenmerk. Tot slot splitsten wij het eindadvies van het AC uit naar taal. Zoals uit tabel 5 blijkt, kregen ongeveer gelijke percentages Vlamingen en Walen een positief advies. Van de Vlamingen ontving 38,4% een positief eindadvies. Bij de Walen was dat 39,3%.

Tabel 5. Percentage Vlamingen en Walen, opgesplitst naar type eindadvies.

	Vlamingen		Walen	
	%	Cumulatief %	%	Cumulatief %
Positief advies (strategisch niveau)	12,8	12,8	15,6	15,6
Positief advies (tactisch niveau)	25,6	38,4	23,8	39,3
"Ja-maar"-advies	37,6	76,0	25,4	64,8
Negatief advies	24,0	100	35,2	100

## Besluit

Uit de onderzoeksresultaten kunnen wij de volgende interessante conclusies trekken. Allereerst beantwoordt het ontwikkelde AC duidelijk aan de gestelde verwachtingen en is het als instrument voldoende valide om predicties naar succes in hogere managementfuncties mogelijk te maken. Diverse elementen ondersteunen de positieve voorspellende validiteit van de procedure. Zo bleek het eindadvies uit het assessment-rapport op een significante wijze te correleren met de jaarlijkse appreciatie door de hiërarchie, met het functieniveau en (in mindere mate) met het salarisniveau.

Ten tweede kan dit bewijsmateriaal voor de goede voorspellende kracht doorgetrokken worden tot het niveau van de verschillende AC-eindscores. De resultaten toonden immers significante relaties met de appreciatie door de hiërarchie in de praktijk en (in mindere mate) met het functieniveau. Verder vonden wij voor bijna alle gedragskenmerken zeer significante correlaties met de totaalscore op de IMC-vragenlijst, die door de hiërarchie werd ingevuld voor de verschillende kandidaten. Enkel communicatiebereidheid, het participatieve aspect van leidinggeven en de beoordeling van waardehiërarchie correleerden niet significant met de beoordeling in de praktijk. Dit kan betekenen dat in de praktijk meer belang wordt gehecht aan een taakgerichte managementstijl (die ook gemakkelijker observeerbaar is voor de chef) dan aan een participatieve, mensgerichte managementstijl. Het feit dat waardehiërarchie niet correleert, heeft waarschijnlijk te maken met het feit dat het AC onvoldoende duidelijke gegevens oplevert om deze dimensie optimaal te beoordelen.

Ten derde zijn deze correlaties des te belangrijker aangezien zij in feite een onderschatting inhouden van de "werkelijke" correlatie. In de mate dat in deze berekeningen alle kandidaten opgenomen zijn, is het duidelijk dat de kandidaten van het laatste jaar of zelfs van de laatste twee tot drie jaar nog weinig of geen kans gekregen hebben om hogere graden te verwerven of om opgenomen te worden in de directie. Ook de salarisevolutie van deze personen is noodzakelijkerwijs beperkt, gezien het korte tijdsverloop tussen het AC en de registratie van de gegevens.

Ten vierde liggen sommige cijfers voor de voorspellende kracht van dit Belgische AC (bijvoorbeeld de correlatie met het IMC-totaal) iets hoger dan die uit voorgaand internationaal onderzoek. Een mogelijke verkla-

ring voor deze goede voorspellende validiteit ligt in de kwalitatief goede opbouw van het AC. Zo waren de assessoren uitsluitend ervaren en gedegen getrainde psychologen. Ook waarborgde de grondige functieanalyse van de senior-managementpositie, die aan het AC voorafging, dat de assessoren op functierelevante gedragskenmerken beoordeeld werden. Verder waren vele AC-opdrachten specifiek op maat gemaakt. Daarom is het van belang dat practici aan deze componenten aandacht besteden bij het ontwerp en de implementatie van AC's. Indien practici op deze manier – dus professioneel en gedegen – een AC ontwikkelen en daarbij ook rekening houden met de cultuur van de organisatie, dan is de validiteit van de AC-methode als selectie-instrument aantoonbaar performant. Dat wordt door deze studie aangetoond. Een andere mogelijke verklaring voor de opvallend goede resultaten van dit AC is dat er sprake is van zogenaamde "criterium-contaminatie" (Klimoski en Brickner, 1987, blz. 246-248). Doordat de hiërarchie betrokken was bij het feedbackproces en dus op de hoogte kon zijn van het rapport, is het namelijk mogelijk dat de hiërarchie met deze informatie in de latere jaarlijkse evaluaties rekening heeft gehouden. Dit zou dan de hogere samenhang tussen AC-scores en criteria kunnen verklaren. De kans op een zogenaamd "kroonprinseffect" bestond ook, aangezien dezelfde hiërarchie ook de kandidaten nomineerde in het AC. In een veldstudie over AC's is het bijzonder moeilijk deze bronnen van contaminatie te beperken of in de hand te houden.

Ten vijfde komt onze evaluatie van de billijkheid van het AC van de BBL ook in grote lijnen overeen met de resultaten van voorgaand onderzoek. Wij vonden nagenoeg geen verschillen tussen beoordelingen van mannen en vrouwen. Hetzelfde gold voor de beoordelingen van Walen en Vlamingen. Dit bevestigt dat AC's billijke selectie-instrumenten zijn. Billijkheid gaat echter verder dan het met elkaar vergelijken van de AC-resultaten van verschillende sociale groepen ("distributieve rechtvaardigheid"). Daarom verdient het de voorkeur dat toekomstig onderzoek naar de billijkheid van AC's zich ook richt op procedurele aspecten ("procedurele rechtvaardigheid"). Hieronder verstaan wij bijvoorbeeld de nominatieprocedure bij het AC, het feedbackgesprek of de instructies aan de deelnemers (zie Blanksby en Iles, 1993, blz. 35). Ook is het belangrijk om in de toekomst billijkheid via differentiële predictie te onderzoeken (zie Schmitt en Chan, 1998, blz. 205). In dit artikel hanteerden we de zogenaamde "naïeve" benadering van billijkheid, waarbij we analyseerden of bijvoorbeeld mannen en vrouwen gemiddeld verschilden qua AC-scores. Bij een meer genuanceerde benadering als differentiële pre-



dictie is de vraag of deze mogelijke verschillen (bijvoorbeeld tussen mannen en vrouwen) in hun gemiddelde AC-scores ook gerelateerd zijn aan gelijkaardige verschillen in de criteria (bijvoorbeeld scores op de IMC-vragenlijst).

Tot slot blijft het een open vraag of deze positieve resultaten veralgemeend kunnen worden naar andere AC's in België. Daarom moet ons onderzoek een aanmoediging vormen voor bedrijven en consultancy-bureaus om hun AC op een systematische manier te analyseren.

## Bibliografie

- ALBERS, M.B.M. en W.M.M. ALTINK (1993), "Assessment centers in Nederland: overzicht van enkele studies naar de psychometrische eigenschappen", *Gedrag en Organisatie*, nr. 6, blz. 241-253.
- BARON, H. en K. JANMAN (1996), "Fairness in the assessment centre", *International Review of Industrial and Organizational Psychology*, nr. 11, blz. 61-114.
- BLANKSBY, M. en P. ILES (1995), "Recent developments in assessment centre theory, practice and operation", *Personnel Review*, nr. 19 (6), blz. 33-44.
- CHAN, D. (1996), "Criterion and construct validation of an assessment centre", *Journal of Occupational and Organizational Psychology*, nr. 69, blz. 167-181.
- FLEENOR, J.W. (1996), "Constructs and developmental assessment centers: Further troubling empirical findings", *Journal of Business and Psychology*, nr. 10, blz. 319-333.
- GAUGLER, B.B., D.B. ROSENTHAL, G.C. THORNTON III en C. BENTSON (1987), "Meta-analysis of assessment center validity", *Journal of Applied Psychology*, nr. 72, blz. 493-511.
- GOFFIN, R.D., M.G. ROTHSTEIN en N.G. JOHNSTON (1996), "Personality testing and the assessment center: incremental validity for managerial selection", *Journal of Applied Psychology*, nr. 81, blz. 746-756.
- GOLDSTEIN, H.W., K.P. YUSKO, E.P. BRAVERMAN, D.B. SMITH en B. CHUNG (1998), "The role of cognitive ability in the subgroup differences and incremental validity of assessment center exercises", *Personnel Psychology*, nr. 51, blz. 357-374.
- HUCK, J.R. en D.W. BRAY (1976), "Management assessment center evaluations and subsequent job performance of white and black females", *Personnel Psychology*, nr. 2, blz. 13-30.
- ILES, P.A. en I.T. ROBERTSON (1989), "The impact of personnel selection procedures on candidates", in: P. HERRIOT, ed., *Assessment and selection in organizations*, Chichester, John Wiley, blz. 257-271.
- JAFFEY, C., S. COHEN en R. CHERRY (1972), "Supervisory selection program for disadvantaged or minority employees", *Training and Development Journal*, nr. 26, blz. 22-28.

- JANSEN, P.G.W. en B. STOOP (1994), "Assessment center graduate selection: decision processes, validity, and evaluation by candidates", *International Journal of Selection and Assessment*, nr. 2, blz. 193-208.
- KRIEK, H.J., D.N. HURST en J.A.E. CHAROUX (1994), "The assessment centre: Testing the fairness hypotheses", *Journal of Industrial Psychology*, nr. 20, blz. 21-25.
- KLIMOSKI, R.J. en M. BRICKNER (1987), "Why do assessment centers work? The puzzle of assessment center validity", *Personnel Psychology*, nr. 40, blz. 243-260.
- LIEVENS, F. en H. GOEMAERE (1999), "Assessment centers in Vlaanderen: Onderzoek naar de kwantiteit en kwaliteit van toepassing", *Tijdschrift voor HRM*, nr. 3, blz. 53-68.
- LIEVENS, F. en W. VANDER MEEREN (1998), "Assessment centers: vluchtige mode of volwaardige methode?", *Management en Organisatie*, nr. 52, blz. 24-41.
- MOSES, J.L. (1973), "The development of an assessment center for the early identification of supervisory potential", *Personnel Psychology*, nr. 26, blz. 569-580.
- MOSES, J.L. en V.R. BOEHM (1975), "Relationship of assessment center performance to management progress of women", *Journal of Applied Psychology*, nr. 60, blz. 527-529.
- RAMOS, R.A. (1992), "Testing and assessment of Hispanics for occupational and management positions: A developmental needs analysis", in: K.F. GEISINGER, ed., *Psychological Testing of Hispanics*, Washington, DC, American Psychological Association, blz. 173-194.
- RITCHIE, R.J. en J.L. MOSES (1983), "Assessment center correlates of women's advancement into middle management: A 7-year longitudinal analysis", *Journal of Applied Psychology*, nr. 68, blz. 227-231.
- RUSSELL, C.J. en D.R. DOMM (1995), "Two field tests of an explanation of assessment centre validity", *Journal of Occupational and Organizational Psychology*, nr. 68, blz. 25-47.
- RUSSELL, G. (1975), "Differences in minority/nonminority assessment center ratings", *Assessment and Development*, nr. 3, blz. 3-7.
- SCHMITT, N. en D. CHAN (1998), *Personnel Selection: A theoretical approach*, Thousand Oaks, Sage.
- SCHMITT, N. en T.E. HILL (1977), "Sex and race composition of assessment center groups as a determinant of peer and assessor ratings", *Journal of Applied Psychology*, nr. 62, blz. 261-264.
- SHORE, T.H. (1992), "Subtle gender bias in the assessment of managerial potential", *Sex Roles*, nr. 27, blz. 499-515.
- TASK FORCE ON ASSESSMENT CENTER GUIDELINES (1989), "Guidelines and ethical considerations for assessment center operations", *Public Personnel Management*, nr. 18, blz. 457-470.
- TZINER, A., S. RONEN en D. HACOEN (1993), "A four year validation study of an assessment center in a financial corporation", *Journal of Organizational Behaviour*, nr. 14, blz. 225-237.
- VALLET, N., S. DEMAREST en E. HENDERICKX (1995), "Het rekruteringsbeleid in Vlaanderen: de resultaten van een survey-onderzoek", *Economisch en Sociaal Tijdschrift*, jrg. 49, blz. 89-110.

- VAN LEEST P.F., R. MAAS, T. NIENHUIS en W.J. WIERSMA (1991), "Validiteit van een Nederlands assessment center", *Gedrag en Organisatie*, nr. 4, blz. 474-482.
- WALSH, J.P., R.M. WEINBERG en M.L. FAIRFIELD (1987), "The effects of gender on assessment centre evaluations", *Journal of Occupational Psychology*, nr. 60, blz. 305-309.
- WEIJERMAN, E.A.P. en M.P. BORN (1995), "De relatie tussen sekse en assessment center beoordelingen", *Gedrag en Organisatie*, nr. 8, blz. 284-292.

#### Abstract

*Assessment Centers in Belgium: the Results of a Study on Predictive Validity and Fairness*

*In Belgium assessment centers are increasingly popular. Despite this popularity, assessment centers have not been systematically evaluated in Belgium. Therefore, this study examines the predictive validity and fairness of an assessment center of a large bank. The sample consisted of 252 middle level managers (201 men and 51 women, mean age = 38 years). The results reveal that the assessment center fulfils the expectations. Assessment centers are valid enough to predict success in higher managerial positions. With respect to fairness this assessment center also scores quite well. Virtually no significant differences were found in ratings of men and women. Finally, we found no significant differences in ratings of Flemish and French speaking Belgians. Probably, these positive results in terms of predictive validity and fairness are due to the high-quality design of this assessment center.*

## Appendix A. Beschrijving van de gemeten gedragskenmerken.

Leiderschap	
1. Communicatiebereidheid t.o.v. de groep (CO)	Schenkt bij vergaderingen aandacht aan de manier waarop men vergadert, verwoordt doelstellingen duidelijk, licht beslissingen toe, kan het belang van een taak aantonen, etc.
2. Participatief leiden van ondergeschikten (PL) (Groepswerking)	Vraagt naar de mening van anderen, laat medewerkers participeren in het nemen van beslissingen, speelt informatie door, geeft valoriserende feedback, etc.
3. Actie-crisis-interventie (AK)	Bepaalt normen voor te behalen resultaten, kan snel beslissen en het roer omgooien als dat nodig is, weet een coherente tactiek op te bouwen in crisissituaties, intervineert tijdig, etc.
4. Integriteit en stressbestendigheid (IS) (Conflicthantering)	Steunt beslissingen in functie van een gemeenschappelijk doel, reageert rationeel en beheerst onder stress, weet gevoelsmatige situaties objectief te beoordelen, etc.
Communicatie en commerciële zin	
5. Doeltreffende communicatie (DC)	Kan zijn ideeën goed structureren, reageert vlot op onverwachte vragen, stelt synthetische verslagen op, drukt zich doeltreffend uit, weet een tekst overtuigend te brengen, etc.
6. Open commerciële omgang (OC)	Gaat vlot en zelfzeker met klanten om, ontspannen houding, diplomatie, speelt doelgericht in op problemen van klanten, drukt zich uit in termen van klantenvoordeel, etc.
Ondernemingsgerichtheid	
7. Doelgericht verandering bevorderen (VB)	Voorziet alternatieven, ontwikkelt actieplannen, stuurt een planning bij, werkt innovatieve procedures uit, bepaalt richtlijnen in termen van resultaten, etc.
8. Tactische planning (TP)	Analyseert problemen alvorens naar oplossingen te zoeken, onderzoekt de interne en externe mogelijkheden en beperkingen, weet consequenties van beslissingen correct in te schatten, stelt prioriteiten, kan anticiperen op veranderende omstandigheden, creëert structuren die nodig zijn om resultaten te behalen, maakt optimaal gebruik van beschikbare middelen, etc.
9. Strategische planning (SP)	Duidelijke kijk op de commerciële realiteit, onderzoekt mogelijkheden en beperkingen, heeft een ruime kijk op opportuniteiten, vooruitziend, structureert complexe gehelen op een doeltreffende manier, heeft een goed overzicht over het functioneren van de organisatie in haar totaliteit, etc.

Intelligentie en persoonlijkheid	
10. Intelligente, systematische aanpak (IA)	Concentreert zich op het essentiële, weet gegevens op een logische manier te groeperen, trekt sluitende conclusies, controleert de kwaliteit, kan relevante informatie vlot assimileren en toepassen, weegt de voor- en nadelen af, systematische en weldoordachte benadering, etc.
11. De manager als doorzetter en initiator van acties (DI)	Kan een inspanning volhouden tot het doel bereikt is, kan zich in iets vastbijten, toont zich ambitieus, geeft zelf de eerste aanzet tot actie, neemt spontaan initiatief, etc.
12. Emotionele stabiliteit (ES)	Kan een tegenslag vlot verwerken, beheerst zijn emoties doeltreffend, kan de dingen relativeren, zelfvertrouwen, etc.
13. Waardehiërarchie (WH)	Onderkent zijn eigen waarde en handelt consequent naar sociale normen en bedrijfsnormen, etc.

## Appendix B. Beschrijving van de gebruikte simulatieoefeningen (SHL).

Naam	Beschrijving
1. Boom Duval	In-tray (in-basket), die diverse functieaspecten van de algemeen manager van een scheikundige divisie uit een multinationale onderneming simuleert. De algemeen manager beheert vijf verschillende bedrijven en zetelt in de beheerraad van de holding, zodat zijn rol duidelijk meer strategisch dan operationeel is. Deze in-tray richt zich op kandidaten voor strategische managementfuncties en telt 15 items.
2. H.M.F. Financiën	In-tray voor de evaluatie van managementpotentieel, die geen beroep doet op voorafgaande managementervaring. Deze oefening beoogt een planning, die meer operationeel en tactisch is. De toebedeelde taak is die van public-relationsmanager in een plaatselijk kantoor van een internationaal financieel bedrijf. De oefening is ingedeeld in drie verschillende taken: planning en doeltreffend tijdsmanagement, concrete besluitvorming en het opstellen van een rapport. Deze in-tray is op een brede basis bruikbaar voor functies met een meer administratieve inslag en telt 25 items.
3. Noordgas	Analyse- en presentatieoefening over de strategische opties die door een middelgroot productiebedrijf genomen moeten worden in het licht van belangrijke veranderingen op de markt. Van de kandidaten wordt verwacht dat zij zich (als extern consultant) de informatie eigen maken en die zo goed mogelijk analyseren, om daarna hun strategie op een heldere en overtuigende manier voor te stellen. Voor de voorbereiding krijgen zij 75 minuten. De presentatie duurt 30 minuten.
4. ARF Verzekeringen	Oefening in "people management", waarbij de deelnemer de rol krijgt van een directeur die geconfronteerd wordt met een conflictsituatie. Deze conflictsituatie bevindt zich reeds in een vergevorderd stadium en dient hoognodig met de betrokkene besproken te worden. De kandidaat heeft 30 minuten om zich voor te bereiden en 30 minuten om het probleem te bespreken met "zijn medewerker".
5. Sanifusie	Groepsoefening, die zich op een strategisch beleidsniveau situeert. Deze oefening beoogt het beheersen van de kosten in een middelgroot bedrijf, waarbij iedere deelnemer een verschillende rol toegewezen krijgt. Er ligt heel wat stof voor conflict in deze oefening, die tevens een beroep doet op de numerieke aanleg van de deelnemers.
6. Marco Motors	Groepsoefening met betrekking tot een praktische probleemoplossingssituatie in groep, waarbij aan de kandidaten geen specifieke rol wordt toegewezen. In deze oefening is de groep verantwoordelijk voor de praktische indeling van de beschikbare ruimte in een nieuw gebouw voor een productiebedrijf. Het ter beschikking stellen van materiaal dat gebruikt kan worden bij het ontwerpen van de inplanting, verleent een extra dimensie aan de discussie.