



## A Multicriteria Decision Analysis Model for Faculty Evaluation

Carlos A. Bana e Costa, Paulo A. F. Martins, Mónica D. Oliveira, Amílcar  
Sernadas, Carlos A. Mota Soares

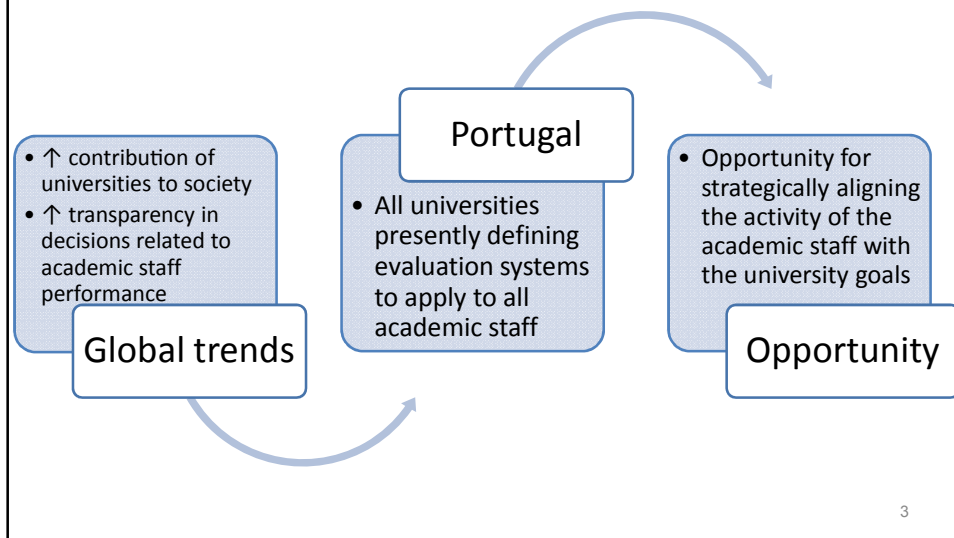
*Instituto Superior Técnico*, Technical University of Lisbon, Portugal

June 21st, 2010

### AGENDA

- Context
- Model building process
- Faculty evaluation model
- Application of the model to IST
- Discussion

## Universities: a changing environment



## Review



- Few studies attempting to evaluate the overall activity of the academic staff (Elmore, 2008)
- Different opinions:
  - Possible to measure with some precision faculty performance and performance measurements can be used in university management (Arreola, 2007)
  - Scientific activities cannot be fully measured given current knowledge and available indicators, and use of measurement tools might affect researchers' autonomy and lead to undesirable effects (Adler & Harzing, 2009)
- Methodological difficulties in faculty performance evaluation
  - “existing metrics do not capture the full range of activities that support and transmit scientific ideas” (Lane, 2010)
- Major flaws in both substance and process in previous studies

## Key challenges in faculty evaluation



Develop comprehensive evaluation systems

...based on methodologically sound procedures

...capable of reflecting differences between  
academic staff

...taking into account the university mission

...applicable to all faculty members  
and scientific areas while  
respecting their  
specificities.

...following a request: model to be used by *Instituto Superior Técnico*  
(IST), an engineering school with 778 faculty members working in a  
wide variety of scientific domains

5

## Key questions not previously addressed in a comprehensive and systematic manner within faculty evaluation literature



1. When STRUCTURING the faculty evaluation model:

- How to design a model reflecting the strategic objectives of the school and useful for human resources management?
- How to define a coherent set of evaluation criteria projecting, in the various areas of academic activity (pedagogical, scientific, etc), stakeholders' values and concerns about academic careers and institutional policies?
- How to describe, as objectively and unambiguously as possible, the performance on each one of the criteria, taking into account and adequately integrating its quantitative and qualitative dimensions?
- How to care for specificities of each one of the scientific domains of the school?

6



INSTITUTO  
SUPERIOR  
TÉCNICO

2. When modelling the MEASUREMENT of academics' VALUE:
  - How to convert individual performance into perceived added value to the school?
  - How to assign relative weights to the criteria adequately reflecting value trade-off judgements between criteria?
  - How to appropriately aggregate added value on multiple criteria, within and across areas of activity, respecting the autonomy of each faculty member to choose to invest more in some activities rather than in others, while not allowing extreme performance compensation phenomena inconsistent with achieving an adequate balance among objectives?
3. How to set BOUNDARIES for the RATING CATEGORIES imposed by law so that the classification of each faculty member may reflect her or his intrinsic value to the school?

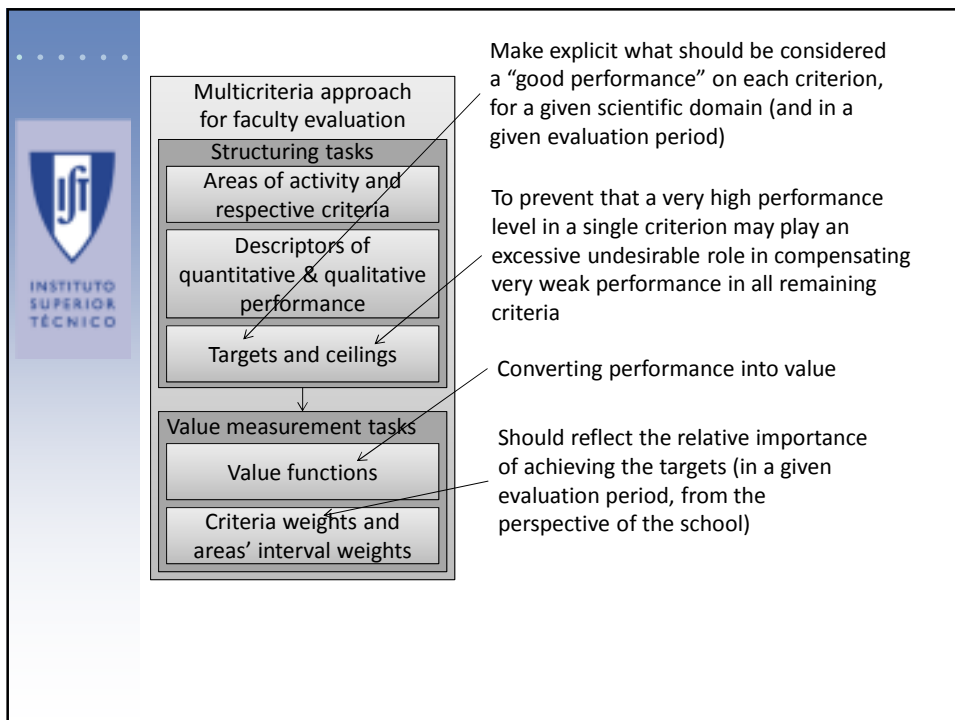
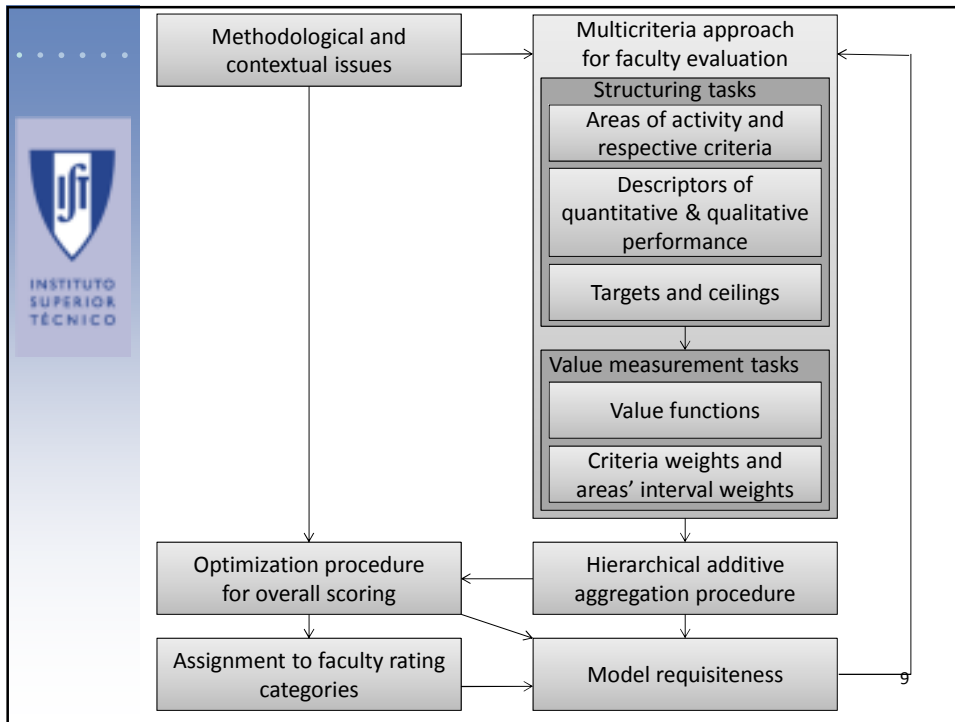
7

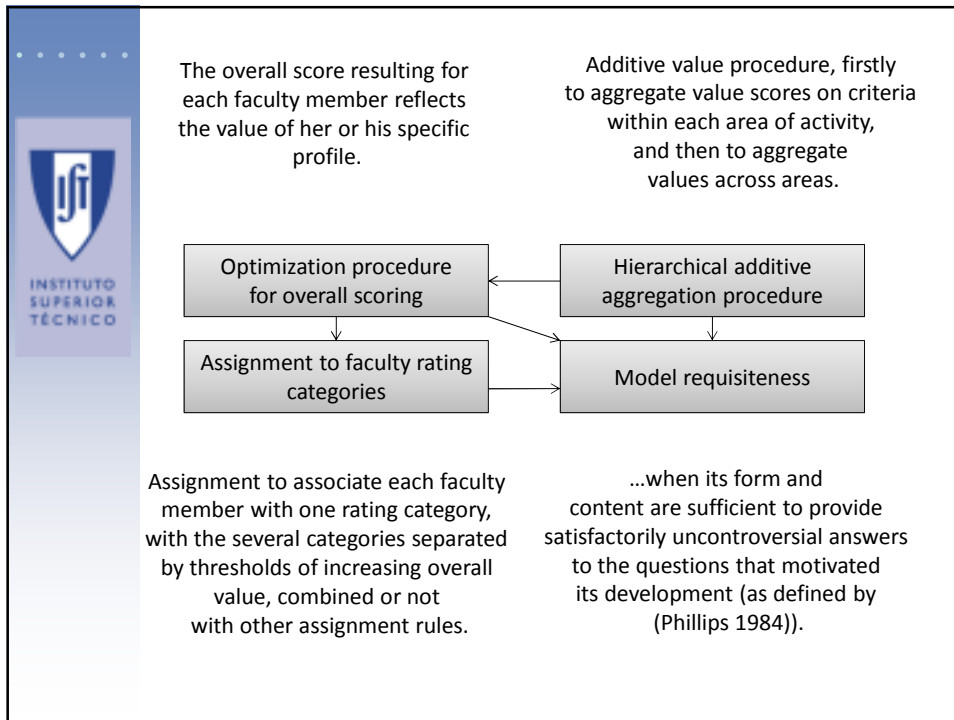


INSTITUTO  
SUPERIOR  
TÉCNICO

## MODEL BUILDING PROCESS

8





## Methods to build multicriteria value measurement models

- Several theoretically sound methods (von Winterfeldt and Edwards 1986) (Kirkwood 1997) (Belton and Stewart 2001)
- We propose the use of MACBETH, the Measuring Attractiveness by a Categorical Based Evaluation Technique:
  - Asks only for qualitative pairwise comparison judgements of difference in value between stimuli (Bana e Costa, De Corte et al. Forthcoming)
  - Theoretical foundations and has been extensively applied in various evaluation contexts (Bana e Costa, De Corte et al. 2005) (Bana e Costa, Lourenço et al. 2008)
  - Interactive application with the M-MACBETH software (Bana e Costa, De Corte et al. 2003)

12



INSTITUTO  
SUPERIOR  
TÉCNICO

## PROPOSED FACULTY EVALUATION MODEL

13



INSTITUTO  
SUPERIOR  
TÉCNICO

$d$  - Faculty member  
 $i_j$  - Evaluation criteria  $i$  from area of activity  $j$

$p_{i_j}^d$  - Performance  
 $Q_{i_j}^t$  - Quantitative performance  
 $Q_{i_j}^l$  - Qualitative performance

$$P_{i_j} = Q_{i_j}^t \times Q_{i_j}^l$$

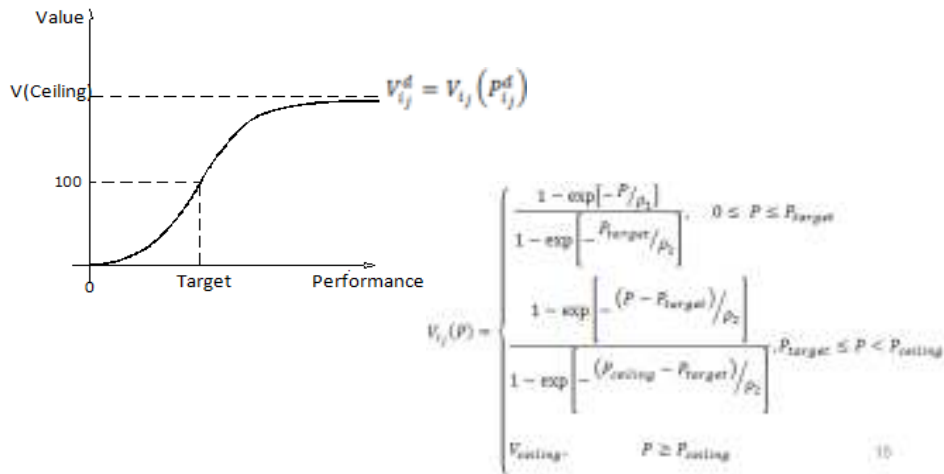
$$V_{i_j}^d = V_{i_j}(P_{i_j}^d)$$

14

## Building value functions

S-shaped value function with:

- Two branches defined by exponential functions;
- Exponential function respecting the delta property (or the constant trade-off attitude condition (Kirkwood, 1997));
- Desirable properties for faculty evaluation...



### Hierarchical additive procedure

$$V_j^d = \sum_{i_j=1}^{N_j} V_{i_j}^d \times w_{i_j}$$

With:

$$\sum_{i_j=1}^{N_j} w_{i_j} = 1$$

$$w_{i_j} > 0, \forall i, j$$

$$V_{i_j}^d = 100$$

$$V_{i_j}^d = 0$$

### Optimization procedure for overall scoring

$$V^d = \max \sum_{j=1}^M V_j^d \times w_j$$

Subject to:

$$\underline{w}_j \leq w_j \leq \overline{w}_j$$

$$\sum_{j=1}^M w_j = 1$$

$$w_j \geq 0$$

16





INSTITUTO  
SUPERIOR  
TÉCNICO

## APPLICATION OF THE MODEL TO IST

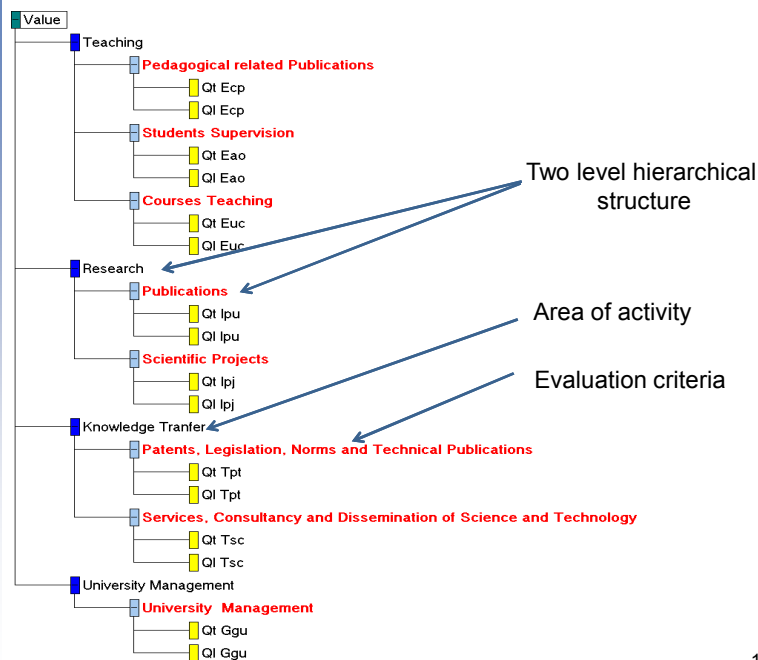
Model applied for retrospective evaluation:

2004-2007 & 2008-2009


17



INSTITUTO  
SUPERIOR  
TÉCNICO



18



Value judgements for building quantitative descriptors of performance

Type of publication	Equivalent units
International book	5.5
Article published in a type A journal	3
Article published in a type B journal	1.75
Chapter in an international book or Edition of international book	1.0
Article published in a type C journal	0.3
Article published in conference proceedings	0.2

### Qualitative descriptor of performance

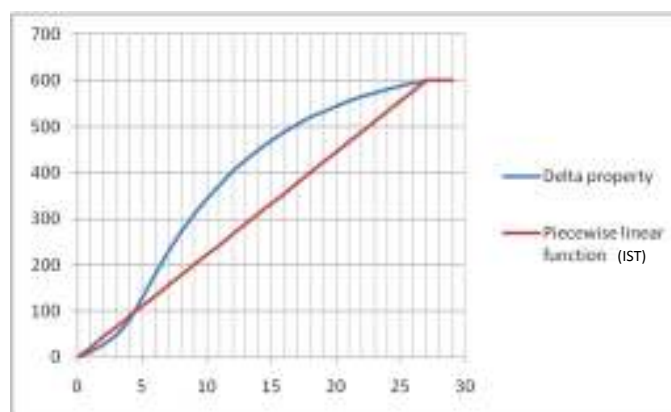
Levels of performance	Description	Value
Highly positive	<i>There is at least one 'strong' determinant point and no 'weak' determinant points</i>	1.5
Positive	<i>There are no 'strong' nor 'weak' determinant points, and 'strong' points more than compensate 'weak' points</i>	1.25
Neutral	<i>There was no identification of neither 'strong' nor 'weak' determinant points, or the 'strong' points are balanced with 'weak' points</i>	1.0
Negative	<i>There are no 'strong' nor 'weak' determinant points, and 'weak' points more than compensate 'strong' points</i>	0.75
Highly negative	<i>There is at least one 'weak' determinant point and no 'strong' determinant points</i>	0.5

## Examples of targets and ceilings

Evaluation criteria		Target	Examples of targets (to be interpreted with quality = 1)
Pedagogical related publications	Ecp	1.5	1 book chapter and 1 pedagogical text
Students supervision	Eao	6	2 supervision of MSc thesis per year
Courses teaching	Euc	9	9h of teaching courses per week with normal evaluation by students

Evaluation criteria	Pedagogical related publications	Students supervision	Courses teaching
	Ecp	Eao	Euc
<i>V(ceiling)</i>	500	300	300

## Value functions



22



INSTITUTO SUPERIOR TÉCNICO

### Building weights



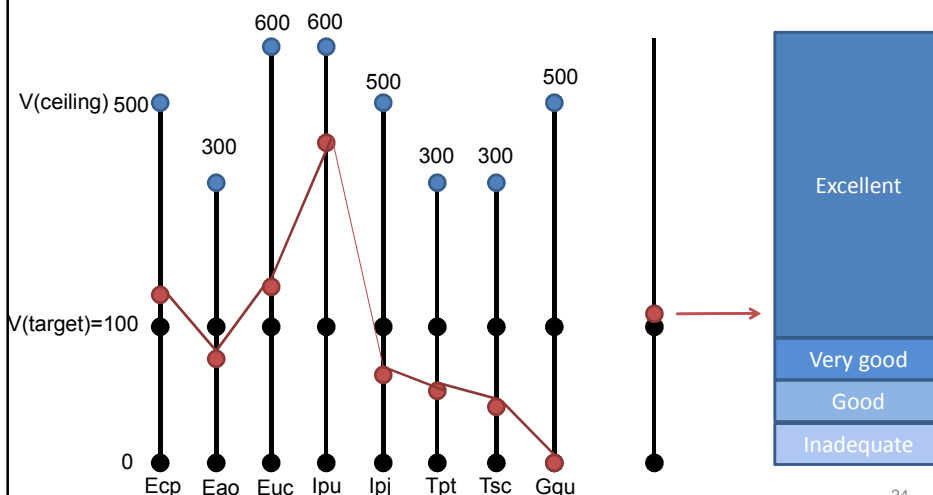
Area of activity	Pedagogical			Research		Knowledge transfer		University management
Weight by area of activity	20%-40%			40%-60%		0%-20%		0%*-20%
Evaluation criteria	Ecp	Eao	Euc	Ipu	Ipj	Tpt	Tsc	Ggu
Weight by evaluation criteria	1/6	2/6	3/6	6/8	2/8	1/3	2/3	1
Weight by evaluation criteria (normalized)	6.7%	13.3%	20%	45%	15%	6.7%	13.3%	20%

### Model outputs

#### 1. STAFF MEMBER PROFILE

#### 2. STAFF SCORE

#### 3. FACULTY RATING CATEGORY





INSTITUTO  
SUPERIOR  
TÉCNICO

## DISCUSSION

25



INSTITUTO  
SUPERIOR  
TÉCNICO

### Effectiveness of the model yet to be confirmed by large scale application!

1. Only the implementation of the model will show whether it is effectively dealing with differences across scientific domains, and whether adjustments are required.
2. Concerns with the calibration of the model when different targets and other values across scientific domains are used.
3. The incentives motivated by the model adoption have not been studied in detail.
4. It is not clear which is the level of acceptability of the model within the school.

26

## Improvements to the model (I)



INSTITUTO  
SUPERIOR  
TÉCNICO

- ✓ Being informed by literature in specific areas, e.g.,
  - Should evaluation of teaching be done only by students, or peers should also review the content of teaching? If peer review of the context of teaching is required, which methods for evaluation are available and have been validated?
  - Which is the best indicator for the impact of publications on the community?
- ✓ Using detailed information on the performance of IST academic staff:
  - The use of high-quality data and of scientific metrics might contribute to build a sounder model and to a higher level of acceptance.

27

## Improvements to the model (II)



INSTITUTO  
SUPERIOR  
TÉCNICO

- ✓ Using participatory mechanisms.
- ✓ Developing multiple criteria interactive analysis tools for the collection and analysis of model inputs and outputs.
- ✓ Developing multicriteria methods, for example:
  - Which procedures should be used for validating the chosen descriptors of performance?
  - Should thresholds be linked with targets and ceilings?
- ✓ Testing the model within schools other than engineering.

28

## Some questions for discussion...



INSTITUTO  
SUPERIOR  
TÉCNICO

- Performance appraisal: which viewpoint?
  - Rules for dividing the score for publications by the number of authors: individual vs. university viewpoints
- Value functions: avoiding complexity?
- Which incentives are being created? Simultaneous to faculty evaluation, which other tools might be used?
  - Regulating levels of teaching activity
  - Rewards to teaching

29

## References



INSTITUTO  
SUPERIOR  
TÉCNICO

- Adler, N. J. and A. W. Harzing (2009). "When knowledge wins: Transcending the sense and nonsense of academic rankings." *Academy of Management Learning and Education* 8(1): 72-95.
- Arreola, R. A. (2007). *Developing a Comprehensive Faculty Evaluation System: A Guide to Designing, Building, and Operating Large-Scale Faculty Evaluation Systems*, Anker Publishing Company.
- Bana e Costa, C., J. C. Lourenço, et al. (2008). "Development of reusable bid evaluation models for the Portuguese Electric Transmission Company." *Decision Analysis* 5(1): 22-42.
- Bana e Costa, C. A., J.-M. De Corte, et al. (2003). MACBETH. Working Paper LSEOR 03.56. O. R. Department. London, London School of Economics of Political Science.
- Bana e Costa, C. A., J.-M. De Corte, et al. (2005). On the mathematical foundations of MACBETH. Multiple Criteria Decision Analysis: The State of the Art Surveys. J. Figueira, S. Greco and M. Ehrgott, Springer. 76: 409-442.
- Bana e Costa, C. A., J.-M. De Corte, et al. (Forthcoming). MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique). *Wiley Encyclopedia of Operations Research and Management Science*. J. J. Cochran, Wiley.
- Belton, V. and T. J. Stewart (2001). *Multiple Criteria Decision Analysis: An Integrated Approach*, Springer.
- Elmore, H. W. (2008). "Toward objectivity in faculty evaluation." *Academe* 94(3): 38-40.
- Kirkwood, C. W. (1997). *Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets*. Belmont, California, Duxbury Press.
- Lane, J. (2010). "Let's make science metrics more scientific." *Nature* 464: 488-489.
- Phillips, L. D. (1984). "A theory of requisite decision models." *Acta Psychologica* 56: 29-48.
- von Winterfeldt, D. and W. Edwards (1986). *Decision Analysis and Behavioral Research*, Cambridge University Press.

30