

# **Theory based evaluation seminar**

**Athens, June 21-23 2017**

**Frans L Leeuw**

**Maastricht University & Applied Research Center  
for Security & Justice, the Hague, the Netherlands**



# Frans L Leeuw

► *Professor, Law, Public Policy and Social Science Research, Maastricht University & Nat Institute Applied Security & Justice Research, The Hague, the Netherlands*

*Former:*


- *Professor Evaluation Studies, Utrecht University,*
- *Ass Professor Leyden University,*
- *Director Program Evaluation & Performance Auditing, Netherlands Court of Audit,*
- *Dean Netherlands Open University,*
- *Chief Inspector, Education Inspectorate,*
- *Fulbright Scholar U of North Carolina at Chapel Hill, USA*
- *Advisor World Bank, EU and other organizations*
- *Former President of The European Evaluation Society and the Netherlands Evaluation Society*



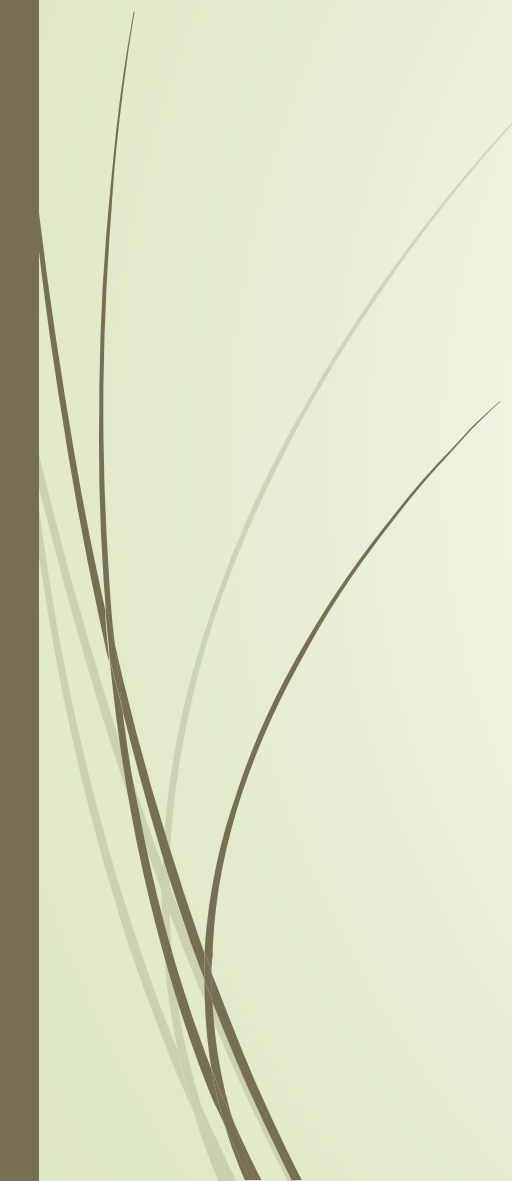
## **Introduction June 21 till 14.30**

Results from questionnaire & what to do

Sharing practical experiences and  
learning a la World Bank's Eval Group  
Blogs &  
Relevance of and challenges for TBE



Results from the questionnaire: what are your expectations and goals we have to take care of?



"Difficulties to construct or reconstruct of the programme theory and the logic of intervention in correlation with this. Difficulties in conceiving logical models of change in relation with programme theory"

"- adapting the public procurement legislation to the specificity of ToRs  
- identifying the most suitable indicative evaluation methods for different themes and questions"

TBIE is a challenging methodology especially in the framework of ERDF interventions, where its application has not been used widely  
Ensuring data collection, quality and adequation of data with needs. Identifying TBIE scope, enabling factors.

Defining use of evaluation process and findings.

I have still no experience of using TBIE, but as I was given responsibilities in managing evaluations and writing ToR, it worries me what to demand from evaluators when evaluating programs mainly when this programs allow differentiations in interventions, as those of LCSD.

Lack of knowledge and experience of the technical and methodological aspects of TBIE, having found very hard to draft ToR for TBIE and assess proposals (and evaluation outputs) technical quality, specially in a national market where it seems practitioners lack some of that knowledge and experience as well.

Theory based impact evaluation is not commonly used commonly in Structural Funds in Hungary yet. Therefore I would like to collect more information on the issue and about its possible practical implementation.

Being able to judge which parts of the planned impact evaluation are suitable for TBIE and how best to plan for and design a TBIE, including writing TORs.

Our biggest difficulty as the Coordinating Body for ESIF 2014-2020 is that we have no experience with evaluations up to this point (writing ToRs and management of the evaluation process from start to end). Our biggest challenge is the responsibility we have for the evaluations at the level of the Partnership Agreement, of which several are planned to be impact evaluations (counterfactual or theory-based).

Challenge on how to do the best and appropriate ToR for TBIE for the Luxembourgish ERDF-OP

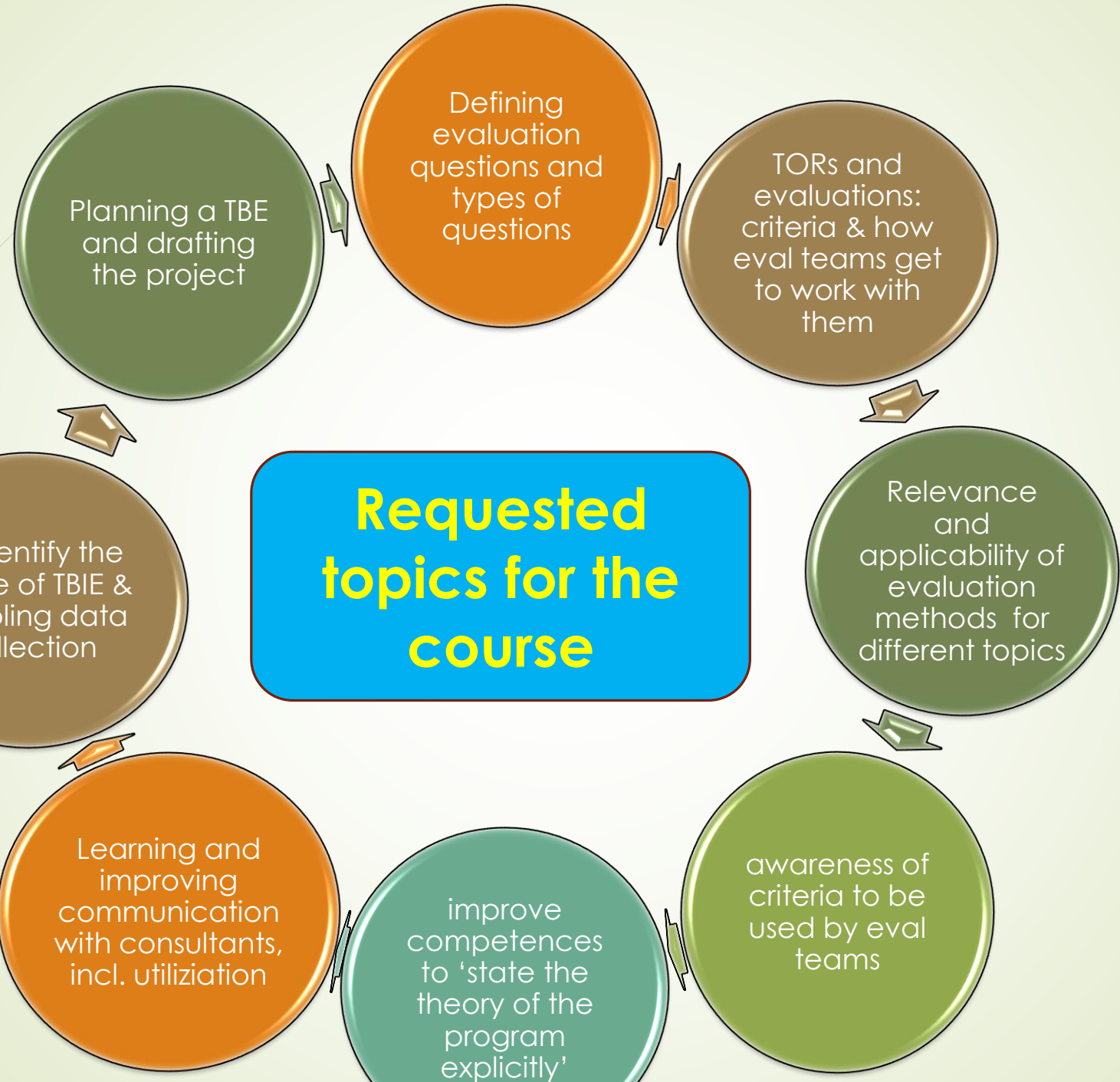
"-The selection of the proper method depending of the type of intervention.

-Combination of different methods in order to achieve an qualitative and quantitative evaluation

-Planning and designing the proper questions

-Defining the judgement criteria"

**Requested topics for the course**





## **General Introduction 14.30- 17.00**

Where did TBE start

Approaches, Mechanisms & examples of theories

Complexity and how to handle it

Adjudicating between rival theories

Role of theories in formulating policy programs

Criteria to assess and improve program/intervention theories

## 22 June Morning

**Presentation of the first TBE, planned by Tecla Livi on Combating House Poverty, role of Home Agencies /Italian Metropolitan Cities National OP.**

➤ **Presentation of the second TBE, planned by Oliver Swab on the planned evaluation of the Investment & Jobs Priority under the Berlin ERDF OP..**

**General discussion on both , ins and outs, opportunities & challenges, learning, management, utilization....**



# June 22 afternoon

- Five working groups; discussing your OWN PLANNED TBE'S

The discussion will centre on

- defining an appropriate theory of change/policy theory for the evaluation;
- identifying the main mechanisms involved in the intervention giving rise to effects and what is known about these mechanisms;
- defining the methodological route to follow to articulate, test and improve the theory and
- considering the research evidence and data likely to be available.

Each of the groups will be moderated by an expert and will elect a representative to report back on the outcome of the discussion.

# June 23 morning

- ▶ Reporting back by the 5 working groups on planned TBEs & outcome of discussions & remaining issues and discussion
- ▶ Next:

The six key principles of a theory-based impact evaluation are:

- ~~1. Map out the causal chain (programme theory)~~
- ~~2. Understand context~~
- **3. Anticipate heterogeneity**
- **4. Rigorous evaluation of impact using a credible counterfactual**
- **5. Use mixed methods and do rigorous analysis**
- **6. Results visualization & transfer / use focused**

**These items will also be included in this part**

o different types of data (sources) and TBE (Leeuw, 2016), including **Big Data**

o Attention will be paid to a **number of pitfalls** when working with TBE and traps to avoid

- **New challenges and opportunities** for TBE amongst others related to the digital arena



## June 23 afternoon

- Continuation of discussion before lunch
- Links with Terms of Reference training of last year
- Wrap up, conclusions and looking ahead

So far so good for the  
program of the next  
days.....

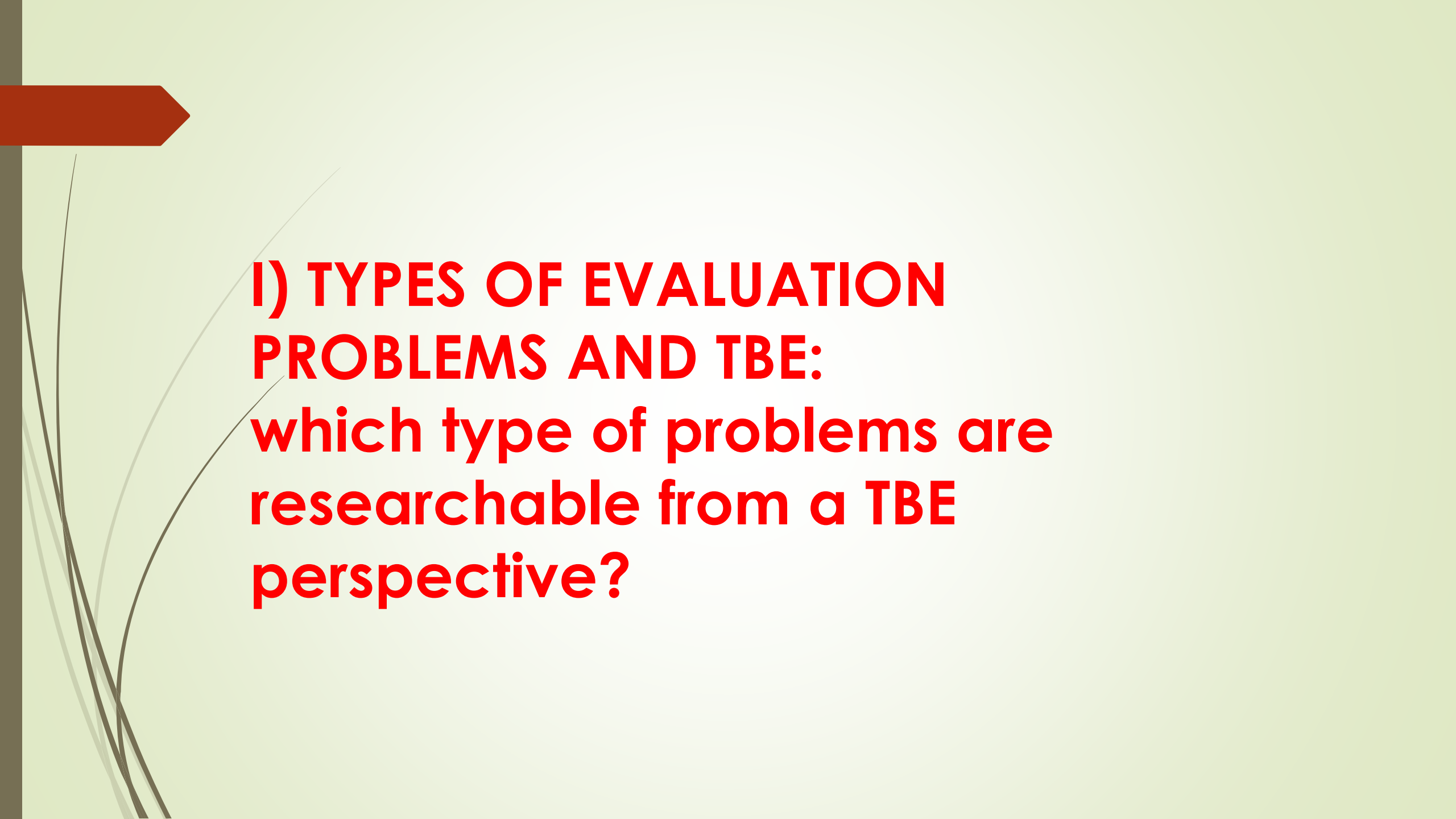


any suggestions, questions

...



***Let's start!***



**I) TYPES OF EVALUATION  
PROBLEMS AND TBE:  
which type of problems are  
researchable from a TBE  
perspective?**

# typology of research problems

## **Type 1: Descriptive research problems:**

*Type 1.1. Research problems focusing on exploring and describing the state of affairs of phenomena: How do people and small and medium businesses handle their (legal) conflicts? How many people are illegally downloading music, movies and books?*

*Type 1.2. Research problems focusing on exploring or making comparisons (of any kind): How does a sample of countries score on indicators a-k of the rule of law and human rights?*

*Type 2: Exploratory research problems: what is going on in the world of regulating internet of things, ISP's and central governments?*

## **Type 3: Causal (or explanatory) research problems:**

*Are web based information campaigns trying to reduce the level of digital piracy changing consumer choice behaviour (i.e. prevent and/or reduce illegal downloading of –e.g.—books/music/movies) capable of doing that and how can this be explained?*

## **Type 4: Normative research problems:**

*Should governments stimulate the privatisation of all public transport? Should there a be policy on preventing and curbing fake news on social media?*



#### **Type 4: Evaluative research questions:**

*Type 4.1 Evaluative research problems of an ex ante character: Will the law on preventing commercialization of surrogacy be effective in terms of realizing its goals?*

*Type 4.2: Evaluative research problems focusing on implementation processes: To what extent is the Leyden-Oregon Parenthood Managing Intervention working according to plan?*

*Type 4.3: Evaluative research problems of an ex post character: What are the effects and side effects of (EU) investments in R & D between 2003 and 2013 in country Q and P?*

#### **Type 5: Design-oriented research questions:**

*Type 5. 1. Research problems regarding the development of designs: How can common property regimes / common pool resources (CPR) be organized in such a way that they contribute in an efficient and effective way to the sustainability of natural resources?*

*Type 5.2. Research problems regarding the piloting of designs: How can a CPR-design focused on preventing deterioration of fish stock in the Mediterranean be tested and what are the results?*

## Table: 4.1 Four Key Questions in Impact Evaluation

1. To what extent can a specific (net) impact be attributed to the intervention?
2. Did the intervention make a difference?
3. How has the intervention made a difference?
4. Will the intervention work elsewhere?



# the & Research Problems: what is doable and what is relevant?

- ▶ Descriptive: yes. Theories focus on variables to be described. Monitoring studies often are (largely only) descriptive. Trends studies too. Theories also needed for explanation of findings.
- ▶ Exploratory: yes, to some extent. If the problem is Unknown Unknown, chances are small that theories will help, but better something than nothing. Subsumation may contribute.
- ▶ Explanatory (causal): yes, certainly.
- ▶ Evaluative: yes, certainly
- ▶ Design: yes. Design problems are a kind-of evaluative problems but then forward looking



How To Distinguish Adequate From  
Less **Adequate** Empirical Research  
Problems?

# Inadequate formulations of RP's

- **Failure 1: working on ill-formulated research problems, if not erroneous research problems**

[vague terms, wrong background knowledge]


- **Failure 2: working on problems lacking clarity and/or brevity**

[26 sub-questions; no logical relationships; 'bags of questions']


- **Failure 3: working on tautological research problems**


[ problems answered by definition]

- **FAILURE 4: WORKING ON NORMATIVE PROBLEMS WITH NO EMPIRICAL DIMENSION.**



## Practical suggestions:

- Understand the types of research problems that exist; “they can help you think about what type of question(s) you are asking” (White, 2009:52);
  - Understand what type of research problem it is you plan to work on;
  - Understand the different failures with regard to problem formulation and know ways how to prevent them:
  - Know and check background knowledge against which you formulate the research problem and pay attention to what “is already known about the issue”.
- 



Understand that descriptive research questions often need to be answered before explanatory or evaluative ones can be asked and studied;

Be precise in terms of

- the scope of your problem;
- the definitions of concepts;
- the time frame for the study,
- the geographical location(s);
- the level of abstraction of the problem and
- the 'unit of analysis' (= natural persons, corporate actors, periods, places or others).



II)

## Theory-based evaluations: where and since when?

**“TBE regards the programme as a ‘conjunction’ of causes that follows a sequence. It follows the pathway of a programme from its initiation through various causal links in a chain of implementation, until intended outcomes are reached. The process is built around a ‘theory’ – a set of assumptions about how an intervention achieves its goals and under what conditions.”**

(Stern et al, 2012)

*Theory Based Evaluation: A wealth of approaches and an untapped potential*

Marielle Riché

European Commission



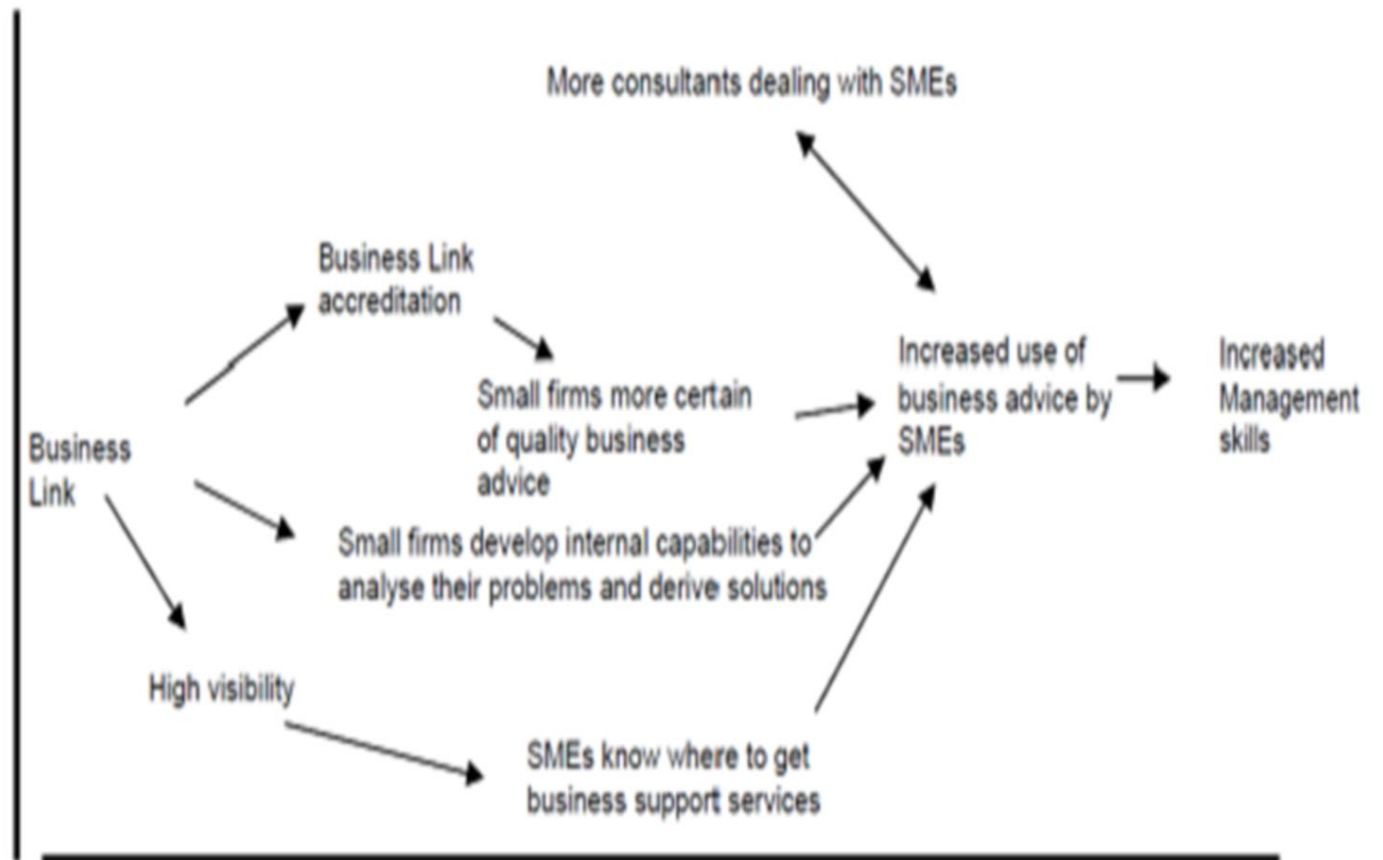
# Fields where tbe is 'active'

- ▶ Traffic (regulation)
- ▶ Education
- ▶ Labour force policies
- ▶ Crime & Justice (beh interventions, naming & shaming, penal sanctions, mediation, etc.)
- ▶ Social cohesion and ERDF: business links, support to enterprises
- ▶ Health policies
- ▶ Legal arrangements in general (from wills and contrats to truthcommissions and treaties...)
- ▶ Family policies
- ▶ Sport pathways to .....you can name it, incl. digital policies & Homo Digitalis..

# Appetizer: a (not too good!) example

- BL is a type of small business support activities that many (European) governments have implemented. It can be seen as a type of brokerage.
- BL activities are believed to increase economic productivity and job growth.

Figure 1.1: Programme Theory for Business Links





**Can you see what is missing  
in the reconstructed  
“theory”?**

# That implies that things like this

## The Ready by 21 Theory of Change



## Logical Framework (Log Frame)

The Goal refers to the overall problem we are trying to address

The Summary outlines the project's objectives: what it hopes to achieve and how.

|            | SUMMARY | INDICATORS | EVIDENCE | ASSUMPTIONS |
|------------|---------|------------|----------|-------------|
| Goal       |         |            |          |             |
| Purpose    |         |            |          |             |
| Outputs    |         |            |          |             |
| Activities |         |            |          |             |

What outputs are needed to achieve the Purpose?


What change or benefit we want the project to achieve so that the goal can be reached

How will the outputs be delivered or produced





***are not to be considered program theories.....***



# HISTORY

## 60s & 70s

Great Society Program LBJ / education training; pet theories/ Logical Framework [with “the Killer Assumption”-approach]; program logics; early beginning in Europe (Sweden, Netherlands, Germany, UK)...

## 1980s:

policy theory concept, learning from this type of work (double loop, single loop; “Can Governments Learn”-studies), link with utilization and ## of studies increasing

## 1990s:

Realist(ic) Evaluation started; And some other approaches (surfacing & challenging strat assumptions), more studies, journals became interested; first TOR / lookalikes asked for this work

## 2000's:

- methods more on the agenda, EU getting seriously involved; 'killer assumption' killed, TBE going global, contribution analysis arriving; more policy fields part of the TBE world; TBE in more parts of the world part of TORs; in some political arena's policy makers getting themselves to work with 'theory' ...and articulate their assumptions

## 2007- 2017

- EU/MS more and more active; guidance, helpdesk etc; complexity & TBE ; Pawson's Science of Evaluation incl different methods (Rameses), guidance & oversight books (Evalsed: EU; Nonie: Leeuw & Vaessen; Defid: Stern et al); stronger links with counterfactual evaluation, global spread continues...but also 'Die Unvollendete'-problems and the *Fake HandBags* problem; first moves towards 'mechanisms [compound] libraries & M-experiments (Ludwig et al).



### Scorbutic (scurvy) example:

RCT & Theory ***without knowing...dr James Lind***. By 1747 he had become surgeon of HMS Salisbury in the Channel Fleet, and conducted his experiment on scurvy while that ship was patrolling the Bay of Biscay



# III) Several Approaches

## The THEORY OF CHANGE APPROACH

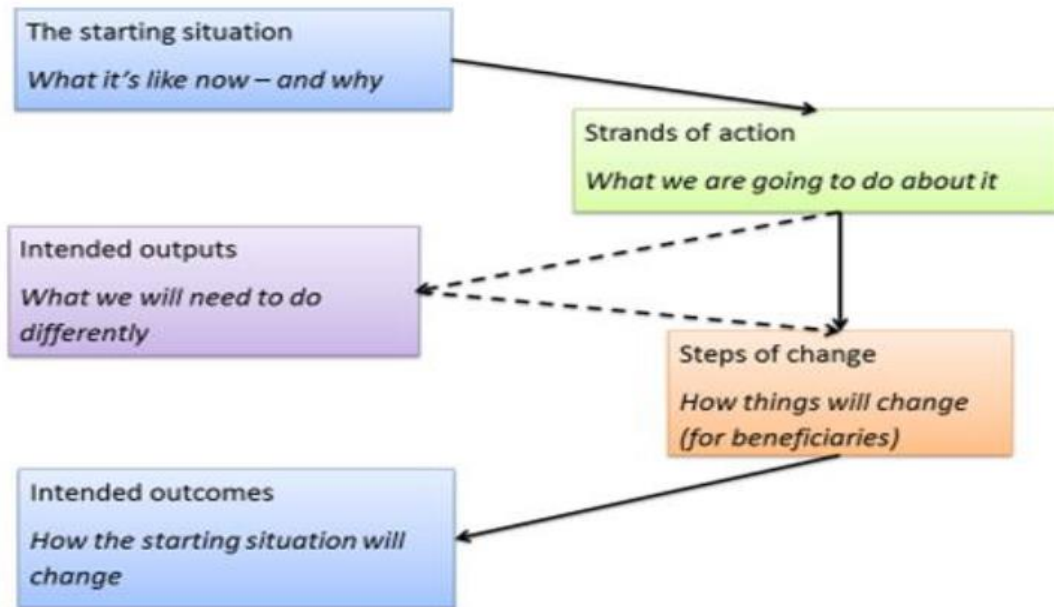


Figure 3 Characteristics of a theory of change

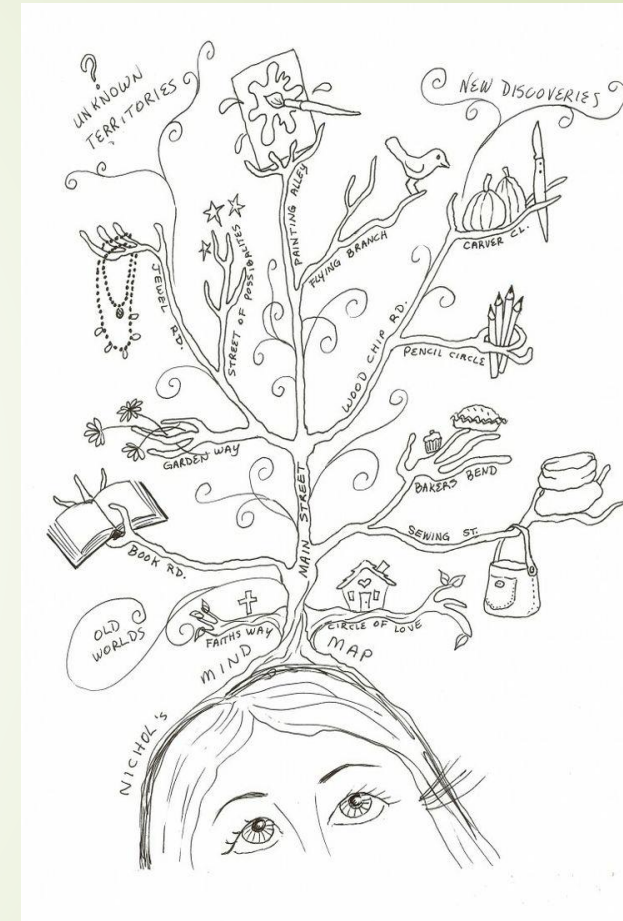
Toc articulates how a project or initiative or ... is intended to achieve outcomes through actions, taking into account its context.

## ***Intervention logic, intervention theory, program logic, program theory:***

- Articulates, broader than only in a change site, how interventions aim to realize their goals
- the **program THEORY** claims or tries to (and indeed) focus on Mechanisms ( and the relationship with context & outcomes), while the **program LOGIC** often is not more than boxes and arrows connecting goals, subgoals, 'activities' or 'measures' over a time line and expected outcomes.

**Aka weak  
vs strong  
approach**

**The Mental Map approach:** all this but only in the heads of people...



The **Realist approach** focuses on CMO's:  
will come to that later

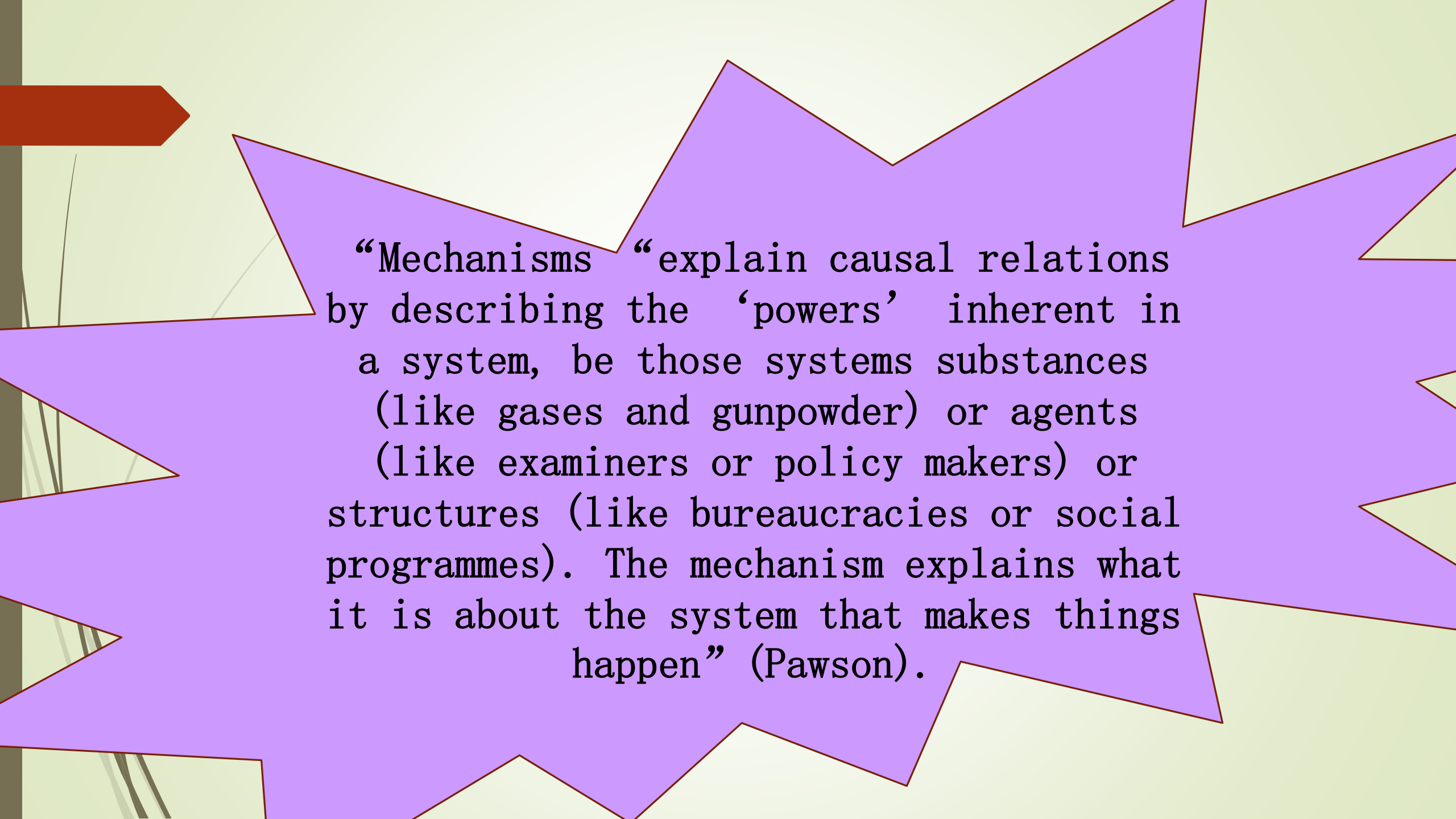
## **IV) A CRUCIAL ELEMENT IN THE PROGRAM THEORY ('STRONGER') APPROACH IS THE ROLE PLAYED BY MECHANISMS**

**Policies, measures, activities, interventions, messages etc ARE NOT EQUAL TO BEHAVIORAL, SOCIAL MECHANISMS.**

**A common mistake is for evaluators to conflate the term mechanism with program activity [ *Leading To Fake Handbags-evaluations* ]**

# What are mechanisms?

- Mechanisms are drivers/ engines, aka “cogs and wheels” that can be triggered by policy interventions and may help realize the policy goals. They are ‘behind’ societal and legal arrangements, like policies, contracts, treaties etc
- Mechanisms are defined as “underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest” .



“Mechanisms “explain causal relations by describing the ‘powers’ inherent in a system, be those systems substances (like gases and gunpowder) or agents (like examiners or policy makers) or structures (like bureaucracies or social programmes). The mechanism explains what it is about the system that makes things happen” (Pawson).

Examples “found” in recent (European) evaluations:

satisficing;

bounded will power,

reinforcement,

Incentives;

social capital;

the shadow of the future, crowding out,

cognitive and socio-neurosciences mechanisms like the fundamental attribution error; attachment, sensation seeking & arousal levels;

fight or flight, the slippery slope;

blame avoidance; performance feedback;

Reputational costs,

and the bandwagon effect.

# A few examples of TBEs

## Ex-post evaluation of the Europe for Citizens Programme 2007-2013

Final report



Article

## Are social mechanisms usable and useful in evaluation research?

Evaluation  
2016, Vol. 22(2) 209-227  
© The Author(s) 2016  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/1356389016643900  
evi.sagepub.com  
SAGE

**Erica Melloni**

Istituto per la Ricerca Sociale, Milano, Italy

**Flavia Pesce**

Istituto per la Ricerca Sociale, Bologna, Italy

**Cristina Vasilescu**

Istituto per la Ricerca Sociale, Milano, Italy

## Support to large enterprises

### Final Report Work Package 4

*Ex post evaluation of Cohesion Policy programmes  
2007-2013, focusing on the European Regional  
Development Fund (ERDF) and the Cohesion Fund (CF)*



## Figure 1: Main objectives of the Programme for Active European Citizenship

Source: Council decision 2004/100/EC

Promoting and disseminating the values and objectives of the EU

Bringing citizens closer to the EU institutions and encouraging more engagement with those

Involving citizens in reflection and discussion on the construction of the EU

Identifying links and exchanges between citizens from participation countries (particularly town twinning)

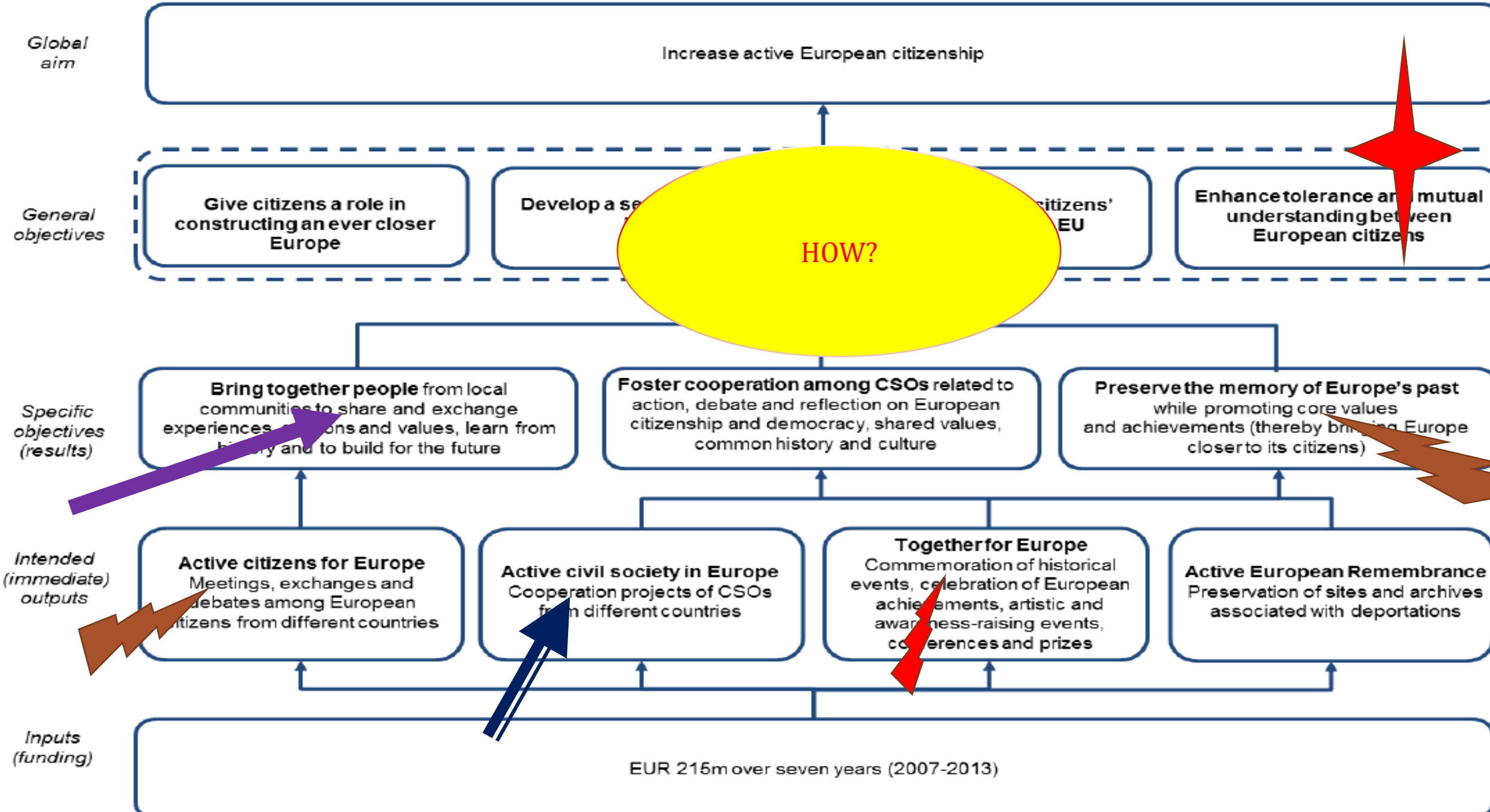
Stimulating initiatives by bodies engaged in the promotion of active and participatory citizenship

The Community Action Programme had a budget of EUR 72 million over the three years of its existence. It supported a wide range of activities in the field of active citizenship through two types of grants:

- Operating grants to co-finance permanent work with an aim of general European interest in the field of active citizenship
- Grants to co-finance specific actions.

In total, the Community Action Programme 2004-2006 funded over 30 organisations, as well as more than 250 projects by NGOs, associations and federations, and trade unions. Over 2,800 town twinning projects received funding from the programme.

Figure 2: ECFP intervention logic



**Table 2.** The working method: three examples.

| Exemplar case   | Policy outcomes  | Mechanisms  | Context features  | Contribution of the approach in informing policy                           |
|---|--|---|---|--|
| AlpEnergy project   | Reduction and postponement of Saint Denis citizens' energy consumption in non-peak periods   | <p>Instant feedbacks on energy consumption levels</p> <p>Competition and 'bandwagon effect' among citizens</p>  | <p>Very small mountainous area</p> <p>High sensibility towards renewable energy issues</p> <p>Local political agenda focusing on energy projects</p>  | Learning from good practices to disseminate findings to different contexts |
| Gender mainstreaming in committees and delegations of the European Parliament | Effectiveness of FEMM's activity in fostering the mainstreaming of the gender equality principle in the EP legislation   | <p>Reputation and/or actor's certification of FEMM's rapporteurs</p>  | <p>22 committees participating in the legislative procedure of the EP</p> <p>FEMM committee in charge of promoting gender equality and gender mainstreaming</p>   | Providing policy advice to enhance the effectiveness of gender policy      |
| Ex-ante evaluation of the ESPON 2014-2020 Programme                           | <p>Fostering a greater awareness, usefulness and use of territorial information for European and regional policy makers</p> <p>Streamlining the management of the programme and reducing administrative burden for beneficiaries</p> | <p>Perception of opportunity of different target groups to generate and maintain engagement to the Programme</p> <p>Certification of the newly established EGTC</p> <p>Repeated interactions among stakeholders favouring trust, mutual learning and commitment</p> | <p>Broad audience of potential stakeholders for the territorial evidence produced by ESPON</p> <p>Large composition of the steering committee of the programme – 32 States</p> <p>Heavy management burden and procedural complexities</p> | Support policy makers in the design and structuring of a new Programme     |

Mechanisms

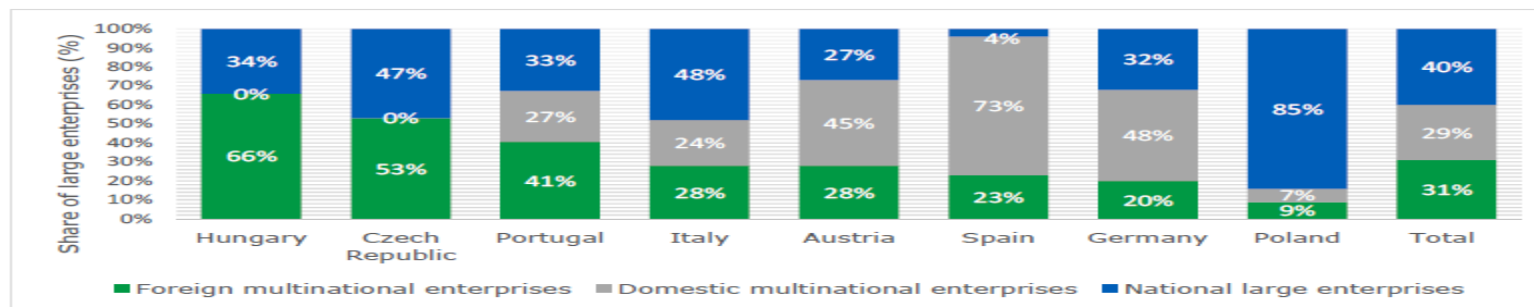
# Support to large enterprises

## Final Report Work Package 4

*Ex post evaluation of Cohesion Policy programmes  
2007-2013, focusing on the European Regional  
Development Fund (ERDF) and the Cohesion Fund (CF)*

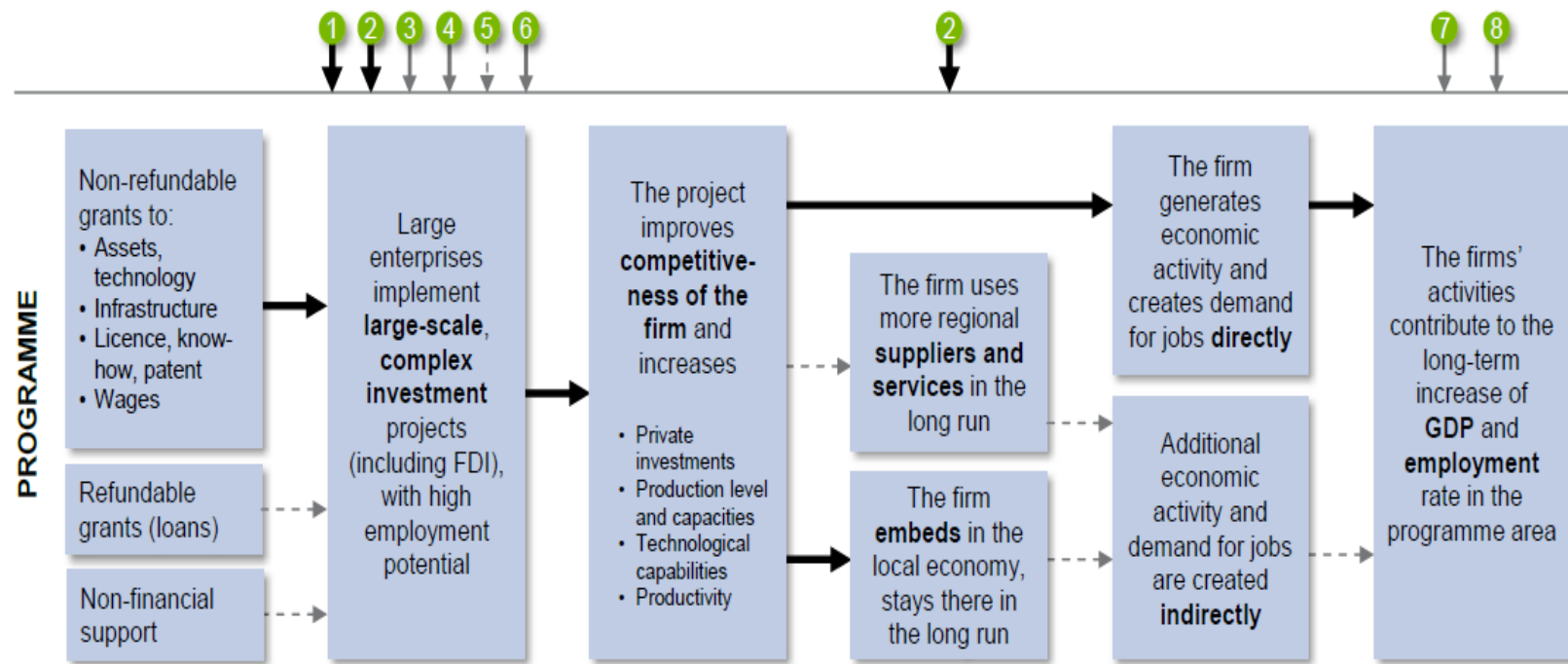
Cohesion Policy invested an estimated EUR 6 billion in support of large enterprises over the 2007-2013 period.

Figure 1: Distribution of large enterprises based on scope of operation and origin



Source: KPMG/Prognos (2016), based on the monitoring systems and public company databases.

## 7.1. Theory of Change 'LE1': Large-scale business investment



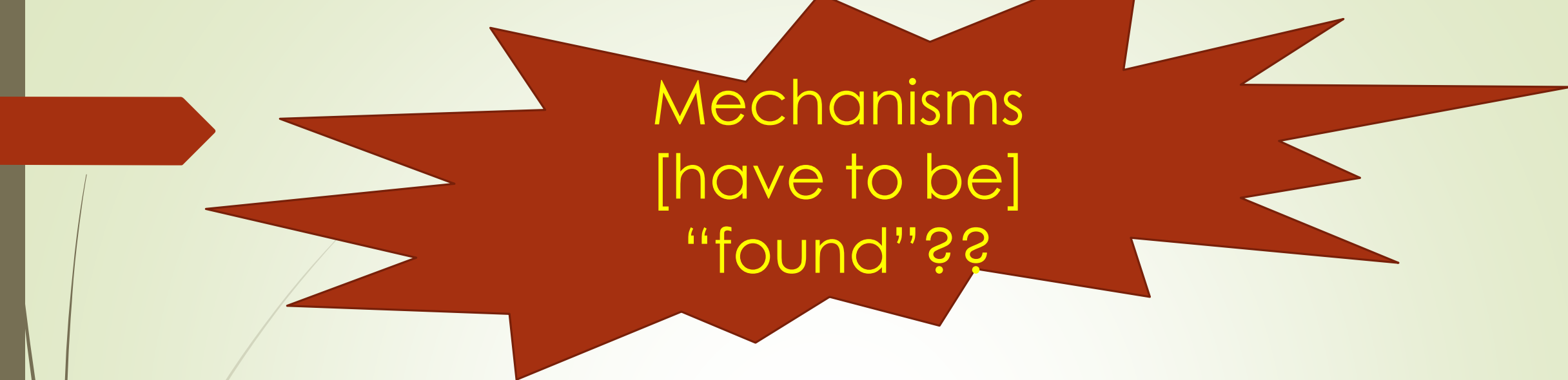
### Assumptions and external factors

1. Tax incentives are competitive (internationally)
2. Company strategy supports long-term stay in the country
3. Developed basic infrastructure (motorways, airport access, ICT infrastructure)
4. Business / industry „heritage“ is present in the area
5. Supportive local government (permits, procedures)
6. Labour market supplies labour in required number and qualification levels
7. Investment is large enough to influence the labour market
8. General economic conditions enable growth

### Indirect effects and wider benefits

- a. Increased demand for „quality“ jobs in the area
- b. Attracting other companies/investors in the region
- c. Improved local transportation and ICT infrastructure
- d. Improved social infrastructure (education, culture etc.)
- e. Spillover of improved business practices, skills, knowledge, R&D and efficient technologies (local enterprises)
- f. Spread of improved working culture (working conditions, wage levels, timely wages, values, stability etc.)
- g. Greater workforce mobility
- h. Crowding-out of SMEs from labour market (skilled labour)
- i. Distort market equilibrium (effect on SMEs & non-supported)





# Mechanisms [have to be] “found”??

- ❖ Mechanisms are usually hidden (so they have to be ‘articulated’)
- ❖ Mechanisms are part of the ‘stratified reality’ (so they may be ‘layered’)
- ❖ Mechanisms are sensitive to variations in context (but some contextual variations are more important than others)
- ❖ Mechanisms generate outcomes (though also ‘effets pervers’)

# Mechanisms are (usually) hidden

Pawson (2013)

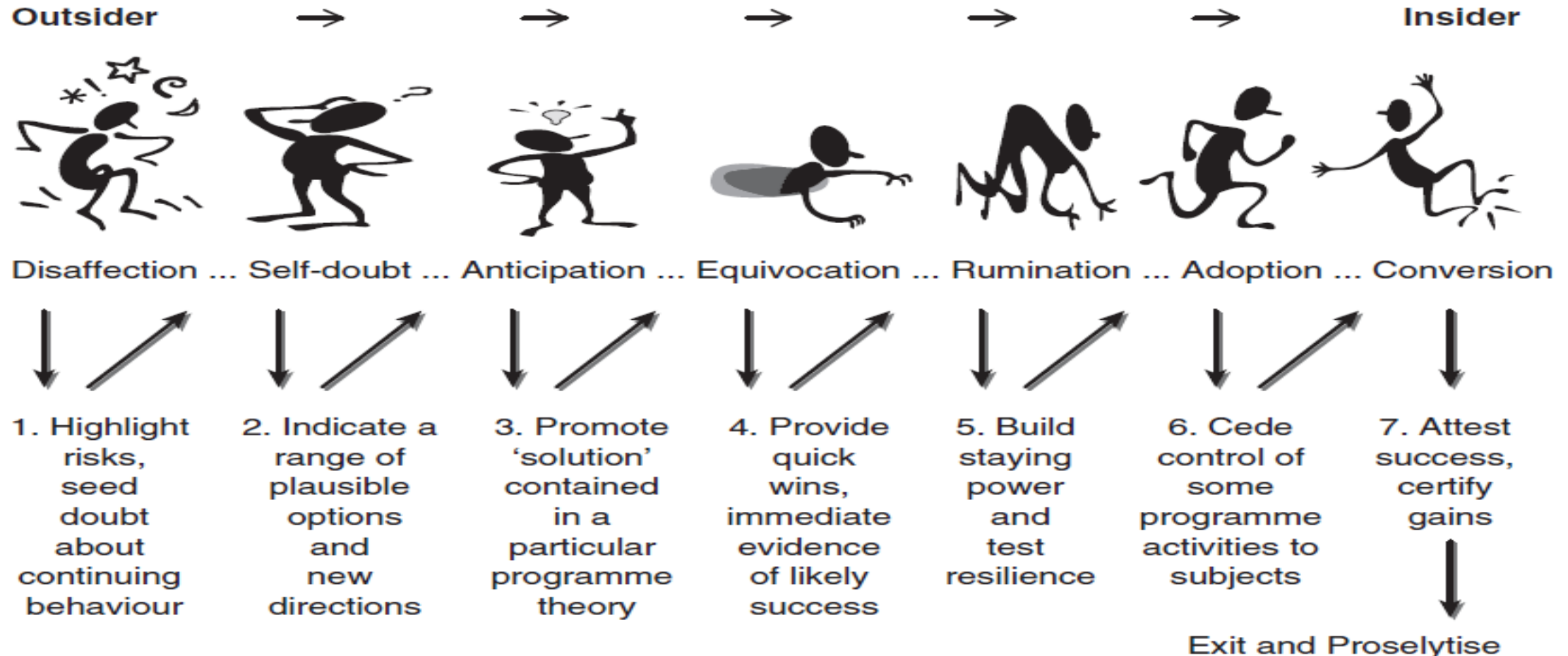


Figure 6.3 A Conceptual Platform for Behavioural Change Interventions

# CRUCIAL IS UNPACKING THE POLICY BLACK BOXES BY SEARCHING FOR *MECHANISMS*





## CUES TO FIND THE 'HIDDEN MECHANISMS'?

- Methodological 'rules' / guidance [ coming soon!]
  - Reflect on the character of the policy program, intervention, law in terms of these and similar principles:
    - What about (whose) **resources** influenced by the policy?
    - What about (whose) **relationships** (social capital) influenced?
    - What about (whose) **power** influenced
    - More in general: what about **VICTORE**?
-

# VICTORE

## Box 3.1

### A complexity checklist

- 1 Volitions. Map the choice architecture of the programme. What choices do subjects have to make to achieve the ambitions of the programme? These choices will vary significantly across the totality of subjects, so the map should cover the full range and balance of volitions. Minds tend to be changed slowly so the volition map should also sketch the pathways of persuasion, the sequence of choices a subject has to make in moving from outsider to insider status.
- 2 Implementation. Map the implementation chains of the programme. Programmes come to life over many months and years and a great diversity of 'traffic' flows

*(Continued)*

*(Continued)*

through them. Implementation maps might begin to chart: flows of resources; chains of responsibilities (individuals and institutions); reception and transmission points for subjects; as well as the different theories of change that lie behind each stage, strategy and tactic.

- 3 Contexts. Map the pre-existing contexts in which the programme is embedded. Consider for whom and in what circumstances the programme might work. Contexts vary from the micro to the macro, so the map might include profiles of stakeholders and their characteristics; the interpersonal exchanges through which the programme is delivered; the organisational settings in which it takes place; the wider societal location of the programme.
- 4 Time. Map the history of the family of programmes of which the intervention under study is a member. What has happened previously will shape what happens next. Temporal mapping might include: previous experiences of programme subjects and communities on similar interventions; previous experiences of stakeholders in delivering similar interventions; the successes and failures of previous attempts, of whatever kind, to address the given policy objective.
- 5 Outcomes. Map the monitoring systems that are likely to be applied and have been applied to programmes like the one under study. Consider which measures are likely to be contested, how stakeholders might differ in their interpretations, and whether behaviour might change as a result of being monitored rather than as a result of the intended action of the programme.
- 6 Rivalry. Map the pre-existing policy landscape in which the programme is embedded. Other, contiguous programmes and policies may share or oppose the ambitions of the intervention under study and can override the actions of stakeholders and subjects under study. Consider how generic implementation strategies such as 'continuous improvement' will continue to modify the delivery of the programme.
- 7 Emergence. Map the potential emergent effects, long-term adaptations, societal changes and unintended consequences associated with the programme. Consider whether the spread and duplication of the programme might blunt its effectiveness. How will the programme be able to maintain a balance between recruitment, retention and exit?

# Mechanisms are part of the 'stratified reality' / 'complex / complicated world' (so they [often] may be 'layered')

## Evaluation

<http://evl.sagepub.com/>

Using complexity-consistent theory for evaluating complex systems  
Gill Westhorp  
*Evaluation* 2012 18: 405  
DOI: 10.1177/1356389012460963

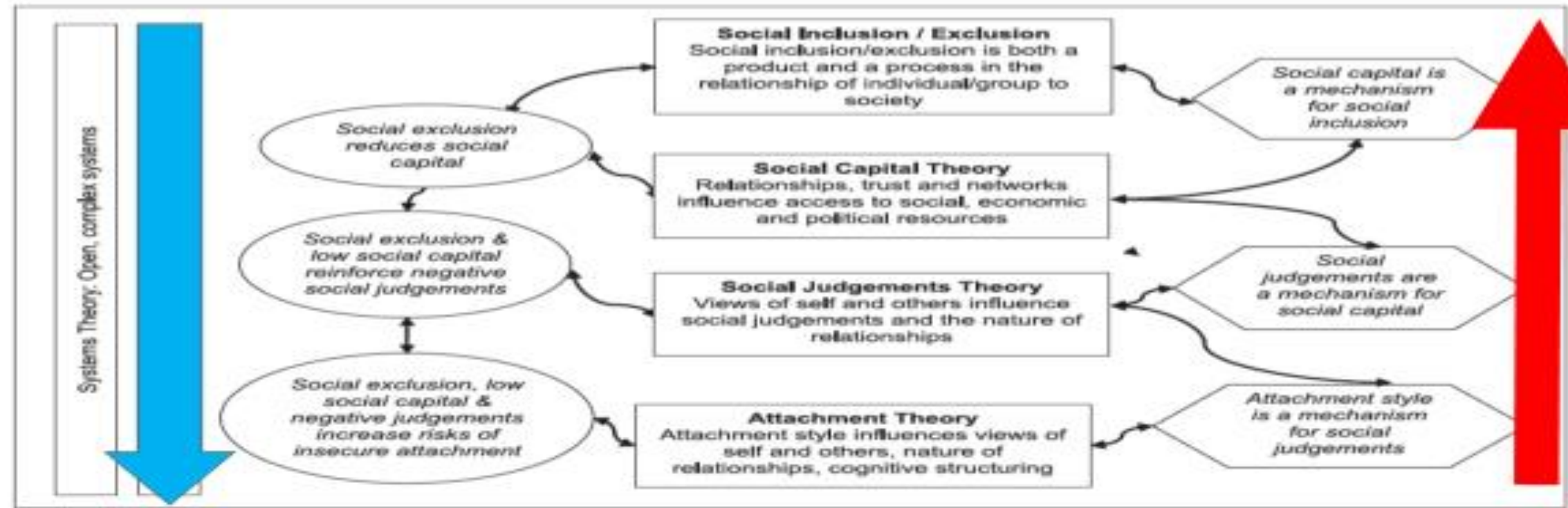


Figure 1. Theory map.

The theory map places attachment theory as the bottom rung of a 'theory ladder'. It posits that attachment style influences the nature of social judgements that an individual makes, which in turn influences the nature of relationships that they form, which in turn influences the extent and nature of social capital available to them, which in turn influences their life outcomes at the societal level (employment status, housing status and so on). That is, it posits attachment style as a mechanism generating differences in social judgements, which in turn act as a mechanism generating differences in social capital, which in turn act as a mechanism generating social inclusion and exclusion. It should be noted here that there is no claim that these are the sole mechanisms – or even, necessarily, the primary mechanisms – generating the outcomes at the next level. The claim is simply that these *are* mechanisms and that they therefore, in at least some contexts, contribute to those outcomes.

However, it should also be noted that the generative process operates downwards as well. That is, social and economic status and the social norms at play within particular groups influence the social capital available to members of those groups, which influences their experiences of seeking assistance through others, which influences both the nature of their relationship experiences and the nature of future social judgements they make, and their relationship experiences in turn have the potential to influence the modification of attachment style over the life course. This bi-directional influence is consistent with concepts of emergent order and downward causation in realist systems theory:

The idea of downward causation is anti-reductionist and posits . . . that influences can be expected to occur in both directions – upward from subsystems and downward from the whole; every level constrains others. (Mark et al., 2000: 156)

**Mechanisms are sensitive to variations in context (but some contextual variations are more important than others)**

*position 1: Contexts are more important than M's*

*position 2: Some contexts are more important than (some) others*

*position 3: demi-regularities & path dependencies tell us that contexts are not that important. Contexts, despite the claims of being different (all the time) usually are not.*

*Position 4: the only thing that counts are bio-social-psycho generative mechanisms “.*

***MECHANISMS GENERATE OUTCOMES (THOUGH ALSO  
'EFFETS PERVERS')***

OUTCOMES CAN BE EVERYTHING, NOT  
ONLY THE DESIRED OR EXPECTED  
POLICY RESULTS ... THEY CAN BE  
RATHER STRAIGHT FORWARD OR  
MULTIPLEX, MULTILEVEL  
(LIKE.....NOP...)

**METHODS TO ARTICULATE,  
TEST  
AND IMPROVE  
PROGRAM/  
INTERVENTION THEORIES**

# ➤ POLICY SCIENTIFIC APPROACH ( A K A THE REALIST)

- Reading and hearing between the lines;
  - VICTORE;
  - And Argumentational analysis focusing largely on ‘warrants’
- > All geared to find **hidden and unhidden mechanisms**, context information etc.
- Criteria to test the program theories are:
  - Logical Consistency and Empirical content (“validity”)
  - Central are also: CMO’s and making use of repositories.
-



## ➤ ELICITATION APPROACH

- As policies and programmes are developed and implemented by organisations, the 'mental models' or 'cognitive maps' of people in these organisations, i.e., their theories, are important for understanding the anticipated impact of their policies or programmes. The emphasis should therefore be placed on organisational cognitions.
- One of the central questions is the relationships between these cognitions and the results of organisations. All stakeholders have 'cognitions' (theories) about the organisation and its environment. These maps partly determine their behaviour.
- Their content concerns the organisational strategies, their chances of success, the role power plays, their own roles and the relationships with the outside world. Parts of these maps or theories are implicit and are tacit knowledge.

### Techniques:

- Look at the concrete record of strategic intentions, through, for example, a study of the documentation which is designed to direct behaviour;
- Look at decision-making in action; get involved in the organisation (an anthropological observer approach). Watch decision-makers, listen to stories;

- Work with managers on strategic breakdown situations. Become immersed in the thinking and the social process of 'strategic fire fighting';
- Use well-designed trigger questions in interview situations so that 'theories in use' can be detected.
- Follow interviews with feedback to individuals and to the team. The 'elicitation cycle' is built on responses to designed trigger questions. The process uses six techniques:

- o Create an open-ended atmosphere in the interview;
- o Do away with formal language and create a 'playful' atmosphere in which it is easier to deviate from the formal phraseology and the official script;
- o Do 'set the interviewees up against themselves';
- o Create dialectical tension by asking the interviewees to adopt unusual roles;
- o Listen very carefully for internal inconsistencies in what is being said;

Apply data/content-analysis programmes or other text analysis programmes to the interview reports and documents; and

Confront the results of these content-analysis activities with relevant (social) scientific research.

# ➤ CONTRIBUTION ANALYSES

- Causal inferences (claims about causes and effects) can be made by testing the theory of change for an intervention against what has been observed, and assessing the influence of other external factors.
- This figure below sets out seven iterative steps in a contribution analysis. Each step in this process adds to the contribution claim and helps address the weaknesses identified at the previous step. The result of a contribution analysis should be a reasonably credible “contribution story” (i.e., the narrative description of the theory of change and its supporting evidence).

**Step 1: Set out the cause-effect issue to be addressed**

**Step 2: Develop the theory of change**

**Step 3: Assess the resulting contribution story**

**Step 4: Gather the existing evidence on the theory of change**

**Step 5: Reassess the contribution story and challenges to it**

**Step 6: Seek out additional empirical evidence**

**Step 7: Revise and strengthen the contribution story**



# A SPECIAL CASE: ADJUDICATING BETWEEN RIVAL THEORIES

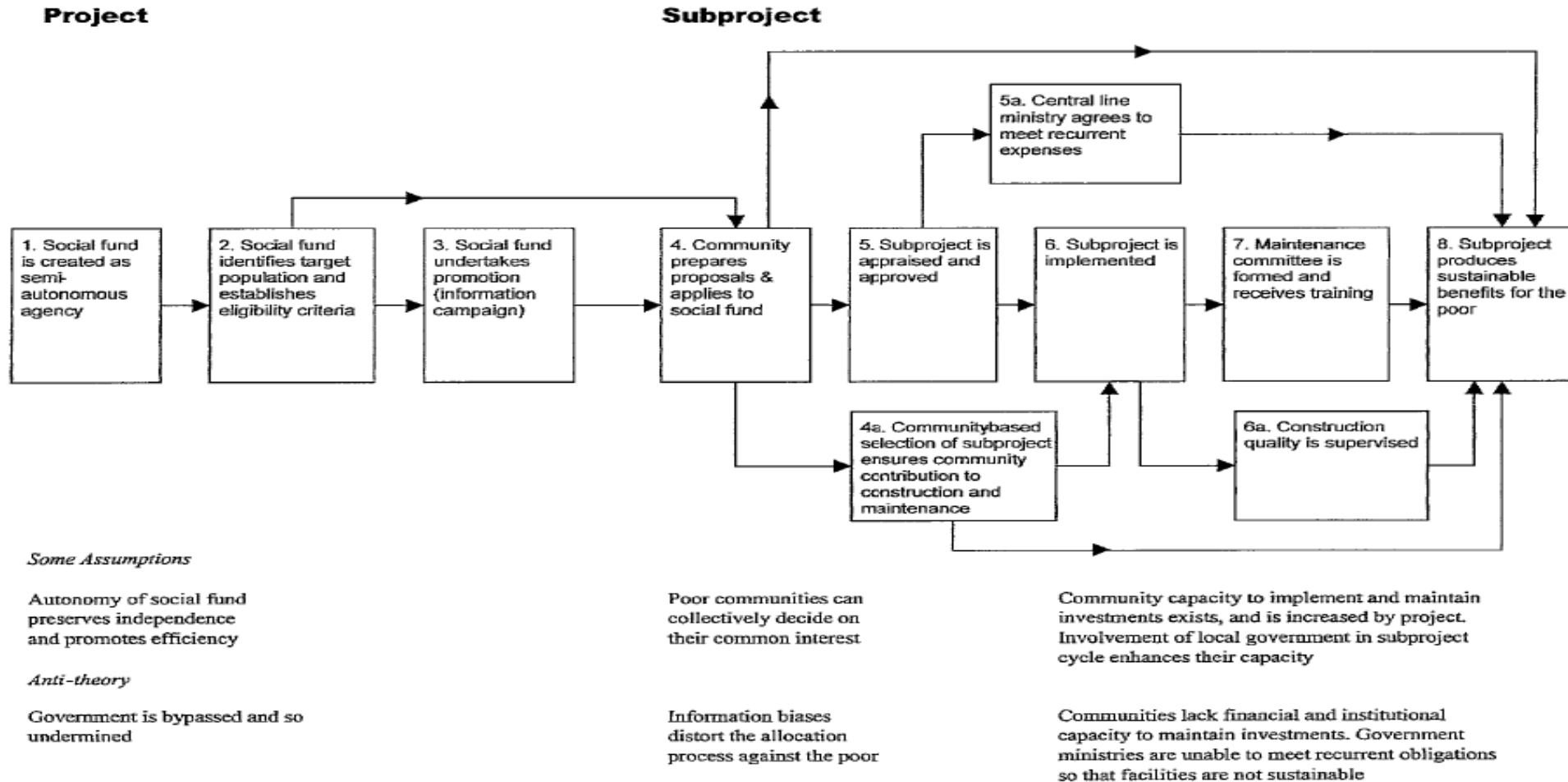


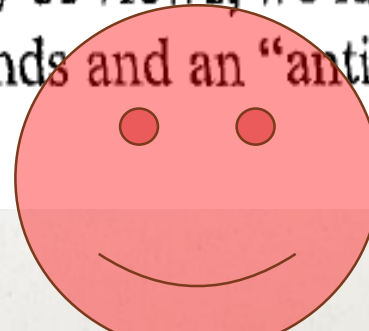
Figure 1. Theory of social funds: community-based model.

Part of assessing the empirical content / validity of this program theory was the handling of the problem “whose program theory” was articulated.

The question may arise as to “whose theory,” as there may be competing views about how a program works. That is certainly the case for social funds. For example, some believe that the social funds’ invitation of subproject proposals from community groups results in the selection of projects which mainly benefit the local elite (a phenomenon known as “elite capture”). Or, rather than building government capacity in planning and implementing small-scale infrastructure projects, social funds might undermine such capacity by taking both resources and functions from government agencies. Given this diversity of views, we identify both a “theory” derived from the writings of the supporters of social funds and an “anti-theory” based on the views of critics.

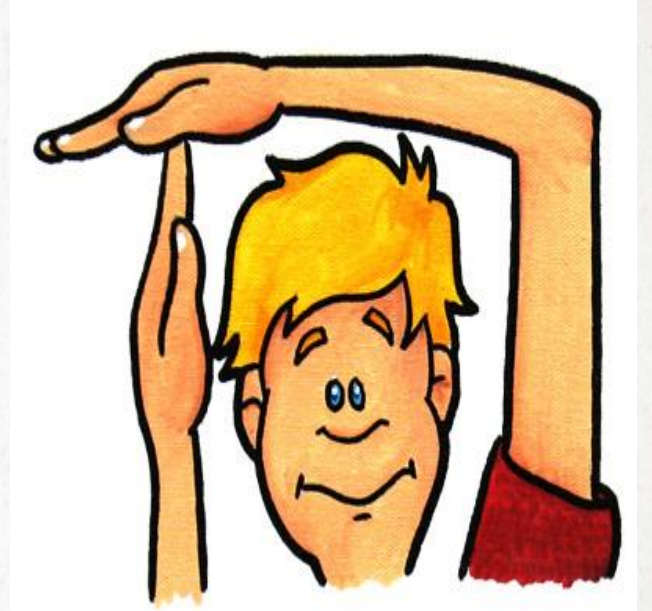
2

1



*just a thought*

**TAKE A  
BREAK**



**A LITTLE BIT ANOTHER STORY:**

**WHAT IS THE ROLE OF THEORIES  
IN DEVELOPING AND  
IMPLEMENTING POLICIES?**

# THERE ARE DIFFERENT ROUTES / OPTIONS

## ➔ **OPTION 1) Developing a policy & theories on the basis of as sound as possible knowledge:**

- ➔ Suggestions on which mechanisms the policy should 'build', or: which mechanisms policy makers can 'trust'
- ➔ These suggestions can be derived from knowledge repositories like Campbell, Cochrane, 3ie, the many What Works repositories, but – with care– also Google Scholar
- ➔ Behavioral insights team (BIT, a.o. UK) can help ...
- ➔ Cases: Perla Maldita & drugs policy



|                    |   |
|--------------------|---|
| <b>Messenger</b>   | we are heavily influenced by who communicates information   |
| <b>Incentives</b>  | our responses to incentives are shaped by predictable mental shortcuts such as strongly avoiding losses |
| <b>Norms</b>       | we are strongly influenced by what others do  |
| <b>Defaults</b>    | we 'go with the flow' of pre-set options  |
| <b>Salience</b>    | our attention is drawn to what is novel and seems relevant to us  |
| <b>Priming</b>     | our acts are often influenced by sub-conscious cues   |
| <b>Affect</b>      | our emotional associations can powerfully shape our actions   |
| <b>Commitments</b> | we seek to be consistent with our public promises, and reciprocate acts                                 |
| <b>Ego</b>         | we act in ways that make us feel better about ourselves   |

To what extent will public information campaigns focused on arousing fear within (future) cocaine swallows help reduce trafficking?

**Would you risk your life?**

**BOLLITA**  
Perla Maldita

LEA  
LUNA  
BLOU

The basic assumption behind the policy was that when information about the dangers of the internal concealment method is made public on airports through folders, affiches, advertisements etc, this information will reach future, potential (young) swallowers and their significant others;

If this information reaches the potential swallowers /significant others, they will experience fear;

The more fear they experience, the larger the probability that they will act upon the information diffused;

Acting upon the information received will lead to refraining from swallowing.

*The evidence is nil to zero.....*



➔ **OPTION II) DOING EX ANTE EVALUATIONS AS SOON AS THE OUTLINE OF A POLICY IS 'AVAILABLE'**

**Prospective evaluations** a la GAO USA: comparing the likelihood of success of policy (strand/ law/ regulation....) A and B on the basis of existing evidence on Mechanisms (and contexts) from repositories and/or through new research.

*The Surfacing & Challenging Strategic Assumption Approach* (Mason & Mitroff; applied in Business research) can be recommended.

## OPTION III) GENERALIZING FROM IMPLEMENTATION STUDIES AND FINDING (DEMI)REGULARITIES

### *Generalizing from implementation research<sup>6</sup>*

Over the years (synthesis) studies on implementation failures and factors have been published in, among others, the following domains: health (Carrol et al., 2007; Greenhalgh et al., 2004); energy (efficiency) (Harmelink et al., 2008); prevention (Durlak and DuPré, 2008); and crime and justice (Barnoski, 2004; Lipsey, 2009; Nas et al., 2011).<sup>7</sup> These and other studies focus on implementation fidelity and its underlying causes. Implementation fidelity is the degree to which an intervention or program is delivered as intended (Carrol et al., 2007). Nas et al. (2011) studied 20 recent Dutch evaluations of behavior-modification interventions and sanctions focusing on crime prevention and reduction. In order to find out which implementation factors and failures are prominent, they synthesized these 20 evaluations. Table 2 lists the ‘hit list’ of implementation factors and failures that were found.

**Table 2.** Implementation failures reported in 20 process evaluations of interventions focusing on preventing/ reducing crime in the Netherlands (Nas et al., 2011: 17).

| Implementation failure  | Number of process evaluations reporting the failure |
|---|---|
| No Guidance document how to implement the intervention  | 15/20   |
| Guidance document unclear   | 10/20   |
| Implementers follow their own ideas   | 5/10  |
| Lack of acceptance of the intervention by implementers and the organizations  | 10/20   |
| Human resources problems:   |   |
| • not enough personnel to deliver the intervention  | 10/20   |
| • quality of personnel  | 9/20  |
| • personnel lacking experience to deliver)  | 3/10  |
| Participants: Not enough participants to work with the intervention <sup>1</sup>  | 9/20  |
| Problems dealing with inter-organizational collaboration between organizations active in the Security and Justice chain | 7/20  |

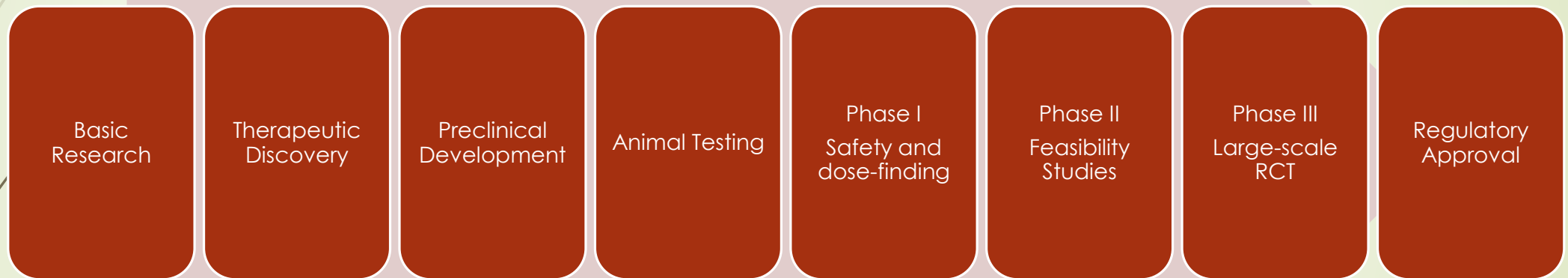
<sup>1</sup>This is known as the Lasagna problem.

THE

Realists'

WARNING

# EBM: Pawson's medicine development 'pipeline'




————— typically 10-14 years —————>



# EBP: Pawsons' Pathway to [social] policy programs RCT





Thanks! Wish you very inspiring and challenging nest two days

➤ [Frans.leeuw@maastrichtuniversity.nl](mailto:Frans.leeuw@maastrichtuniversity.nl)

Or

➤ [flleeuw@cuci.nl](mailto:flleeuw@cuci.nl)

Or

➤ [fransleeuw@gmail.com](mailto:fransleeuw@gmail.com)