



The Changing Role of Evaluation in a Changing Society

June 14, 2012

Centrum für Evaluation
Saarbrücken

Professor Peter Dahler-Larsen, PhD
Dpt of Political Science and Public Management
University of Southern Denmark



Evaluation as a construction

- Evaluation is assisted sense-making (Mark et al.)
It is *constructed* by means of activities, institutions, politics, norms, values, expectations...
evaluation imaginaries (T. Schwandt)
- Evaluation *helps construct* decisions, policies, practices... the social order becoming *fragile* (N. Stehr)
- Evaluation is a conceptual *construct*. A concept (more than a word) is politically and historically controversial. Concepts seize the future (Koselleck)



Outline

- Evaluation: What are we talking about?
- Five issues
 - Popularization
 - Outcome orientation
 - Systematization of evaluation
 - Changing relations: evaluation – research
 - Changing patterns of utilization
- Summing up



A formal definition of evaluation includes (Shadish et al.)

- A systematic and methodological inquiry...
- Value-based...
- Looks at interventions, activities etc...
- Aiming at use, such as improvement...



Evaluation as...

- Evaluation: a product such as a report
- Evaluation: an evaluation process
- Evaluation: *a field of activity*



Evaluation





Evaluation

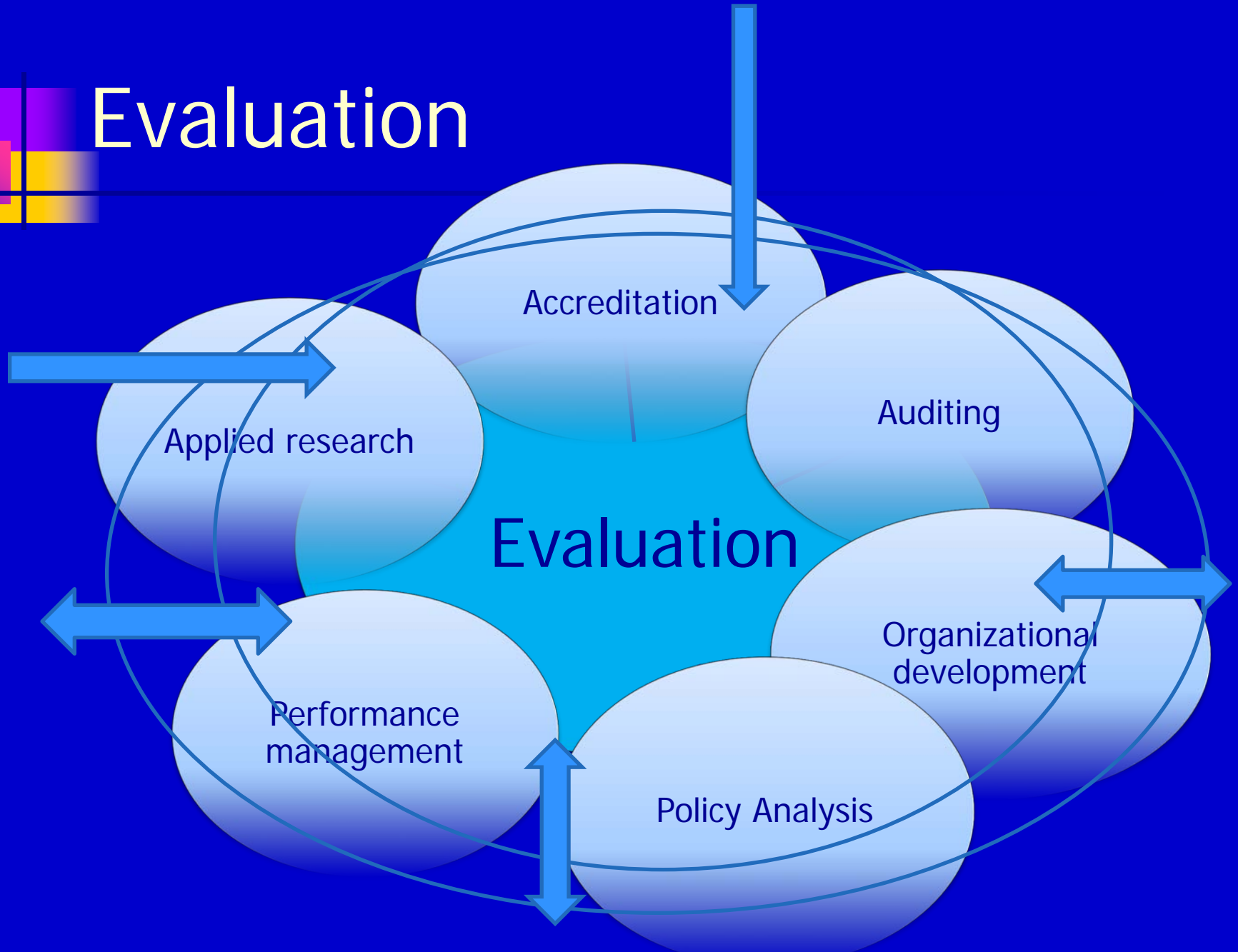




Evaluation



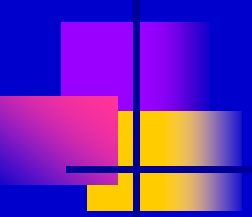
Evaluation





Evaluation as...

- Evaluation: a product such as a report
- Evaluation: an evaluation process
- Evaluation: a field of activity
- Evaluation: *the historical, political, social construct of evaluation*



...unplanned long-term processes moving in a discernable direction – with spurts and counter-spurts to and fro. They arise from the interweaving, the conjunction, co-operation and confrontation of many planned activities (N. Elias)

... a living tension between reality and concept (R. Koselleck)

Do we include people?

Who?



- Evaluation thinkers ("The Great Men") (M. Alkin)
- Evaluation developers, teachers, advisors, evaluation journal editors, evaluation center leaders, evaluation society spokespersons and authors of handbooks and guidelines, theory weavers (N. Stame)
- Evaluation practitioners



Should we talk about...

- Evaluation
- Evaluation
- Évaluation
- Evaluación
- Evaluazione
- Evaluering
- Utvärdering... or
- Naliliineq?



Outline

- Evaluation: What are we talking about?
- Five issues
 - Popularization
 - Outcome orientation
 - Systematization of evaluation
 - Changing relations: evaluation – research
 - Changing patterns of utilization
- Summing up



Issues

- A question, point or concern to be disputed or decided; a main matter of contention; a sticking point or grievance; a belief at variance (*R. Stake*)
- *Issues* are to be expected, because
 - evaluation is constructed
 - evaluation constructs



Outline

- Evaluation: What are we talking about?
- Five issues
 - Popularization
 - Outcome orientation
 - Systematization of evaluation
 - Changing relations: evaluation – research
 - Changing patterns of utilization
- Summing up



Popularization

- Making evaluation *popular*, supporting it with norms, values, institutional support
- Bringing evaluation to the *populus* (people) making them objects/consumers/participants in evaluation



Popularization in different eras

- Modernity (Myth of *Progress*)
 - Difference: experience vs expectation (Koselleck)
 - Autonomy
 - Rationality
- Reflexive modernization (Myth of *Development*)
 - Side-effects
 - Contingency
 - Multiple perspectives
- Audit society (Myth of *Assurance*)
 - Comprehensive, mandatory surveillance
 - Reporting mechanisms
 - Management systems



Popularization in different eras

- Modernity (Myth of *Progress*)
 - Difference: experience vs expectation (Koselleck)
 - Autonomy
 - Rationality (*experimentalist eval*)
- Reflexive modernization (Myth of *Development*)
 - Side-effects
 - Contingency
 - Multiple perspectives
- Audit society (Myth of *Assurance*)
 - Comprehensive, mandatory surveillance
 - Reporting mechanisms
 - Management systems



Popularization in different eras

- Modernity (Myth of *Progress*)
 - Difference: experience vs expectation (Koselleck)
 - Autonomy
 - Rationality (experimentalist eval)
- Reflexive modernization (Myth of *Development*)
 - Side-effects
 - Contingency
 - Multiple perspectives (*illuminative, responsive, participatory e.*)
- Audit society (Myth of *Assurance*)
 - Comprehensive, mandatory surveillance
 - Reporting mechanisms
 - Management systems



Popularization in different eras

- Modernity (Myth of *Progress*)
 - Difference: experience vs expectation (Koselleck)
 - Autonomy
 - Rationality (experimentalist eval)
- Reflexive modernization (Myth of *Development*)
 - Side-effects
 - Contingency
 - Multiple perspectives (illuminative, responsive, participatory e.)
- Audit society (Myth of *Assurance*)
 - Comprehensive , mandatory surveillance
 - Reporting mechanisms
 - Management systems (*performance indicators, auditing, evaluation machines*)



Popularization: Paradox

- Evaluation has become popular, but
- The imaginaries which support evaluation are in tension with each other
 - Not all beliefs in evaluation are rational
 - Evaluation is spreading before evaluation competence
 - Difficult to identify the "evaluator"
- Risk of evaluation fatigue



Outcome orientation:

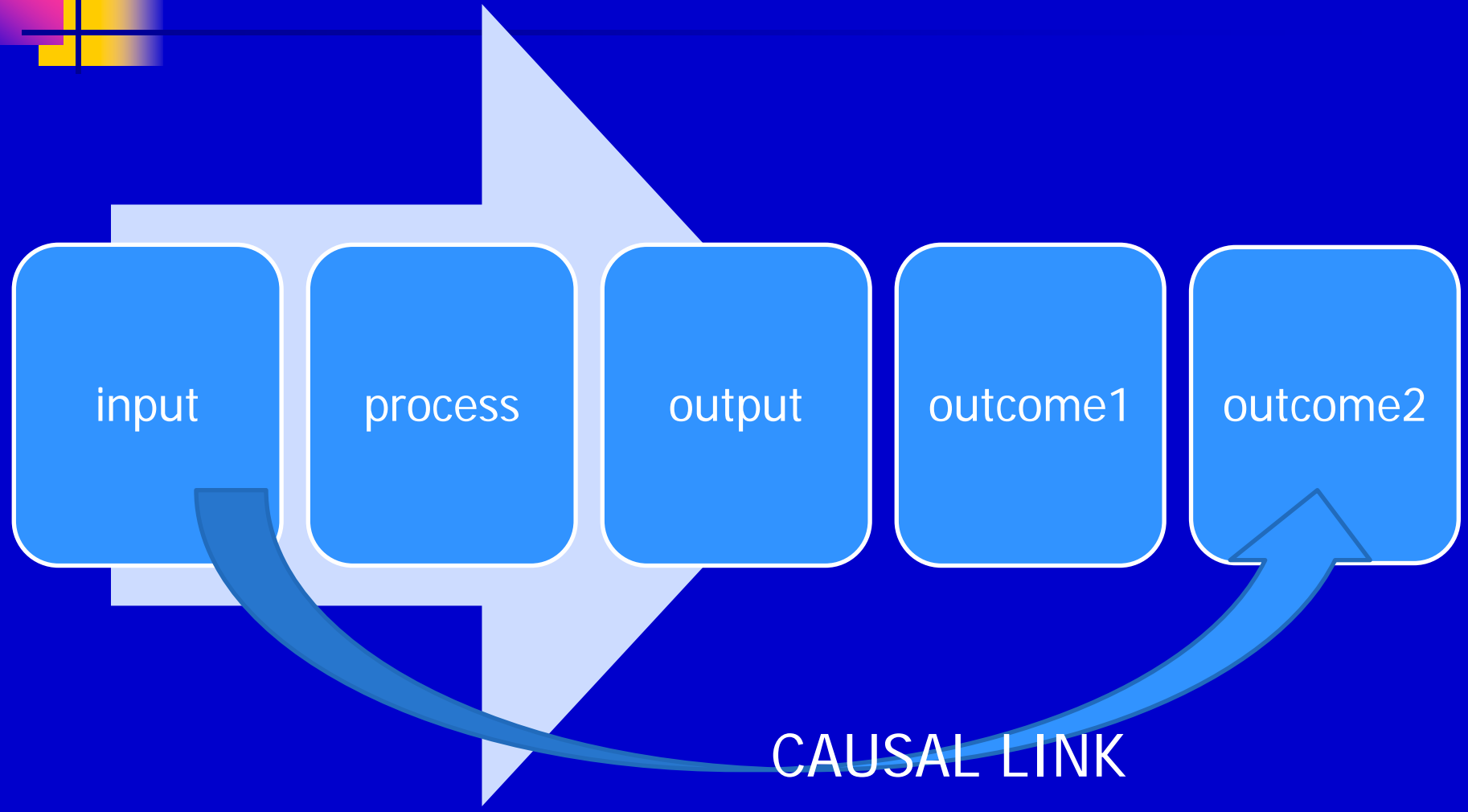
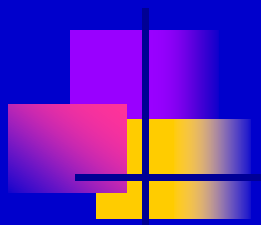
Five motivations

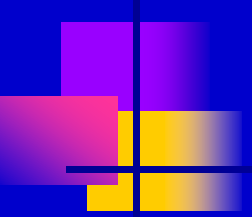
- Redefinition of sense and purpose: We are here for the citizens
- Anti-bureaucracy: Emancipate professionals! Let managers manage!
- Improve the knowledge base
- Effects should be measured (weak NPM)
- Effects should be rewarded (strong NPM)



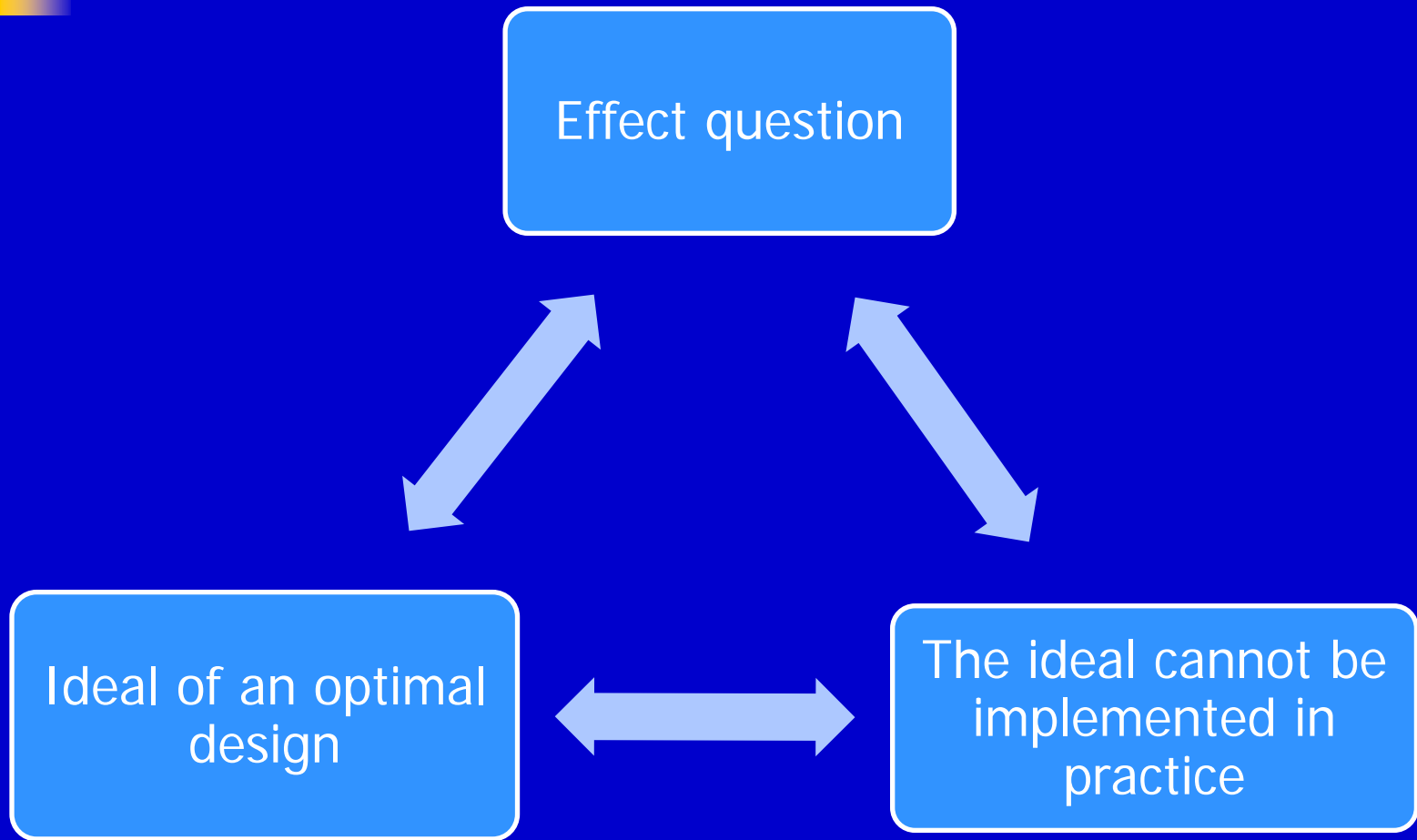
Terminology

- Results, outcomes, impacts are sometimes referred to without further specification
- In evaluation terminology, the term *effect* is reserved for situations where a causal link is assumed or demonstrated

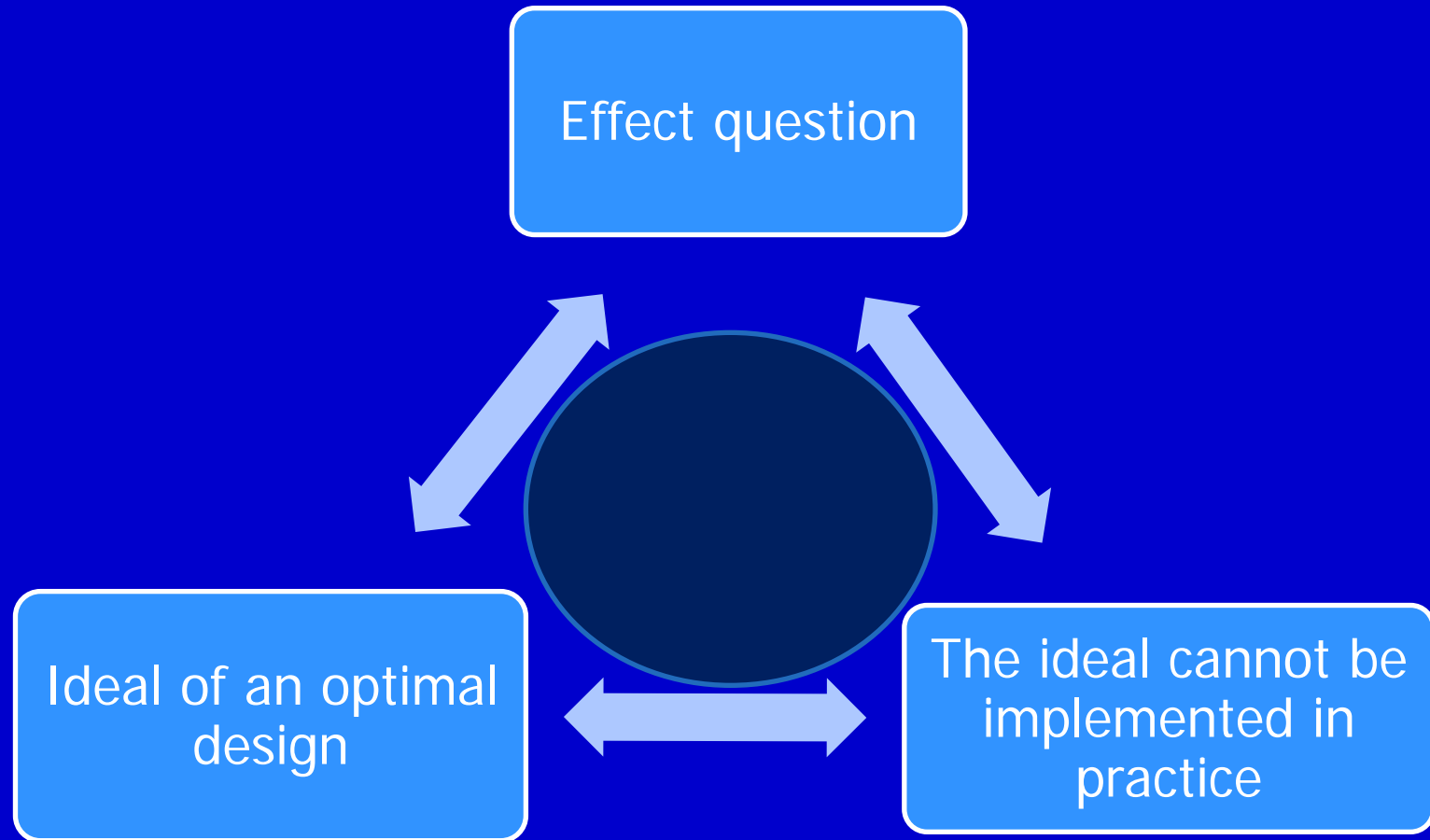


- 
- As a result, in order to measure effects, some methodologies seek to isolate interventions and results, and an evidence hierarchy is mobilized, such as:
 - Reviews of RCTs
 - Randomized controlled trials
 - Studies with generic or statistical controls
 - Pretest-posttest (reflexive control)
 -

The Bermuda Triangle of evaluations



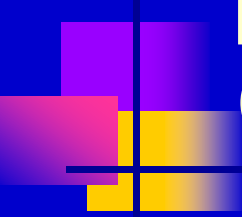
The Bermuda Triangle of evaluations





Because...

- Evaluator appears too late
- Little control over intervention and context
- No adequate comparison
- Large scale, small N
- Context and intervention impossible to separate
- Contexts are too different



Effectiveness of external inspection in improving health care organisational behaviour, health care professional behaviour, or patient outcomes (review)
(Flodgren, Pomey, Taber & Eccles, Cochrane 2011)

- 9901 records surveyed
- 15 studies reviewed
- 2 studies included (UK and SA)

"It is not possible, or even desirable, to compare or synthesise results from studies in which the conditions during which health care is provided are so different...basic necessities like soap and paper towels were not available in more than half of the included hospitals."



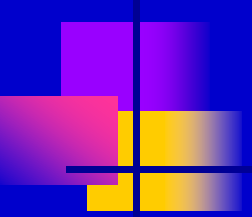
What evaluators do to establish *effects*

- Motivate early engagement in evaluation thinking
- Discuss what counts as effects
- Plan interventions that are evaluable
- Standardize and control interventions
- Find comparison groups
- Involve experts
- Conceptualize intervention theories (specific/necessary signs)
- Engage with professionals/clients to gauge relevance of effect indicators
- Argue about credibility and use of particular study in situation
- Build evaluation capacities



What evaluators do to establish effects: *verbs only*

- *Motivate*
- *Plan*
- *Standardize*
- *Find*
- *Involve*
- *Conceptualize*
- *Engage*
- *Argue*
- *Build*



Whatever counts as effects in a given situation is based on a large set of activities, relations, and arguments



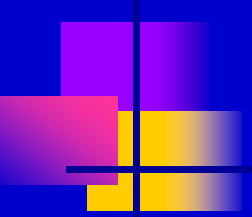
Example: Bolius - a Danish non-profit org which enhances life quality of 400,000 house owners

- Vision 1: Bolius takes responsibility for its effects. Bolius measures *effects* by isolating how often Bolius information packages lead house owners to take action on five dimensions of quality of life



Example: Bolius - a Danish non-profit org which enhances life quality of 400,000 house owners

- Vision 1: Bolius takes responsibility for its effects. Bolius measures *effects* by isolating how often Bolius information packages lead house owners to take action on five dimensions of quality of life
- Vision 2: Bolius works together with house owners. Bolius demonstrates its relevance through an understanding of the needs of house owners and through the quality of the dialogue with them



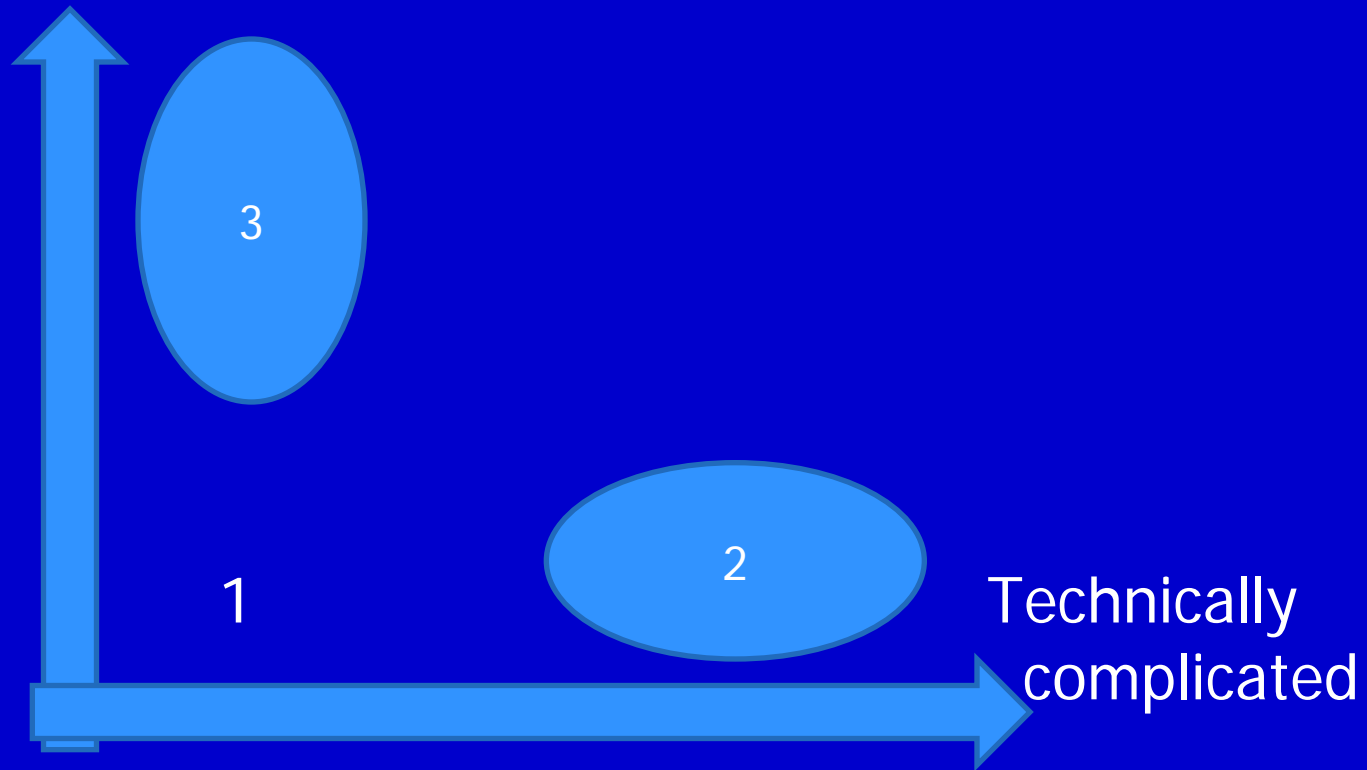
A lack of control of the intervention may not be a small problem to be fixed by "getting the design right", but a sign of the social, political and organizational characteristics of the intervention itself.... a *complex* intervention

There are genuine sources of complexity in our type of society (globalization, technology, crisis, cultures...)

Complex interventions

(Stacey, Zimmerman, Patton)

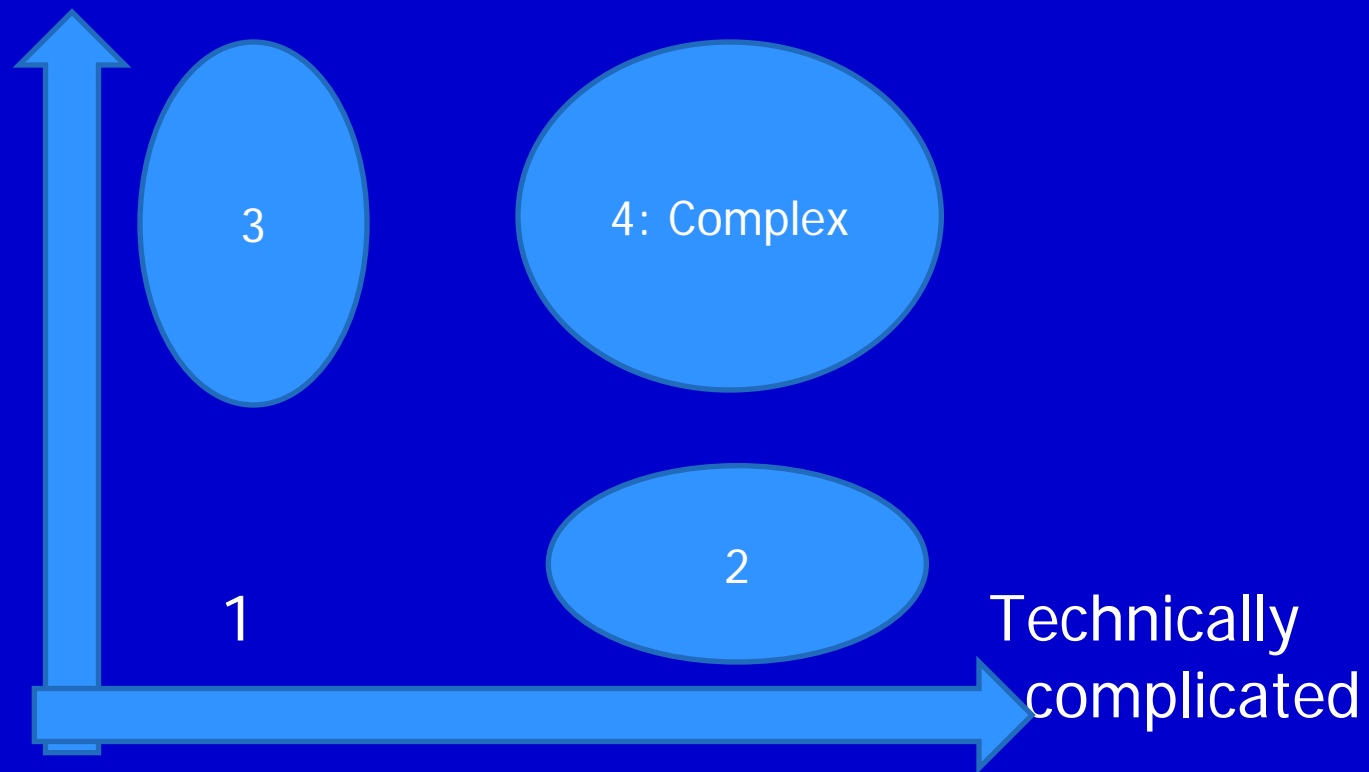
Socially complicated





Complex interventions

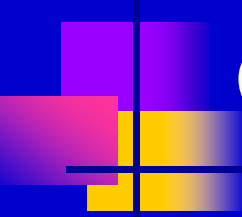
Socially complicated



Characteristics of complex interventions



- Contested key terms: safety and quality
- Several disciplines, several paradigms
- Meaningfulness of intervention matters
- Uneasy relation: Problem structure, accountability structure and intervention structure
- Dynamic intervention (politics and organizational learning).
- Organizational aspects matter



Under complex conditions, evaluators focus on

- Facilitating interpretation of less than certain results
- Ongoing attention & rapid feedback
- Emphasizing intervention-context responsibilities
- System change
- Tipping points



Outcome orientation: Paradox

- In order to measure effects, some methodologies seek to *isolate* interventions and results
- But the complementary side: *contexts and interactions* is of increasing importance both theoretically, practically and politically



Systematization of evaluation

(Leeuw and Furubo 2008)

- Integration of evaluation into routines which
- Organize (in super-scales)
- Streams of information that are
- Comprehensive and mandatory
- And have managerial implications



Systematization works together with

- Evaluation policy
- Evaluation strategy
- Evaluation capacity
- Evaluation culture
- Evaluation machines



Evaluation machines

- Permanence
- Prospective approach to quality
- Standardization: Abstract, general operations
- Mandatory
- Objective evaluation based on handbooks, guidelines, indicators, IT

Systematization of evaluation: Paradox



- One the one hand evaluation becomes rational, well-planned in large scales, but on the other:
- Evaluation machines confuse responsibility
- Adiaphorization due to social distance (Bauman)
- Microquality and defensive quality
- Increasing costs (no evaluability assessment!)
- Democracy?



Changing relations between evaluation and research

- Common roots: Evaluation research
 - Academics helped define the field
 - Assumptions of rationality: division of labour with politics/practitioners
- Establishing evaluation separate from research
 - Pragmatic participatory: use (M. Patton, B. Cousins)
 - Transformative evaluation: values/politics (D. Mertens, J. Greene)
 - The market for evaluation
 - Integration into organizations & management
- Partial reconnections: Evaluation – research
 - Theory (*Pawson and Tilley*)
 - Methodology is "back with a vengeance" (Evidence movement)



Promises of new evaluation – research links

- Research offers a set of institutional rules (protection and quality)
- Research is a rich set of resources (theories, methods, skills, roles, and ways of arguing)
- Research (e.g. in the sociology of knowledge) has developed advanced views on uncertainty, perspectivity, positionality, values, etc.
- Research on evaluation (Mark)

Evaluation-research relation:

Paradox



- Some distance from research has helped evaluation establish itself as a field...
- But research is a rich source of inspiration for evaluation (methodology, theory, rules, roles) and
- Research on evaluation would be welcome



Changing patterns of utilization

The utilization problem (meaning too little use) has been (the) main driver of new models and approaches in evaluation



New insights

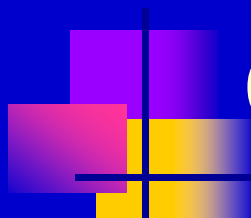
- Knowledge society: knowledge is a *productive force* and the social order is *fragile*
- A contextualist, situational, nuanced view on types of use
- An attention to consequences of evaluation systems, such as
 - imposed use (Weiss)
 - performance paradox (van Thiel and Leeuw)
 - constitutive consequences (Dahler-Larsen)



Factors influencing the use of evaluation (Ledermann 2012)

- How surprising are the results?
- How is the quality of the evaluation (perceived)
- Contextual factors

Evaluation argumentation in different contexts (Valovirta 2002)

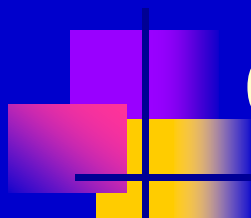


Level of conflict



Pressure for change

Evaluation argumentation in different contexts (Ledermann 2012)



Level of conflict

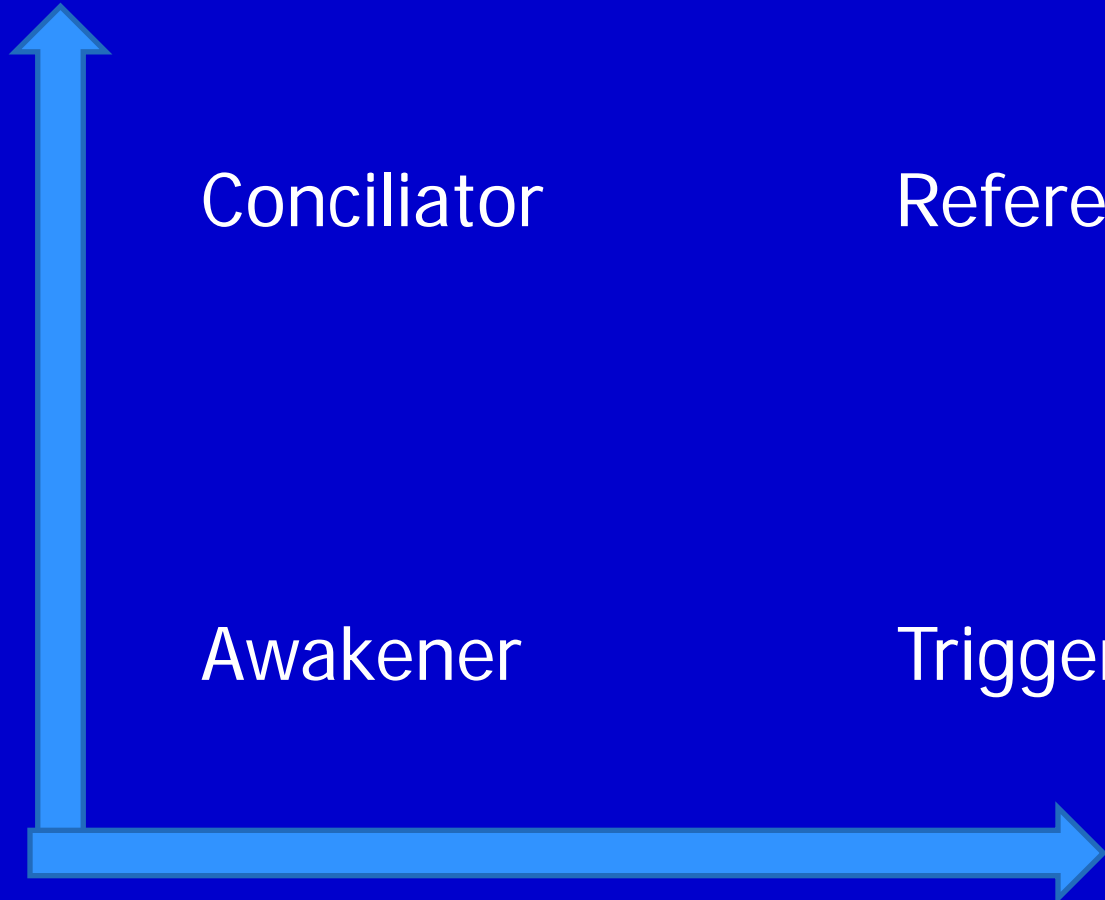
Conciliator

Referee

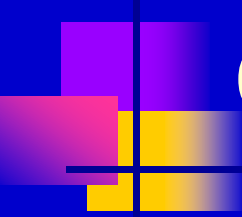
Awakener

Trigger

Pressure for change



Expanding typology of forms of use/consequences

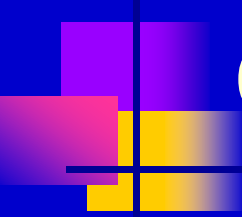


- Accountability/control (summative)
- Learning/development (formative)
- Enlightenment
- Strategic
- Tactical
- Symbolic
- *Proces*
- *Constitutive*

Lessons on the importance of proces



- Pragmatic-participatory evaluators:
 - Process use is an important source of learning/reflexivity
- Organizational learning theorists:
 - Integrate evaluation!
 - Follow-up is too late!
- Transformative and deliberative evaluators:
 - Evaluation process and democratic values are inseparable



Expanding typology of forms of use/consequences

- Accountability/control
- Learning/development
- Enlightenment
- Strategic
- Tactical
- Symbolic
- *Process*
- *Constitutive*



Constitutive consequences

- Content
- Time
- Social relations
- World view
- Impact on other methodologies or forms of knowing

Constitutive consequences

(example from study of test system of Danish for immigrants,
Dahler-Larsen 2012)

- Content: *Teaching to the test*
- Time: *No time for broader helping role*
- Social relations: *Weak students? Or slow?*
- World view: *"Effectiveness" or language in use*
- Impact on other methodologies or forms of knowing: *Other forms of evaluation ignored*

Changing patterns of use:

Paradox



- As the contours of the utilization problem changes (non-use not main problem)...
- New consequences of evaluation are discovered which perhaps make the very concept of "use" obsolete or insufficient



Outline

- Evaluation: What are we talking about?
- Five issues
 - Popularization
 - Outcome orientation
 - Systematization of evaluation
 - Changing relations: evaluation – research
 - Changing patterns of utilization
- Summing up



Summing up:

- These issues are important:
 - Popularization
 - Outcome orientation
 - Systematization of evaluation
 - Changing relations: evaluation – research
 - Changing patterns of utilization
- Each of them: Tension and paradox are not likely to disappear! (I hope!)



Summing up:

- These issues are important:
 - Popularization
 - Outcome orientation
 - Systematization of evaluation
 - Changing relations: evaluation – research
 - Changing patterns of utilization
- Each of them: Tension and paradox are not likely to disappear! (I hope!)
- Congratulations, CEval! 10 years!



Literature

- Dahler-Larsen, P. (2012): *The Evaluation Society*. Stanford University Press
- Dahler-Larsen, P., Ozga, J., Segerholm, C., & Simula, H. (Eds.) (2011). *Fabricating Quality in Education: Data and Governance in Europe*. London and New York: Routledge.
- Dahler-Larsen, P. (2010). Defining Quality in Evaluation. In *International Encyclopedia of Education 3rd Edition*, Penelope Peterson, Eva Baker and Barry McGaw (eds.), (3rd ed.) Pergamon Press.
- Dahler-Larsen, P. (2009). Learning-Oriented Educational Evaluation in Contemporary Society. In Ryan, K. E., & Cousins, J. B. (Eds.), *The SAGE International Handbook of Educational Evaluation* (p. 307-322). Los Angeles: Sage Publications.
- Boyle, R., Breul, J. D., & Dahler-Larsen, P. (2008) (Eds.), *Open to the Public. Evaluation in the Public Sector*. New Brunswick, New Jersey: Heinemann.
- Dahler-Larsen, P. (2006). Evaluation after Disenchantment: Five Issues Shaping the Role of Evaluation in Society. In Shaw, I., Green, J., & Mark, M. (Eds.), *The Sage Handbook of Evaluation* (p. 141-160). Sage.
- Dahler-Larsen, P. (2005). Evaluation and Public Management. In Ferlie, E., Lynn, L., & Pollitt, C. (eds.), *The Oxford Handbook of Public Management* (p. 615-642). Oxford: Oxford University Press.

■