Evaluation, Quality and Research

Rotterdam Jan 31, 2017

Professor Peter Dahler-Larsen, PhD
Dept of Political Science
University of Copenhagen
The Evaluation Society

- Evaluation machines
- Socio-political investment in indicators
- The knowledge society: Knowledge is a production factor and the social order becomes fragile (Stehr)
Outline

- Introduction: Evaluation Society
- Models: Role of evaluative data
- Concept of Quality
- Constitutive effects of indicators
- Implications
The role of evaluative data: Three models

- Metrological model
- Deliberative model
- Constitutive model
Metrological model

- Underlying concepts assumed to be clear
- Improve measurements through technical expertise and machinery
- Measurements are instruments for decisions
Deliberative model

- Multiple experts have multiple criteria
- Judgments are situated and dynamic
- Limited transparency, limited resources
Constitutive model

- Concepts are interpretable and partly defined by measurement
- Measurements are active players (language, interaction, imagination)
- Measurements provide stability and objectivity to aspects of reality
How the models relate

- Deliberative
- Metrological
- Constitutive
How the models relate II

- Deliberative
  - Metrological
  - Constitutive
How the models relate III

Deliberative

Metrological

Constitutive
Quality:
History of the concept

- Ancient Greece: Epistemological term
- 1600: Aesthetic and moral term
- 1900: Statistical term: Correspondence with norms
Criteria


- Clarity (e.g. transparency, precision, crispness, tightness)
- "Quality" (craftsmanship, depth, soundness)
- Originality
- Significance (of topic, of impact)
- Methods
- Feasibility
- Informal, evanescent criteria (elegance, intelligence, personal qualities, moral qualities)
Increasing number of criteria in the knowledge society

- Relevance
- Impact

... which have an indeterminate relation with "quality"
How the models relate III

- Metrological
- Constitutive
- Deliberative
Constitutive effects

- When indicators become a way of seeing reality
- When an indicator system defines what it claims to measure
- When it shapes a landscape of meanings in which action takes place
Why not just *unintended effects*?

- Whose intentions?
- At what time?
- Only official intentions?
- How clear?
- How can intentions be captured empirically?
- Solidarity with the system architect?
What helps make constitutive effects real

- Limited organizational attention (H. Simon)
- Pressures to make an institution look good
- Local "wise men": If you want to score on the indicators, follow my advice
- Affective reactions (self-esteem, status, strategy)
- Ad hominem arguments!
- Institutional "lock-in"
Constitutive effects in the following domains

- Content (research topic, type of publication, field..)
- Time
- Social relations
- World view (model for research)
Constitutive effects I

Definition of “field”

- Bibliometric indicators should be used only after control for field effects
- That requires a definition of field!
- Institut for Statskundskab
  - “Political Science”?
  - “Public Management”? 
L-index: % of total Google Scholar citations of Non-English publications

L-index among 42 professors in Pol Sci in Denmark by subfield

- Public Admin
- Com Pol
- Pol Behav
- Int Relations
- EU studies
Constitutive effects of

- Counting only the impact of English publications?
- Normalizing/not normalizing by "subfields"?
- L-index itself?
Constitutive effects: Reactions towards BFI
(Danish bibliometric research indicator)

<table>
<thead>
<tr>
<th></th>
<th>Optimists</th>
<th>Pessimists</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguity</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sources of ambiguity

- The ambiguity of the term quality
- Uneven implementation of measurement regimes
- The balance between metrological/deliberative/constitutive model
- Which of several indicators are most important?
- Min and max requirements? (indicators vs eval criteria)
- Do collective goals apply to individuals?
- Will there be control for age, gender, subfield, teaching, managerial roles, workplace contributions?
- Money versus indicator scores?
- How stable are indicator systems?
- Should I use a bad indicator for a good purpose?
Constitutive effects III:
Reactions towards BFI
(Danish bibliometric research indicator)

<table>
<thead>
<tr>
<th>Clarity</th>
<th>Optimists</th>
<th>Pessimists</th>
</tr>
</thead>
<tbody>
<tr>
<td>Producer</td>
<td>Stressed</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ambiguity</th>
<th>Optimists</th>
<th>Pessimists</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pragmatic</td>
<td>Confused</td>
<td></td>
</tr>
</tbody>
</table>
Constitutive effects: Sum up

To establish a system which measures quality of research is an active intervention into the meaning of research, timing, fields, and the people involved.
Implication

Evaluation of indicators

- One thing is to document growing production (higher score on indicator)
- Another thing is to conclude that the indicator was a success because it helped increase production
- A conflation of the two is a logical leap that supports a self-fulfilling prophecy
What to do

- Do not unknowingly amplify the importance of indicators
- If there is one dominant indicator, enjoy weakening it
- Evaluate indicators now and then
- Pay attention to constitutive effects
- Discuss the normative assessment of constitutive effects
- Act responsibly and sometimes ignore evaluation
- Keep the future open for the unknown
Literature

- Michèle