Working Paper

Vol 7 no 1 year 2014

Apples and Oranges:Synthesis without a common denominator

Bojan RADEJ

Ljubljana, february 2014, rev. from march 2011
Slovenian Evaluation Society
Tabor 7, Ljubljana, Slovenia
info@sdeval.si, http://www.sdeval.si

CIP - Kataložni zapis o publikaciji
Narodna in univerzitetna knjižnica, Ljubljana

316.4(0.034.2)
303:316.4(0.034.2)

RADEJ, Bojan
Apples and oranges [Elektronski vir] : synthesis without a common denominator / Bojan Radej. - 1. izd. - El. knjiga. - Ljubljana : Slovensko društvo evalvatorjev = Slovenian Evaluation Society, 2014. - (Delovni zvezek SDE = Working paper SES ; vol. 7, no. 1)

Način dostopa (URL): http://www.sdeval.si/Publikacije-za-komisijo-za-vrednotenje/Meso-Matrical-Synthesis-of-the-Incommensurable.htmll

ISBN 978-961-93348-4-3 (html)

272329728

Slovenian Evaluation Society,
Working paper, vol. 7, no. 1 (February 2014)

»Apples and Oranges: Synthesis without a common denominator«

Bojan Radej, Slovenian Evaluation Society,
bojan.radej@siol.net

Abstract: Evaluators of large-scale and multi-domain policy interventions have an aggregation problem (Scriven). Leopold et al (1971) recognised it and left evaluation results in a disaggregated form. But he failed to observe that cross-sectional impacts are weakly commensurable. Ekins and Medhurst (2003, 2006) have appropriately acknowledged this, but failed to implement the principle consistently. When this inconsistency is removed, initial Leopold matrix can be translated into an input-output matrix of impacts on the meso level of evaluation, which leads to correlative synthesis. In the second part of the paper, the aggregation problem is also in horizontal perspective so as to extend Dopfer's vertical classification of meso 1, 2 and 3 levels of synthesis with meso 2a and 2b levels. Conclusion is that precondition for neutral evaluation is not only an objective analysis but also a consistent synthesis of findings. Synthesis is a part of political arithmetic and far from trivial when dealing with social complexity. Meso synthesis is extensive into itself.

Keywords: Social complexity, Aggregation, Impact assessment, Meso level, Secondary meaning.

History of this text: Version 5 (2007-2014). First published in Časopis za kritiko znanosti, vol. 34/227 (spring 2007) in Slovenian language and in IER's Occasional paper No. 7/2007 in English. Present text is extended from the English version 2011 (Radej, Sage, Evaluation vol. 17/2), and revised Slovenian version from 2012 (Radej, Golobič, Macur, UL/BF). All previously expressed disclaimers and acknowledgments also apply here.

Proposal for citation: Radej B. Apples and Oranges: Synthesis without a common denominator. Ljubljana, Slovenian Evaluation Society, Working papers, vol. 7, no. 1 (February 2014), 40 pp.
http://www.sdeval.si/Publikacije-za-komisijo-za-vrednotenje/Meso-Matrical-Synthesis-of-the-Incommensurable.html

Language versions. This paper is also available:
- In Slovenian language http://www.sdeval.si/attachments/512_sde_DZ_3-13_%20Jabolka%20in%20hru%C5%A1ke%20%2822avg2013%29.pdf and
- In Spanish (Version March 2011, in the book: Farinós Dasí J. (ur.). 2011. From Strategic Environmental Assessment to Territorial Impact Assessment: Reflections about evaluation practice. Valencia, Publicaciones de la Universitat de València).

Ljubljana, February 2014

1 Aggregation problem

Scientific conception of social reality as essentially simple is fatally inadequate for comprehensive understanding of increasingly complex social issues. Uniform scientific concepts adopted at the beginning of modern age have been uniformly applied in the research of physical as well as social facts. Dominant version of "normal" science (Funtowicz, Ravetz, 1994) is based on methodological reductionism which necessarily leads to micro or macro chauvinism (Turner, 2006) because it produces holistic understanding with the means of bottom-up or top-down simplifications – generalizing either from detail to the whole, such as with aggregation from micro to macro or from general to specific, such as with disaggregation from macro to micro. Reductionist frame of normal science resulted in monumental progress of our knowledge about things but also shaped relationally impoverished societies operating with homogenised systemic values and dysfunctional majoritarian democracies (Hurley, 1999).

Rejection of simplistic aggregation algorithm goes beyond purely technical observation, that even slightly different methods, such as arithmetical, geometrical, multi-criteria, weighted, one- or two-step aggregation, can and do produce substantively different quantitative results (Gutiérrez et al., 2013). More importantly, methodologists of social research (Giddens, 1976) pointed to asocial consequences of dominant aggregation algorithm because it results in extensive social exclusion of untypical contents (such as Hicks, Samuelson in economics; Weblen, Popper in sociology) if it is not compatible and therefore not commensurable with dominant common denominator (Radej, 2013; Radej, Golobič, 2013), such as with market value or with any uniform statistical conception such as per capita indicator or aggregate index.

In the search for a more comprehensive understanding, social reality is increasingly modelled as complex (Geels, 2002; Easterling, Kok, 2002; Bar-Yam, 2004; Sanderson, 2000). Complexity means that opposing and even contradictory explanations of social issues coexist and they are irresolvable in a simplistic and purely analytical way. There is no unifying common denominator for direct comparisons of diverse appearances of a given social fact due to incompatible value systems of various agents with deeply divided social aspirations, lacking shared understanding of what is primary important to a society as a whole. Complexity of social issues arises from different scales (individual – society) and mutually incompatible domains, e.g. economic, social and environmental, from which a given social fact is evaluated. Complexity is then a prominent reason for the surfacing of aggregation problem as a major methodological concern in social research. It calls for reappraisement of traditional aggregation methodology from the aspect of its ability to provide a more holistic and plural understanding.

Authors who apply evolutionary method in social research like Veblen, Schumpeter, Hayek and neo-Scumpeterians (Dopfer, Potts, Foster) among economists or Turner and Sanderson among sociologists abandoned linear and reversible thinking of chauvinists and replaced it with non-

linear thinking which produces irreversible results. They observe micro and macro as two disconnected presentations of a given objective matter. Despite deep disagreement between classical chauvinists and evolution theorists they both observe the aggregation problem only in the vertical direction. This is too narrow for comprehending complex objects of research that also arise horizontally. However, the horizontal aspect of synthesis is hardly ever given due attention in the research of complex social concerns.

The modern scientific tradition treats present aggregation algorithm as essentially unproblematic and thus only of a technical importance for research methodology. On the contrary, assumption of complexity elevates the aggregation problem into the heart of social theory. Ritzer and Smart (2003) say that every epoch of social history forwards different synthesis approaches and these are always an integral part of political arithmetic. Question of synthesis is about standards of consistent understanding of social reality. Social theory is basically a search for an appropriate theory of synthesis that politically consistently orders and explains the social (Ritzer, Smart, 2003).

Evaluation of governmental policy measures' diverse impacts on overall social welfare is used as an illustrative example for elaborating aggregation problem in the research of complex social issues. To evaluate policy impacts is to collect factual evidence of the performance of a policy and make a neutral judgment about the policy's "worth or merit" (Scriven, 1994). The importance of evidence-based policy-making is one of the postulates of the "new public management". In this doctrine, evaluation fulfils an essentially instrumental function in answering the question how effective are chosen policy means in achieving their specified public ends (Schwandt, 1997). Impact evaluation is in ruling doctrine firmly situated within the positivist scientific paradigm in which policy-making is seen only as a mechanical exercise in social technology (Sanderson, 2000).

There is, however, widespread recognition of the failure of impact evaluation to live up to its promises of improving public management and contributing to overall welfare. Governments have been experiencing systemic failures in managing complex social issues and, in particular, in producing forms of knowledge that strategically inform action to improve policies. The Impact Assessment Board (2009) estimates that the majority of impact-assessment studies provided to the European Commission supply the kind of information that does not inform policy makers whether their global objectives can be met. Huitema arrived to similar conclusion that some 60-80% of 260 evaluation studies prepared for the needs of EU climate policy avoid or at least attempt to diminish complexity of the evaluated objects (Huitema in dr., 2011).

A further challenge to the dominant evaluation doctrine is posed by increasing social complexity. The concepts of chaos and complexity are well established in natural sciences but have only recently come to the fore in evaluation research (Geels, Easterling, Kok, Munda, Virtanen,

Rotmans). Theory of social complexity challenges dominant assumptions that underpin linear thinking, such as about commensurability of social issues. In the evaluation of complex issues, judgment comes from many independent sources, through many technical means, using multiple criteria embedded in different value systems. Social complexity refers to multi-domain and multi-level judgements, which are poorly understood in totality if studied only by conventional causal models, which are unilineal – which means that a single determinable cause leads to a single specific and clearly distinctive effect. In complex frame, evaluators necessarily face the question of how to consistently comprehend and integrally report on varied cross-sectional policy impacts which are multineal – where a specific causal factor, e.g. a policy measure, leads to qualitatively diverse social-wide impacts, such as intended and unintended impacts, and in turn each social wide impact is caused by several independent sources which cannot be evaluated reductively by means of causal disentanglement.

A heterogeneous corpus of information produced in impact evaluation of policy measures is rendered sensible through the process of synthesis (Encyclopaedia of Evaluation, 2004). Synthesis as a logical procedure consistently ignores what is not essential and outlines what is set as indispensable from the aspect of the whole. For complex social issues, there are different and incompatible aspects of what is to be counted as indispensable. When the evaluated impacts are not commensurable, they cannot be expressed numerically in terms of each other (Gutiérrez et al., 2013), and so they cannot be unified on a common denominator (Funtowicz, Ravetz, 1994). From a normative point of view, a complete compensability such as economic growth for environmental pollution – and thus complete commensurability of impacts – is often not even desirable (Munda, 2012). With multiple and incompatible values in mind, evaluator cannot obtain synthesis simply by putting all assessed details like a jigsaw puzzle together.

Aggregation problems exist in evaluation at virtually every level, from the initial issues of data construction and model specification to the subsequent issues of how to usefully summarize and apply fragmented results (Blundell, Stoker, 2005), which is central for us. Summation of commensurable elements, obtained with analysis of various independent aspects of an evaluated object, implies mere amassing of the elementary content into a categorical aggregate conclusion. In this way, the whole is comprehended as a larger quantity of a given elementary quality, so that obtained aggregate category is seen purely in terms of the quantity of its parts and is not as an independent quality of the whole. A new quality or a new wisdom about the whole cannot emerge from a tautologous aggregation of commensurable content (Allport, 1928). To regard social world as mechanical and everywhere measurable in terms of a single operational unit is indeed to rule out the diversity of human beings (Faris, 1939) and complexity of social life. Gould (1996a), for instance, refused the theory of culturally motivated biological determinism claiming that value can be assigned to individuals and groups by measuring intelligence as a single quantity; he refused a thesis that social and economic differences between human groups –

primarily races (non-Westerns to Westerners), classes (lower to higher), and sexes (women to man) – arise from inherited distinctions which can be explained as different levels of evolution in this way sometimes even unconsciously justifying superiority of the Western (north European) white man.

Another relevant illustration is put forward in the theory of intersectionality (Kimberlé, 1994), which studies patterns of gender inequality. This theory examines how various socially and culturally constructed domains of discrimination which are incommensurable, such as gender, race, and class interact in multiple areas, contributing to systematic inequality which is larger then the sum of three inequalities observed separately. Theory dismisses the aggregative claim that black women are twice as poor as white women because their condition is the sum of sexism and racism. The complexity of the social inequality cannot be captured by such simple additive arithmetical frameworks (Prins, 2006). In this context, we recall a paraphrase in Baker's (2012) article about moral hazards of measurement claiming that commensurable synthesis is composed of small truths that tells cardinal lies. Human beings are complex organisms living in complex conditions, so that any plausible measure of social processes will almost certainly also have to be complex (Michalos, 2011). Despite that, the aggregation problem is mainly absent from conventional textbooks (Elsner, 2007). List and Polak (2010) extended this observation claiming even further that the aggregation problem in general remains unresolved in social sciences as the literature has mainly concentrated on how to avoid it.

Society as a complex object demands a complex research methodology, including new aggregation algorithm. The starting point is that there are essentially diverse or socially incommensurable viewpoints which analyse and evaluate social reality in incompatible ways. The concept of social incommensurability (Munda, 2004) implies that no objective basis exists for rational choice between alternatives because different principles of legitimacy and social primacy must be reckoned with and reconciled (Wacquant, 1997). When values are irreducibly plural, value conflicts are un-decidable (Mouffe, 2000), i.e. there can be no rational resolution of the conflict between them, and therefore no clear-cut judgment about merit and worth of a given controversial policy judgment is possible. Kuhn (1970) argues that different sciences say economy and ecology, integrate information in different ways. Different theories weigh the appearances of the same world differently, so they support different judgments. This leads to disagreements about appropriate ways to sum up what we know about social reality. When confronted with multidisciplinary issues, even the most competent, honest and disinterested scientists may arrive at different problem framings and conclusions because of systematic differences in the way they summarize available information (Mumpower et al., 1996).

There are two equally important axis of social incommensurability. In our practical example, the evidence about policy impacts is obtained at different scales of assessment so it provides

understanding at different levels: at micro, meso, or macro level. In evaluation, this induces incommensurability in scale, e.g. when a given issue is perceived differently in its local specifics and in its global entirety, such as between self-interest and collective good. This constitutes the vertical axis of complexity, which demands a multi-level approach to impact evaluation. The second axis of social complexity is involved in irreconcilable evaluation scopes, such as in our example environmental, economic and social. Interest groups put forward equally legitimate and so equally (primary) important, but inconsistent policy demands. This sort of incompatibility constitutes the horizontal axis of complexity which requires cross-sectional and multi-criteria evaluation. Incommensurability in scale and incommensurability in scope (or domain) are two coordinate axes of organized (Easterling, Kok, 2002) or ordered social complexity (Foster, 2004).

Introduction of a concept of social incommensurability into aggregation algorithm is needed because we recognize that what one observes and thinks about a given social fact is always predefined in domain (scope) and in scale of her evaluation. The idea that policy impacts should be evaluated from a multiple points of view has long been recognized (Rotmans, 2002; Weaver, Rotmans, 2006), but poorly implemented. Standard approaches are designed for the appraisal of homogeneous policy interventions with commensurable impacts observable from one specific point of view (Elbers et al., 2007; Rotmans, 2002). However, in the real life, governments usually intervene in conditions that are not homogenous, they have a number of primary goals – all of which may hinder, support or reinforce each other with their unintended or secondary impacts.

One of the immediate consequences of complex perception of society is the need to review the foundations of the aggregation methodology in evaluation (Scriven, 1994). In described conditions, aspiration for evaluation of cumulative wide social impact of a complex policy intervention is far from trivial (Veen, Otter, 2002). The lack of explicit justification of the aggregation procedure is the Achilles heel of evaluation efforts (Scriven, 1994). Different algorithms for aggregating yield different results and, more importantly, lead to contradictory policy conclusions (Gutiérrez et al., 2013).

There is an apparent paradigm crisis in policy evaluation (Virtanen, Uusikylä, 2004; Hertin et al., 2007). Foster and Potts (2007) argue that the crisis is in part a result of unresolved aggregation problem. The majority of standard evaluation approaches are struggling with how to appropriately address incommensurable impacts, such as the EU's strategic impact assessment (2001/42/EC), Impact Assessment Guidelines (SEC(2005)791), territorial impact assessment (TIA; ESPON - 3.2, 2006), and ex-ante assessment of the contribution of the EU structural funds to regional sustainability (GHK et al., 2002).

In evaluation theory there is a harsh disagreement over assumptions about the aggregation of the assessed detailed impacts across multiple domains and from micro to macro level. Standard approach is either divisive, when evaluated macroscopically, or leads to relativization, when approached solely from the aspect of diversity of microscopic evaluations. This double *cul-de-sac* called for methodological response. Older methodological approaches, strictly speaking, do not aggregate detailed assessment results. Leopold et al. (1971) were the first to set this rule. They proposed a detailed assessment method at the micro level from which synthesis of results remains absent. Besides, Leopold's matrix is binary because it is only concerned with two domains – economic and environment, and assesses the possible side or secondary effects of the former on the latter (section 2).

On the contrary, Ekins and Medhurst (2003, 2006) proposed a macro evaluation approach from a multi-domain perspective. They claim that the majority of the assessed policy impacts conform to the normatively prescribed critical system thresholds, so they can be evaluated as 'normal' and therefore conditionally commensurable. They propose a partly aggregated version of the Leopold matrix, expanded from two to four assessment domains (social, human, economic, nature; the obtained matrix is named here the Leopold-Ekins-Medhurst matrix or LEM). They allow for partial aggregation of assessed impacts for all assessment criteria within each of four assessment domains that are placed in the columns of the LEM. Policy impacts are aggregated only within each impact domain, but not between them as they are incommensurable. In this way they obtain at the end four composite indicators of a social-wide impact of evaluated policy, one for each evaluation domain. Their work is an important step towards cumulative assessment methodology in the framework of social incommensurability. An analogous logic to LEM is now accepted in various procedures, such as the EU's strategic impact assessment (2001/42/EC), territorial impact assessment (ESPON – 3.2, 2006), and assessment of the contribution of EU structural funds to regional sustainability (GHK et al., 2002).

Nevertheless, LEM's aggregation approach is inappropriate in its second step, when it allows for the summation of all policy measures' impacts irrespective of their source of impact. The case is that a given policy's impacts on different assessment domains are not qualitatively the same and not commensurable. We argue that fragmented assessment results can be aggregated in LEM only partially by the given source and given impact domain. This reorganizes LEM into the square input-output or Leontief (1970) matrix (in section 3), which is situated at the meso level. Meso level is where vertical and horizontal axis of complexity intersect. In this way mesoscopic framework is obtained for evaluation of complex social issues.

Meso-level reasoning is pertinent for researching phenomena in transformation due to their internal evolution, and evolution is one of the defining characteristics of complexity. The complex social system is actually built upon meso – as Veblen and Schumpeter showed – where

the very substrate of social is generated (Goldspink, 2000). Dopfer et al. (2004) go further and develop a three-level classification of meso 1, 2 and 3 describing three distinctive phases in the synthesis from micro to macro. But the horizontal aspect remains largely absent from the standard evolutionary model. To further elaborate this problem, we extend evaluation case study horizontally from three to four domains (scopes) as in the Four-capital model (Ekins). Overlap between four domains has more layers then the one between three domains, it is deeper, and therefore demands adding meso 2a and 2b sublevels to classification of meso levels (section 4). If even more horizontal scopes were added, meso perspective would further extend into itself. Conclusion is that, contrary to the simplistic aggregation algorithm, complex synthesis provides a new insight when descending into dark depths of its secondary, tertiary etc. meanings – as revealed by overlaps between relativised categorical distinctions through which an object is evaluated – not any more from the aspect of the primary content located somewhere in bright heights of the *suma summarum* category assumed as homogenous and having a uniform meaning in all of its diverse appearances and readings.

The aggregation problem is illustrated by a comparative ex-ante evaluation of the development program for the Slovenian region Pomurje for 2007 – 2013 (RP; Radej, 2006). Pomurje is the least developed Slovenian region (at NUTS 2; with 6.6% of the national territory and 4.3% of national GDP) with a strong cultural and ecological identity – more than a third of its territory is protected nature reserves including unique landscape along the River Mura. Its economic capital is fragile but improving since the mid 1990s. Its social capital is very frail and further depleting. For half a century, the region had been surrounded by cold war borders. Along with geostrategic realignments in Eastern Europe in the early 1990s, it has found itself on the main European transport corridor which exposed it unprepared to international flows of people and businesses. The accession of Slovenia to the EU also imposed a more restrictive border regime between Pomurje and Croatia (at that time an EU candidate country), which previously (in Yugoslavia) were traditionally close. These trends have further weakened social capital, leading to continued depopulation, brain-drain, long-term unemployment, prolonged health and social risks for vulnerable groups (the majority of regional population is officially classified as vulnerable). Regional development lags have accumulated in social capital despite the increased inflow of resources earmarked for less-advanced regions from the national and EU budget in the past two decades, because not enough emphasis was placed on genuine local needs in externally imposed regional projects. This suggests that national and regional policy-makers have failed to address critical regional trends and trade-offs between economic, social and environmental domain appropriately. This supposition has been tested in the assessment of the RP and clearly confirmed, but only with newly developed mesoscopic evaluation approach.

2 Micro and macro approaches

First-generation methodologies for matrical impact assessment of complex policy interventions were proposed by geologist Luna Leopold and colleagues (1971). Their approach encompasses only two domains (economy, environment). It is focused at the micro level, only assessing impact of individual economic measures' impacts on selected environmental criteria. The approach is nevertheless appealing because it goes beyond monitoring policy performance – i.e. how a particular economic policy measure impacts economic assessment criteria – and so in a way introduces cross-sectional evaluation between incommensurable domains. Leopold et al. in this way placed assessment of side effects or secondary impacts (of economic policy measures on non-economic, environmental assessment criteria) in the centre of evaluation concerns. Unfortunately, they decisively refused to capitalise on this achievement.

In the finest positivist tradition of analytical researchers, Leopold et al. aim to assess the complex policy issue through a pedantic description of its numerous impacts on the most detailed criteria. They listed in their matrix the 100 most important economic policy measures horizontally, in the rows of the matrix and 88 criteria of environmental impact vertically, in the columns. This created a matrix with 8,800 cells – each further divided into four sections that describe every impact by its size (large/medium/small), direction (positive/negative/neutral), probability (high/low) and the amount of risk (critical or not). In this way impacts are assessed in sufficient detail to enable maximally informed policy decisions.

Appropriately recognizing the incommensurability between economic and environmental domain of an assessment, they rejected the summation of diverse impacts into an aggregate impact indicator. It is claimed that detailed assessment results should be presented disaggregated, leaving policy-makers with full responsibility for the evaluation synthesis and for drawing its policy implications. Refusal of aggregation is essential for neutral evaluation, says Leopold, as it draws a demarcation line between evaluator and policy-maker to protect the former from the value judgments and political interference (Kunseler, 2007). This argument has been accepted as an evaluation standard for decades and is also respected in the EU's Impact Assessment Guidelines (SEC(2005)791).

Rejection of summation in evaluation and shifting this task to policy-makers is problematic. Refusing to summarize is "letting the client down at exactly the moment they need you most" (Scriven, 1994). Fragmented evaluation results make political decisions more informed but not necessarily easier (Diamond, 2005). It is the failure of politicians as social aggregators that calls for policy evaluation in the first place – recall Arrow's impossibility theorem (1951) which proves that rationality and democracy of public choice cannot be simultaneously satisfied. A problem is especially evident in the assessment of complex 'cross-cutting' or horizontal social

concerns (Sanderson, 2000) such as sustainable development, gender equality and social cohesion. Sanderson notes that evaluation that seeks to isolate policy instruments will produce results of limited usefulness due to limited external validity. Assessment which simply produces non-overlapping information tends to underplay inherent system contradictions, legitimizing disregard of legitimate stakeholders' concerns in policy-making (Stake, 2001). Without any explanation of how different parts of a public program work together horizontally, assessment fails to satisfy information needs at the strategic level and produces banal answers to multi-dimensional questions (Virtanen, Uusikylä, 2004). Finally, when assessment results are left horizontally unrelated, it is impossible to substantiate evaluation findings which leaves them exposed to political manipulation.

The selected case study illustrates the issue. Table 1 presents detailed assessment results for RP. They were obtained in a group of experts who convened a workshop and applied the Delphi method to assess the possible impacts of the 47 proposed RP measures on a selected set of assessment criteria (social, economic, environmental). In order to simplify the description of methods, experts' opinions are presented here only in terms of the direction of impacts: be they positive, neutral or negative.

Table 1: Micro view of RP's impacts – Leopold's matrix

Policy impacts by domains of evaluation / Program Measures	Economic		Social		Environment	
	GDP growth	Investm. intensity	Unemp-loyment	Migration	Abate-ment ex-penditure	Sewerage connectio n
1 Development lag	+	+	0	-	+	+
2 Competitiveness	+	+	-	-	-	+
3 Investment promotion	+	+	+	+	0	+
4 Endogenous advantages	+	+	+	+	+	+
5 Entrepreneurship	+	+	-	+	0	0
6 Regional tourist organizational model	0	0	0	0	0	0
7 Pomurje as a tourist destination	+	0	0	+	0	0
8 Destination management	0	0	0	0	+	+
9 Destination marketing	+	0	0	+	0	0
10 Human resources in tourism	0	+	+	+	0	0
11 Quality management	0	+	0	0	+	+
12 Tourist infrastructure investment	+	+	0	0	+	+
13 R&D in tourism	+	+	0	0	0	0
14 Health inequality (criteria)	0	0	0	0	0	0
15 Health promotion network	+	+	+	+	0	0
16 Health inequality – regional	+	0	+	+	0	0
17 Health inequality – vulnerable groups	+	0	+	+	0	0
18 Quality, access to health services	+	0	0	0	0	0
19 Healthy environment	0	+	0	0	+	+
20 Mental health	0	0	+	+	0	0
21 Agriculture modernisation	+	+	-	-	+	+
22 Environmental agriculture	+	+	+	+	+	+
23 Entrepreneurship in agriculture	+	0	+	0	0	0
24 Human development in agriculture	0	0	+	+	0	0
25 Value added growth	+	+	-	-	+	+
26 Products, services – farms	+	+	+	+	+	+
27 Products, services - agro industry	+	+	-	0	+	+
28 Marketing agro-products	+	+	0	0	0	0
29 Rural development –products & services	+	+	+	+	0	0
30 Countryside development	+	+	+	+	+	+
31 Rural entrepreneurship	+	+	+	0	0	0
32 Rural stakeholders' co-operation	+	0	+	+	0	0
33 Water supply	+	+	+	0	+	+
34 Transport infrastructure	+	+	+	+	+	0
35 Alternative, local energy sources	+	+	+	+	+	0
36 Energy distribution network	+	+	0	0	0	0
37 Access to IT services	+	+	+	+	0	0
38 Waste waters, collection & treatment	+	+	+	0	+	+
39 Solid waste management	+	+	+	0	+	+
40 Communally equipped zones	+	+	+	+	+	+
41 Water quality	+	+	+	0	+	+
42 Revitalisation of hot-spots	-	-	0	0	+	+
43 Illegal land-filling, monitoring	+	+	+	0	+	+
44 Nature and culture conservation	-	+	-	0	+	0
45 Energy policy	+	-	+	0	+	0
46 Spatial planning	+	+	+	+	0	0
47 Communication strategies	+	+	0	0	0	0

Source: Radej, 2006.

The assessment results presented in Leopold's fragmented way in Table 1 would be usually summarised following three lines of descriptive reasoning: (i) a prevalence of the program's

positive impacts suggests that a majority of the measures will positively contribute to regional development, which supports its endorsement; (ii) negative impacts focus evaluator's attention on the weakest parts of the proposal, which ought to be improved, compensated or abandoned in the finally adopted program; (iii) neutral impacts (0) are not really problematic and can be skipped in further evaluation.

These three lines of reasoning would suggest that the evaluator and policy-maker should focus their attention on negative impacts. However, this evaluation approach is not appropriate in the context of social complexity where a given positive (and sometimes neutral) impact may not be unproblematic in all relevant contexts; nor may a negative impact necessarily be assessed as a threat, when it is fully compensated to the consenting victim in his preferred "currency" – in monetary, physical or in symbolic form.

The prevalence of positive impacts does not support the conclusion that the proposed policy is adequate in general, but only that the policy proposal has been prepared by a largely competent authority. Proposals prepared by a democratically elected government are carefully scrutinised as well as painfully negotiated among various group interests before they are submitted for formal evaluation.

There are some additional reasons why a dominance of positive impacts shall not by itself lead evaluators to positively assess the policy proposal as a whole. Impacts are sometimes assessed against criteria selected by formally responsible implementation authority itself, which questions their neutrality. Even when this is not the case, impacts are assessed each in relation to its own benchmark, i.e. in isolation from each other. Successful realization of separate policy goals does not by itself guarantee positive society-wide impact of the policy as a whole if its goals (or assessment criteria) are in conflict – a very common situation in policy making.

A prevalence of positive impacts in a Leopold matrix can, at its best, inform policy-makers about their effectiveness observed at atomistic level (micro view) while it does not enable a systematic conclusion about the proposal's merit and worth for the overall society. Only when systematic evidence of positive impacts is obtained, can evaluators assess the appropriateness of the overall proposal. But "systematic evidence" can be identified only at higher levels of evaluation, when detailed results of assessment are properly summarised. So evaluation judgment about overall policy impact does not directly follow from detailed assessment results but is the outcome of their "post-production" with the careful synthesis – either accomplished with a voluntaristic political judgment (as Leopold would have it) or with some logically justified aggregation procedure.

The aggregation imperative sparks two methodological concerns in evaluations, both horizontal in nature and concerned with the conditions under which negative impacts may be compensated with positive impacts. The first concern relates to situations in which different experts cannot

reach consensus on the direction of a given impact (positive or negative). Some approaches, such as CAF (Common Assessment Framework; 2006) suggest that assessed differences need to be discussed with the aim of reaching consensus among participating assessors. However, forcing consensus for every single assessment detail is risky because it may invoke existing asymmetries within the assessment team – such as their different negotiating skills, which can lead to a kind of closed, exclusive process (Connelly, Richardson, 2004) where the dominant actor prevails. Sankey (1995) is instead more in favour of "rational disagreement" where experts verify arguments for their disagreements and discuss them – not necessarily to reach consensus but at least to confirm validity of arguments that stand behind the opposed claims. If results of detailed assessments are conflicting but well founded, disagreement between experts is based on valid arguments and so irresolvable. This however does not mean they are not aggregatable. Alternative to consensus as a precondition for summation is cancelling out the opposing assessments. This approach is also applied in our case study. Expert opinions are not only equally legitimate but also only specific, partial claims – at least when observed horizontally from the wider perspective of the overall program evaluation. Cancelling-out of incompatible assessments is a threat to disagreeing experts. These encourage a symmetrical cooperative effort towards convergence of their assessment differences.

The second difficulty is related to aggregating positive and negative impacts of a given policy measure on various assessment criteria (or similarly, positive and negative impacts of different policy measures assessed against the same evaluation criteria). To resolve this one has to take a position on the fundamental issue of compensability. Compensability refers to the existence of trade-offs, i.e., the possibility of offsetting a negative impact of a given policy measure on one indicator by a positive impact on another indicator (Munda, 2012). For instance, in Table 1, may negative impacts of entrepreneurship promotion on employment in Pomurje be outweighed by positive impacts of entrepreneurship promotion on migration? Or another example which is extensively elaborated elsewhere (Radej, 1995): is it permitted in evaluation to cancel out additional tons of greenhouse pollution (negative impact) with additional purchase of tradable pollution permits (positive impact, because its proceedings finance additional environmental investment at the permit seller's plant)? Greenhouse emissions cause irreversible changes in the climatic conditions and so the economic and climate domain of evaluation are not interchangeable but incommensurable. Thus a trade-off between greenhouse gases and money is not adequate as a general principle. However, climate change is not exclusively macroscopic but complex phenomena that must be observed at multiple levels with different evaluation principles. Trade-offs between income and greenhouse emissions are not incommensurable in every single case, locally, or at least people are not willing to treat them as such.

To incorporate this peculiarity of evaluation, system thresholds have emerged – such as ecological and social safety standards (for a survey of literature see Muradian, 2001). Threshold

marks a tipping point, beyond which a small quantitative change can have a disproportionately large (qualitative) effect on the entire system critically involving high risk to its integrity.

The concept of system thresholds is closely linked to incommensurability of social phenomena. Existence of quantitative thresholds means that there are discontinuities in the measurement of value-based social phenomena and value addition (Mason, 2006) that are consistent with discontinuity between individual and social values. As Wiggins (1997) explains, two values are incommensurable if there is no general way in which A and B trade-off across the range of situations of choice and comparison in which they are normally present. Social phenomena are usually not incommensurable as such but only beyond (or below or both) their threshold values and not necessarily within these limits, where in fact a large majority of social interactions takes place. Within the safety limits an agent either does not sense the qualitative difference between two distinct social conditions or refuses to declare a preference for one or the other (Luce, 1956 in Munda, 2006), as in the case of minor environmental damage that stays within safe ecological standards. System thresholds hence normatively define objective criteria against which policy impacts are conditionally commensurable. Systemic thresholds encircle a Pareto indifferent space of social normality within which concerned parties in a micro context (locally) freely interact and trade-off positive with negative impacts, depending exclusively on specific considerations of those directly affected by trade-offs and their case specific compensability schemes for offsetting voluntary victims. Beyond these thresholds, any further trade-offs are prohibited, even if victims consent can be preserved, because their consequences are perceived as incommensurable on the higher, social level of considerations.

Recognising system thresholds is important for policy impact evaluation as it extends possibilities for synthesis of at least partly commensurable detailed assessments. One of the first macro-evaluation methods recognizing this was strategic environmental assessment (Sadler, Verheem, 1996; SEA Directive, 2001/42/EC), but it gives no indication about how to cumulate environmental impacts and parallel them in an aggregate way to economic ones. The missing link has been contributed by Ekins and Medhurst (2003, 2006) in their novel approach to assessment of the EU structural funds' impacts on regional sustainable development. Ekins previously proposed the Four capitals model which evaluates in parallel economic, social, environmental and human impacts of a given policy proposal. This particular framework of multiple welfare and evaluation domains can be traced back to the Brundtland report (WCED, 1987), and to the conference on sustainable development in Rio de Janeiro (UNCED, 1992, where Munasinghe and Ekins independently proposed analogous idea).

Ekins and Medhurst have developed LEM as a highly compacted form of the Leopold matrix. The columns are reduced from 88 fields of possible environmental impacts to four domains of sustainability, each covered by a smaller number of assessment criteria (only two in the

illustrative example, presented in Table 1). For simplicity, the original Four capitals model of Ekins and Medhurst is reduced in our explication to a three capitals because this is entirely sufficient to discuss the aggregation problem in evaluation of complex social issues. RP's impacts on individual evaluation criteria are already summed up horizontally within each domain. The number of rows in LEM matrix may be as large as in the case of Leopold matrix, but is also reduced in the experimental case, as suggested in Ekins and Medhurst, from numerous program measures (47 in Table 1) to six main regional policies (see LEM in Table 2). LEM presents impacts on a wider range of scores compared to Table 1: from the most robust positive impact with the highest score (+++) to the most negative impact with the lowest score (---), with all other five intermediate possibilities included. When uncertain about how to round-off aggregate impact from Table 1 to Table 2, a decision was taken based on the comparison of the financial weight of the related measures. So in the final row of Table 2 three vertically aggregated impacts of the RP are presented – one composite impact indicator for each of three evaluation domains.

Table 2: LEM's macro view – the RP's impacts on three scopes of sustainability

Evaluation scopes / Sectoral policies	Economic (E)	Social (S)	Environment (N)
Value added growth (rows 1-5)*	+++	-	+
Tourism (rows 6-13)	+	0	0
Health (rows 14-20)	0	+	0
Rural development (rows 21-32)	+++	+	++
Infrastructure (rows 33-40)	+++	++	++
Environment (rows 41-47)	+	+	+
Summary impact of RP (rows 1-47)	++	+	+

Source: Radej, 2006. *Note*: * Summary of rows 1-5 in Table 1, etc.

Table 2 presents the aggregated impacts of each sectoral policy involved in the RP on all three domains of assessment. Infrastructure development will be the most welfare-enhancing policy, followed by rural development policy. The most problematic is the negative impact of value-added growth, based on enhancing cost efficiency, on social segment of welfare. The impacts of health policy and tourism are disappointing – the explanation is that their measures mostly relate to the preparation of plans and regional organization structures. But the overall impacts of the program, presented in aggregate at the bottom row of Table 2, do not appear problematic. The program will improve regional sustainability in all three domains, though more in the economic domain (++) than in social and environmental ones (+). The summary impacts of the RP are therefore unevenly positive but differences are small. Thus the assessment concludes that the program will have a rather weak but overall positive impact on regional sustainability. Such descriptive conclusions usually bring evaluation to its end.

However, sustainability is a complex concept and so vertical as well as horizontal. Main task of sustainability evaluation is not to examine only how sectoral policies achieve their narrowly defined goals, in vertical direction. LEM cannot say much about RP's impact on the sustainability of regional policy-making in terms of its internal conflicts and synergies between policy domains, horizontally. More subtle insight is needed to overcome the aggregation problem arising from social complexity.

3 Meso approach

If detailed assessment results are not aggregated, as in Leopold's approach, the evaluation produces findings that are too fragmented for holistic comprehension of complex issues. In contrast, full aggregation, such as in macroscopic LEM, results in findings that are too amassed to enable categorical judgment about the evaluated issue. Different aggregation approach is needed which negotiates a compromise between microscopic aversion and macroscopic passion for synthesis.

Too narrow understanding of otherwise appropriately observed incommensurability between evaluation domains is the origin of troubles into which LEM plunges Ekins and Medhurst. They did not acknowledge in LEM that the assessed impacts are vertically not fully aggregatable despite their stringent conformity to their associated system thresholds. Column aggregation assumes the homogeneity of the impacts of all different policy measures on a given assessment criteria. Many studies indeed demonstrate that a given policy does not influence different areas of impact in the same way (Schnellenbach, 2005). Policy impacts are by their nature either (i) direct or wanted when operating vertically and affecting the primarily targeted impact area, or (ii) indirect or side effects which operate horizontally and bring up secondary meanings in evaluation. They are indirect when impact on areas which are not targeted with an evaluated policy measure and usually even fall under the jurisdiction of other policy domain (Rotmans, 2006) with possibly divergent goals and evaluation criteria. It is a reductionist and non-neutral position to assume that only the primary description of the phenomenon describes the real (Cat, 2010), because humans create the social world as their 'second nature' (Benton, 1998), where secondary meanings are equally important for holistic evaluation. If evaluated impacts are not horizontally comparable, they do not share common denominators, so they are not vertically fully aggregatable.

Non-neutral impacts are confirmed even for those sectoral policies that had previously been taken as most neutral such as monetary (Lucas, 1972) and tax policy (Leith, Thadden, 2006). For instance, wide social impact of increased interest rates is differently relevant for the owners of capital than for owners of non-economic wealth, and differently important for financial sector than for non-financial sectors. In principle, policy interventions are sectoral and should always be

analysed not in relation to specific goal they set themselves but to the general interest they are supposed to serve (Donzelot, in Burchell et al., 1991). For this reason, holistic evaluation needs to elaborate direct as well as indirect policy impacts indiscriminately. Yet indirect or secondary impacts routinely fade out in policy evaluation because it is assumed that "they are too complex" (Morçöl, 2011) and impossible to track. This is a profoundly upsetting observation for anyone aspiring to understand and obtain capacity to effectively interfere with complex social processes.

The new generation of evaluation approaches (Guba, Lincoln, 1989) or new wave of evaluation studies (Vedung, 2010) increasingly focus on secondary aspects, such as with the emphasis on "horizontal themes" in policy impact evaluation: gender equity, employment, sustainability. However, these are in practice evaluated vertically. For instance, gender equality is not taken as an umbrella criterion to which policy measures need to contribute but only as one additional evaluation criteria, which is added to existing ones – in this way introduction of horizontal criteria leads to fragmentation rather than towards integration. The same sort of simplifying (mis)understanding has been previously blamed on Leopold. So, all three distinguished musketeers of our story, Leopold, Ekins and Medhurst have gotten into trouble in the same way. Iron shirt of normal science led them to comprehend and methodologically resolve quandary of complexity with the enhanced simplification of their evaluation object instead of with a modified logic – where direct and indirect policy impacts are consistently oriented in vertical and horizontal direction of the complex evaluation frame.

Sectoral specialisation, and therefore the non-holistic nature of policy-making, implies not only that a distinction between incommensurable domains as primary important must be preserved (Ostmann, 2006) and evaluated independently, but also that secondary issues should be taken into account as equally important to primary ones. Distinction between primary and secondary impacts allows only partial vertical aggregation in LEM. For example, if we differentiate in evaluation three domains, economic, social and environmental, then economic and social policy's impacts on the environment are not commensurable so they are aggregated separately (economic impacts on environment separately from social impacts on environment).

Partial aggregation rule seems at odds with the strong version of the incommensurability thesis. Martinez-Alier et al. (1998) point out that in situations when there is an irreducible value conflict in public affairs, we can only search for weak comparability as a facilitator of collective discourse. Some authors have explicitly argued against the strong incommensurability thesis (Morgan, 2007; Nola, Sankey, 2000). They proposed making a distinction between relations of strong and weak in/commensurability. Impact is said to be weakly commensurable when specific limitations are imposed in aggregation such as with the partial aggregation rule above. Further, impacts that are weakly commensurable in two or more incommensurable domains of the

evaluation, like "hybrid" socio-economic impacts (compared to social impacts and economic impacts), are "only" weakly incommensurable.

The difference between strong and weak incommensurability can be illustrated with a juicy example. It is entirely possible to mix apples and oranges in fruit juice because their meat is weakly incommensurable, they can be tastefully combined under certain conditions, provided by a recipe; despite such enjoyable harmony it will never be possible to grow an apple tree from the seed of orange, because seeds in their primary essence are strongly incommensurable. So only in the strong case one is not permitted to add apples and oranges.

Weak incommensurability is essential for possibility of complex synthesis in the evaluation. Weakly incommensurable impacts are secondary to both of its constituting domains. Such impacts with hybrid content can be evaluated against two otherwise incompatible sets of criteria so they imply a delicate possibility of translation between them by the means of non-commensurable algorithm of synthesis.

Let's now capitalize in our case study on distinction between strong and weak in/commensurability. Taking into account weak commensurability of majority of impacts (all cross-sectional, located non-diagonally, see below), evaluator needs to regroup all rows, describing policy measures in Table 1 in the same way in which impact areas are grouped in columns – by their three incommensurable domains (see Table 3). This divides the Leopold matrix into three sections vertically and horizontally, resulting in nine sub sections. When detailed impacts are aggregated, partial aggregate is obtained by a given source (row) and area of impact (column) for each sub-section. These sub-aggregates can be neatly organised in a condensed square "input – output" matrix. Leontief (rus. *Леóнтьев*) initially developed it (its "central quadrant") to facilitate inter-sectoral studies in multi-input, multi-output systems. Matrix is suitable for an explicit presentation of the synergies and tensions between sectors, direct and indirect. For instance, if we have agriculture, industry and services as three sectors in national economy, the rows and columns in a matrix illustrate how the sectors are linked through sale and purchases of each others inputs.

The square matrix exists above the micro-level (Leopold matrix) because it is aggregated from it. At the same time, as a set of only partial aggregates, it exists below the macro level (summary row in LEM). The Leontief's matrix presents an intermediate or meso level of the RP's impacts. This approach to evaluation is mesoscopic. The concept of a meso-matrix is important because the complex social system is built upon meso (Dopfer et al., 2004).

Being rooted simultaneously in micro and in macro, meso has two opposite horizons so it exhibits bi-modality (Dopfer, 2006), which enables a hybrid perspective of the evaluated object. As situated on the middle it shares the logic that is characteristic of both poles, and so enables mid-level articulation of oppositions that accompany public choice. Meso introduces what

Malthus (in Cremaschi, Dascal, 1996) described as the "doctrine of the middle" (Dopfer et al., 2004), which provides one with mesoscopic descriptions, »*une déscription médiane*« (Prigogine, 1996). As a "plural-relativistic" view (Geertz, 2000), meso covers many parallel views of one closed reality containing many (pre) existing substantial contexts. With its bimodal characteristics meso is situated in "the un-excluded middle" (Wallerstein, 2004) of the social complexity. This brought Easterling and Kok (2002) to justify an argument that meso is the perspective from where the modelling of social complexity is the most tractable "a priori". This forms a solid ground for claiming that meso perspective is the imperative for neutral and holistic evaluation of complex social issues.

Bimodality of mesoscopic perspective is crucial as it facilitates better understanding of the basis of deep social contradictions (Mertens, 1999). Mesoscopic evaluator would be able, for instance, to intervene in conflicts to help actors understand where their disagreements have epistemological and ethical roots and help expose the incommensurable meaning systems by which these facts are being interpreted (Bovens et al., 2008). This alone can not overcome principal oppositions among protagonists, but it nevertheless helps at mutual legitimisation of their oppositions.

A mesoscopic evaluation refocuses attention from the performance assessment of sectoral policies' effectiveness to the evaluation of trade-offs among evaluation domains. So it is appropriate to present impacts of policy intervention in the perspective of overlap between its policy domains (inputs, in rows) and its evaluation domains (outputs, in columns). Overlap between input and output domains is denoted with the intersection sign '∩' from the set theory. For example, economic policy impact on the social domain is denoted as E∩S (E overlaps S). A mesoscopic perspective of RP's impacts on three domains of regional sustainability is presented in Table 3.

Table 3: The mesoscopic perspective of RP's impacts

Output: Impacts / Input: Policy measures	E	S	N
E	E∩E = +++	E∩S = –	E∩N = +
S	S∩E = +++	S∩S = ++	S∩N = ++
N	N∩E = +	N∩S = +	N∩N = +

Source: Radej, 2006.

Evaluation results in Table 3 are already too highly aggregated to be really relevant for executive manager, who is narrowly concerned only with the effectiveness of a particular RP measure in her jurisdiction. Table 3 is relevant only for the middle and higher ranking decision-makers who are concerned with overall policy consistency and synergy between impacts of different measures.

Contrary to LEM, Table 3 conveys a qualitatively different conclusion of RP's sustainability compared to Tables 1 and 2. LEM and Table 1 are both verticalist and thus simplistically concerned only with the effectiveness of RP in realisation of its primary, nonoverlapping goals. RP's impacts are multiple and incompatible, which calls for horizontal evaluation. To overcome horizontal incompatibility in synthesis evaluator specifically accounts for secondary or overlapping impacts. This novelty is presented only in Table 3, which consists of two classes of relationship. Primary or intended impacts of RP (E∩E, S∩S, N∩N) are located on the diagonal, top left to bottom right and present the vertical, "key dimensions" or domains of RP. The diagonal elements of Table 3 are strongly incommensurable so they can not be aggregated any further and are interpreted as they are, descriptively. Diagonal elements indicate that: regional economic policy would be very successful in achieving its own primary goals (maximum, three pluses); moderately successful in social policy (two pluses) and only weakly effective in environmental protection (one plus). Therefore social, environmental and economic domains of sustainability are not treated indiscriminately in RP. This observation does not match with the previously obtained conclusion from summary row of LEM that suggested a broadly balanced impact of the program on the three assessment domains. Beside, RP's cumulative impact on economic and social domains is assessed higher in the mesoscopic matrix than in LEM.

Secondary impacts or RP's side effects (E∩N etc) are located below or above diagonal and present horizontal RP's evaluation of the "cross-cutting issues".

Insight into RP's synergies is obtained with the correlation of weakly incommensurable partial aggregates. It produces three correlates in Table 4: SE, for the 'socio-economic' overlap between S∩E and E∩S; SN, as the socio-environmental correlate; NE, as correlate between nature and society. A correlation is applied for evaluation of horizontal connectedness between pairs of variables. In our case correlation assesses strength of two directional relationships between say E∩N and its diagonally symmetric opposite, N∩E (in Table 3). Such correlation is dual as it assesses reciprocal connectedness or overlaps between two evaluation domains.

As far as RP's overlapping impacts are concerned, the correlation in SE is strong but damaging for S which indicates socially constraining RP's economic impacts (Table 4). Low correlation in SN emerges from a socially constraining environment protection measures. Here we recall the previous observation that environment policy will not be very effective in pursuing its primary goals – so the RP will, at least in relative terms, further impose a social burden for only slight anticipated improvement in environmental sustainability. A socially constraining economic and environment protection policy further weakens already fragile regional social capital. Beside, the economic domain very poorly integrates with S and with N; similarly can be said about the feeble secondary impacts of N on E and S. Table 4 thus reveals non-social character of RP, which is highly problematic taking into account the baseline conditions. This observation is

entirely absent from the assessment results presented in Table 1 and 2. The conclusion from meso evaluation is that the RP is inconsistently contributing to regional sustainability – this is again less favourable result compared to the conclusion obtained from Table 2.

Table 4: The RP's correlation matrix of synergy between scopes

Policy \ Impact	E	S	N
E	Economic sustainability (E∩E) = (+++) → Strongly coherent impact	Socio-economic sustainability ES = (+++ , -) → Strongly but negatively correlated, in favour of E	NE = (+, +) → Weakly correlated, balanced
S	-	Social Sustainability (S∩S) = (++) → Medium coherent	Eco-Social sustainability SN = (++, +) → Medium correlated, unbalanced in favour of N
N	-	-	Natural sustainability (N∩N) = (+) → Weakly coherent

Source of data: Table 3.

Disagreement in evaluation conclusions derived from Tables 1, 2 and 4 are, of course, not due to different detailed expert assessments at the micro level; the assessment remains the same in all three instances (Table 1). It emerges solely from the summative endeavour – the simplistic methodology is premised on commensurability of impact while complex one is established on the incommensurability of evaluation domains with only weak in/commensurability of impacts on lower levels of evaluation. Different aggregation algorithms lead to different evaluative conclusions and inappropriate aggregation algorithm misleads policy advice. Policy failure to implement more integrative policies is therefore not necessarily a result of intentional such as neoliberal bias, but can be a straightforward consequence of inadequate strategic evaluation of options in the presence of social complexity. This validates the premise given in the introduction that many difficulties arising from inconsistent impact evaluation are caused by methodologically inadequate handling of complexity of social issues. It follows that, in the evaluation of complex social issues, an appropriate two-part synthesis, taking into account overlapping and nonoverlapping information as equally important, is vital to neutral conclusion (see Lipsey, 2009). Even more, a new theory of synthesis is the starting point from which the more effective methodology of evaluation of social issues can be developed.

4 Deepening into the meso

Now that we have achieved an intermediate goal, explaining mesoscopic aggregation procedure for three-part version of Ekins' model, we can return to its original four-part formulation to meet the second aspiration regarding elaboration of the horizontal logic of complex synthesis. Picture 1 illustrates both models in the form of three- and four-part Venn diagram. Four-part model can no longer essentially change previous evaluation findings from Table 4, as they are both obtained in the same correlative way. Still, it can reinterpret them more in depth to demonstrate capacity

of meso to deepen into itself, but also to display practical limits to mesoscopic exploration of complex social phenomena.

We start with Dopfer, Potts and Foster (2004), who developed theoretical basis for understanding the meso logic in economy. Their explanation may be considered a paradigmatic formulation of evolutionary economics as a meso-centred process (Elsner, 2010). Dopfer et al. do not understand meso as a single level but divide it into three sub-levels which describe three core processes in a trajectory of a given system between two consecutive macro states, such as before and after implementation of the evaluated RP. This is a three-phase process of origination (emergence) of a novelty in an economic system, its diffusion (adoption and adaptation) and retention (maintenance and replication). Trajectory involves a process of creative destruction, disturbing an initial order with a new idea and a new population of followers (or "carriers" in Dopfer et al.'s terms) which is then subjected to forces of variation and selection, adoption and adaptation before stabilizing its structure (Dopfer et al., 2004). They refer to these three distinct phases of a trajectory as meso 1, 2 and 3.

Meso 1 describes micro to meso emergence of novelty by the means of grouping of similarities or affinities. In our case study meso 1 relates to partial aggregation of detailed weakly commensurable assessment of impacts, similar by source and area of impact into input-output table. Meso 2 consists of meso to macro (Dopfer et al., 2004) correlative process in which novelty is integrated (or disseminated) into broader context through mixing and blending, which brings about hybrid categories of evaluation – in our case represented by the non-diagonal fields in correlation matrix. Novelty is established as a new structure which constitutes a new aspect of system normality and provides the basis for a new macro order. With meso 3 Dopfer et al. describe meta-correlation between emergent products of meso 2. Meso 3 is the world of stable knowledge concepts, such as skill, routine, competence, capability. In our case meso 3 is covered with interpretation of correlation results. These feed back as modified preconditions for initiating meso 1 process. This is either supportive and strengthen meso 3 achievements or contest them and subsequently induce new emergent processes – so the whole cycle is autocatalytic (Dopfer et al., 2004).

We propose an extension to Dopfer et al.'s classification of meso sub-levels. Dopfer et al. were not sufficiently equipped to accomplish this step. As the mainstream economists, they produced a theory of meso which is conceptualised in the vertical direction of complexity. They stick to the conventional economic (and biologist's) evolutionary view, that coordination (emergence, synthesis) is mainly vertically oriented problem – how to link and translate elementary (micro) inputs into systemic (macro) results (Dopfer, 2011; Arrow, 1951).

It is a common error assuming that complexity of evolutionary process is somehow tied only to vertical processes (Adams, 2009). (Neo)Darwinists understand evolution as a vertical

progression by natural selection. Analogously, evolutionary economists understand competition where winners emerge as superior and the losers are relegated to subordinate status (Krossa, 2006). For the mainstream authors horizontal evolution is not an independent form of evolution, rather they regards it as a part of variation (Gontier, 2006). Coordination as a horizontal aspect of dynamics is conceptualised in (neo)classical as well as in evolutionary economics with *deus ex machina* of invisible hand of market which takes place behind our backs, hidden to our eyes, inaccessible to our minds – leaving everybody without a possibility of substantive interference into the extension of spontaneous social order (Hayek, 1992).

Economists think in a unilineal way and in a uniform hierarchy: they ascribe primary importance to economic issues, while non-economic considerations are seen as secondary, variation, or "external", and so beyond protagonist's immediate concerns. In their verticalism, evolutionary metaphors are not more successful in explaining social processes than physical metaphors of classical economist. In both cases verticalism turns out as a reasonable strategy of domination, command and control (Olsen, 2006) which creates "islands of excellence in seas of underprovision" (Ooms et al., 2008). Dopfer et al. have successfully broadened horizon of (neo)classical economists from micro-macro to micro-meso-macro perspective. However, they have not questioned underlying verticalist hegemony in the mainstream economics, which curtailed their capacities to fully capitalize on mesoscopic innovation.

Contrary to the ruling doctrines, unorthodox authors, who are relying on complexity theory do not see evolution only as a vertical, but also as a massively horizontal process (Riofrio, 2013), such as in the case of horizontal gene transfer (Woese, 2004), hybridization, parasitism, or symbiogenesis (Gontier, 2006). For Heylighen (1989) complex evolution is in general parallel and distributed. Steward (1955) has developed a theory of "multilineal evolution" or co-evolution. He believed that secondary factors, like political systems, ideologies and religion push the evolution of a given society in several directions at the same time, so it can not be understood appropriately with the unilineal assumption.

Even more is horizontality important in social context. For example, Carly Fiorina, former CEO in a leading computer company, said that "value in this era of technology is delivered horizontally, not in vertical silos, by department, by application, by process", which demands a move away from vertical specialization to horizontal integration (2004). Centrality of horizontal perspective in economy is forwarded with alternative economic models, such as horizontal production chains cantered on the quality, origin and identity of each product, advancement of common goods, and in open economy with sharing. The same horizontal effects are achieved in hybrid models such as socio-economic or socio-ecological models. Global models of multi-polar world can be understood only in orthogonal intersection between the horizontal and vertical aspect.

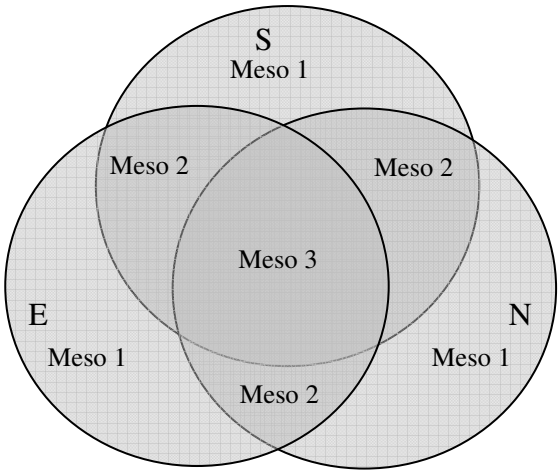
Complex social process simultaneously evolves in vertical and in horizontal direction and they are of different nature. The former defines multiple scopes or domain of evaluation (E, S, N, H), while the later stands for its multiple scales or levels (micro-meso-macro). According to this, evaluated object consists of two axes, and a change in one is related to the change in the other. For example, as given by Bar-Yam (2004) when describing the scale (vertical complexity) and the scope (horizontal complexity) of emergent processes: "Consider observing a system through a camera that has a zoom lens. For a fixed aperture camera, the use of a zoom couples scope and resolution in the image it provides. As we zoom in on the image we see a smaller part of the world at a progressively greater resolution… We must allow a decoupling of scope and resolution, so that the system as a whole can be considered at differing resolutions as well as part by part. For this purpose scale can be considered as related to the focus of a camera—a blurry image is a larger scale image—whereas scope is related to the aperture size and choice of direction of observation."

Now that the horizontal aspect of synthesis is justified as equally important in complex evaluation as the vertical one, we aim to deepen insights into how newly added horizontal matter is changing the procedure of meso synthesis. Dopfer et al.'s initial triadic meso scheme is reconceptualised as complex and presented in Picture 1a, while Picture 1b serves to schematise original Ekins' model consisting of four overlapping horizontal evaluation domains. Comparison of their overlapping areas immediately displays emergence of two new sub-layers of meso structure, meso 2a and meso 2b which indicates capacity of mesoscopic reasoning to extend into its own middle (Prigogine, Stenger, 1982).

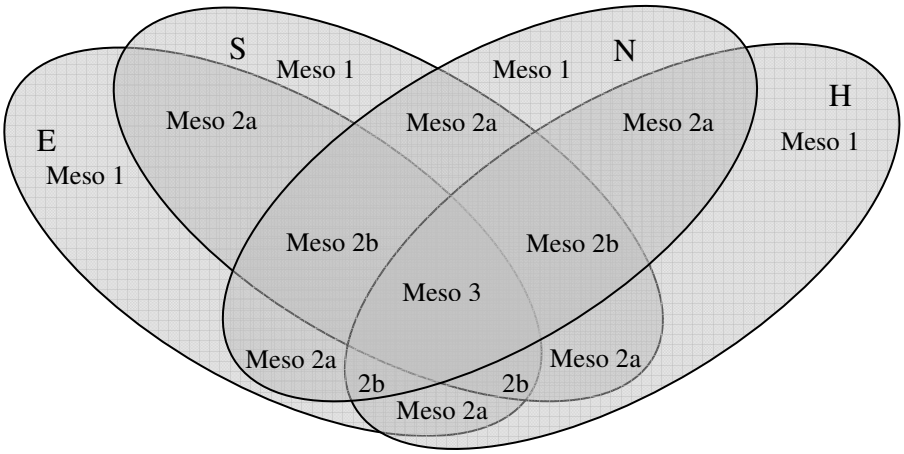
To conclude conceptualisation of the four capital model we return to the case study. Table 5 shows the results of the impact evaluation that are already aggregated at the level of the four capitals (detailed assessment of the effects of RP on H, human capital, are shown only in the original study). Table 5 is split on two; upper part involves a square matrix of impacts (5a, which is analogous to Table 3, in the case of three part model) while lower part presents the correlation matrix (5b, analogous to Table 4).

Picture 1 Three- and four-set Venn diagram of sustainable development

Picture 1a: Three-part Venn diagram of sustainable development



Picture 1b: Four-part Venn diagram of sustainable development



With horizontal addition of the fourth evaluation scope (H), the correlation of partial aggregates from meso 1 to meso 3 is extended by another full cycle. Extension inserts a whole new sub-layer to the synthesis algorithm. RP's secondary impacts are now presented with six dual overlaps (ES, EN, HE, SN, HS, HN), previously only three (Table 4). In addition, four triple overlaps are obtained (HSE, HNE, SNE, HSE, equation (2) below), previously only had one (SNE, or meso 3 in Picture 1a). And finally a quadruple overlap emerges (HNES, meso 3, picture 1b). The horizontal extension thus multiplies an overlapping area of meso 2 (Picture 1a) into two sub-levels: a meso 2a (with double overlaps) and meso 2b (with triple overlaps; Picture 1b).

The deepening of meso logic means that Dopfer et al.'s micro–meso–macro structuration may be radicalised as meso–meso–meso process. In the latter meso is not seen any more only as a transitory, intermediating level between micro and macro. Backed by complexity instead of evolutionary theory, meso approach turns out as an intrinsic approach for researching social issues, as Easterling and Kok have already emphasised.

Table 5: RP's impacts on sustainability of Pomurje region – four-part presentation

Table 5a: Four-part square matrix of RP's impacts

Policy \ Impact	E	H	S	N
E	E∩E = (+++)	E∩H = (+)	E∩S = (-)	E∩N = (+)
H	H∩E = (0)	H∩H = (+)	H∩S = (+)	H∩N= (0)
S	S∩E = (+++)	S∩H = (+)	S∩S = (++)	S∩N = (++)
N	N∩E = (+)	N∩H = (+)	N∩S = (+)	N∩N = (+)

Table 5b: Four-part correlation matrix of RP's impacts

Policy \ Impact	E	H	S	N
E	EE = (+++) → Very coherent	HE = (0, +) → Very weak correlation, unbalanced in favour of H	SE = (+++ , -) → S and E are strongly but negatively correlated, in favour of E	NE = (+, +) → Weakly correlated, balanced
H	-	HH = (+) → Weakly coherent	HS = (+, +) → Weakly correlated, balanced	HN = (0, +) → Very weak correlation, balanced
S	-	-	SS = (++) → Medium coherent	SN = (++, +) → Medium correlated, unbalanced in favour of N
O	-	-	-	NN = (+) → Weakly coherent

Source of data: Radej, 2006.

Synthesis in Four-capital model, this time going into reverse, starts where we aspire to conclude, in meso 3 with the overlap between four evaluation domains:

$$HNES = H \cap N \cap E \cap S.$$

HNES can be rewritten into overlap between three triple overlaps of meso 2b:

$$HNES = HSE \cap HSN \cap HNE \cap SNE. \quad (1)$$

(1) is practical because it translates four-part correlation matrix into four sub-matrices of the third order, which can be solved the same as in Table 4 on the level of meso 2a:

$$HSN = H \cap S \cap N = (H \cap S) \cap (H \cap N) \cap (S \cap N).$$

If further simplified by replacing $H \cap S = HS$:

$$HSN = HS \cap HN \cap SN. \quad (2)$$

Hence, triple, quadruple and overlaps of higher orders can be decomposed in the correlation matrix of double overlaps in meso 2a. In the case study, therefore, the formula of quadruple overlap (1) converts through (2) into the relationship between the bilateral overlaps:

$$HSNE = (HS \cap HN \cap SN) \cap (HS \cap HE \cap SE) \cap (HN \cap HE \cap NE) \cap (SN \cap SE \cap NE). \quad (3)$$

If simply due to idempotency of sets, HS ∩ HS = HS, etc., equation (3) is rewritten shorter:

$$HSNE = (HS \cap NE) \cap (HE \cap SN) \cap (SE \cap HN).$$

The evaluation synthesis of four-part area of the meso 3 is finally obtained with three overlapping factors, consisting of overlaps in meso 2b, which consist of a double overlaps area of meso 2a:

$$(HS \cap NE) \rightarrow \text{Overlap A},$$
$$(HE \cap SN) \rightarrow \text{Overlap B},$$
$$(SE \cap HN) \rightarrow \text{Overlap C}.$$

Three overlaps at the meso 2b sub-level turn into three distinctive composite meso indicators which present competing panoramic aspects of overall impact of RP on regional sustainability. Hence, the result on the highest level of aggregation is not singular as in summary row of LEM but multiple and hybrid. In a complex setting, what counts as a whole is not counted as one, but many. Meso synthesis produces a set of partial wholes (truths) which can not rule each other out. Each partial whole presents an integral perspective of sustainability, so none of them is sufficient indicator or could be left aside and ignored in evaluation. As Argentine philosopher Arturo Andres Roig has written, 'the truth is not found primarily in the totality, but in determinate forms of particularity with power to create and recreate totalities from a place outside the latter, as alterity' (Roig, 1983) – i.e. trough secondary meanings.

A plural outcome of synthesis is consistent with initial conceptualisation of complex social issues as incommensurable. In a way, this conclusion brings us back to the beginning, but the journey has not been in vain. Meso synthesis has been successful in transforming evaluation problem from incommensurable incompatibility at the beginning to weak integrity in hybrid content at the end. Case study has been previously helpful in illustrations of abstract claims, and so it is appropriate to ask it for assistance again.

Overlap A presents the “material aspect” of sustainability, because it describes the nature of the interaction between material (N and E) and non-material (H and S) content of regional sustainability. Analogously, overlap B presents “progressive aspect” while overlap C stands for “productive side” of sustainability.

Material overlap A can be described from the correlation matrix as:

$$HS = (+, +) \rightarrow \text{Weak correlation and}$$
$$NE = (+, +) \rightarrow \text{Weak correlation.}$$

So (HS ∩ NE) denotes one perspective of regional sustainability which is balanced, but only weakly correlated, between material and nonmaterial factors of regional sustainability. Weak link within nonmaterial aspect, HS, is not surprising if we take into account previously acknowledged

strategic weakness of regional social capital. More striking observation is how poorly RP performs in the materialist aspect (NE), which is prioritised in RP. Namely, meso 2b reveals that RP is inconsistently materialistic when prioritising only economic aspect, leaving N relatively aside (Table 5a). As a result, materialist character of RP is narrowed to economic aspects.

The second is a correlate B. It describes the nature of the interaction between progressive (H, E – the most responsive to stimulus on the short run) and conservation factors (S, N – slow in response and important to preserve for long-term future) of regional sustainability. Overlap between (HE ∩ SN) is evaluated as only weakly correlated. RP emphasizes conservation factors of regional sustainability. This is an expected reflex to observed systematic failure of regional authority to address the most urgent regional needs in their priorities.

The third and final correlate C describes the nature of the interactions between the produced factors of sustainability (S, E, which are created through permanent transformations) and non-produced factors (H and N, which are enhanced as they are – which is justified applying liberal and ecological argument). RP induce medium strong correlation between (SE ∩ HN). It significantly more (positively and negatively) affects produced than non-produced factors of regional sustainability.

From previous elaboration of three part model (Table 4) we are already aware of RP's asocial character, which can be confirmed with further observation of RP's internal inconsistency. RP attributes privilege to economic goals in primary as well as in secondary aspect, but not consistently. The same holds for RP's conservative aspect, which shows aversion for externally induced development, but the preconditions for endogenous development are nevertheless diminished by the poor cohesiveness between regional factors of sustainability.

Meso would deepen further when a fifth horizontal element, say culture, is added to the model. However, extension of meso may not necessarily help us better understand social complexity. Meso branching needs to stay within the hierarchy of "the moderate span" (Simon, 1962) – beyond which representation of system gets again too much complicated. How would you, for instance, conceptualise triple overlap between nature, culture and human vs. overlap between culture, economy and society? It is not that these overlaps do not exist, but they can always be decomposed to more simple, dual overlaps, for which three part presentation entirely suffices.

The complex methodology is not meant to propose a model which studies in one place all primary important substantive issues centrally, because this would collapse the assessment back to the old simplistic cumulative logic. Meso 1 to 3 simplify evaluation in a complex way. To retain complex logic, one needs to accept its multiplicity. To include culture in the study of regional development, one would do better to develop independent meso model with incommensurable cultural categories of evaluation, such as autonomy, demography and tradition

and so study these matters independently from the concept of sustainable development, or only in partial overlap with it.

5 Conclusions

This paper has comparatively explored the aggregation problem under social complexity, on the example of summation of conclusions in the policy impact evaluation. A range of methodologies is available explaining different approaches on how to summarise collected evidence. A meso synthesis is proposed as an intermediate possibility to prevailing micro and macro based approaches, based on simplistic assumption of commensurability of policy impacts. What elevates mesoscopic approach beyond the standard one is that it evaluates its object holistically, which means (i) with primary as well as secondary meanings (ii) both in horizontal as well as in vertical direction of social complexity.

Initially, Leopold et al. accurately perceived environment and economy as two principal and thus incommensurable domains of valuation; but they overlooked that their assessment perspective is verticalist, while the object of evaluation is spanned in horizontal direction. They were concerned only with secondary impacts, which are weakly commensurable, so they could have been partially aggregated at least from micro to meso level. On the other side, Ekins and Medhurst accurately observed weak commensurability of detailed assessment results. They appropriately allowed for partial aggregation only on the output side of the assessment, but failing to apply the same principle consistently also on the input as causal side of the evaluation model. Without construction of square matrix by source and area of impact, they have had very poor chances to locate weakly incommensurable impacts as a step which emancipates horizontal axis of evaluation; and this would be needed to carry on a synthesis from meso 1 to meso 2 level(s). So, LEM has properly distinguished vertical from horizontal axis of evaluation, but it missed to organise them orthogonally as simultaneous and equally important. In both standard cases difficulty has arisen from inconsistent organisation of primary/secondary evaluative meanings with horizontal/vertical axis of evaluation, which is, as linked to inconsistent application of incommensurability of values in evaluation of complex social object.

Mesoscopic evaluation demonstrated that social incommensurability is not an irresolvable obstacle to more holistic reasoning in public domain. Imperative of incommensurability establishes itself only as a safety mechanism which reminds evaluator that social issues are complex and cannot be explained as singular in their entirety from any specific point of view. However, strong and principal differentiations in evaluation of complex affairs are crucial for only a small number of the principal concerns. Even though contemporary societies are built on incommensurable oppositions which cause strong social fractures, the majority of issues important for the understanding of everyday social life are weakly in/commensurable. Social

members are in majority linked only by weak ties (Granovetter, 1983) which are of secondary importance to them.

Secondary meanings are in particular important in policy impact evaluation. When there is no straightforward mechanism to install an optimal public policy, a policy proposal that is the most secondary effective ought to be selected and implemented (see Demsetz, 1969). Despite this, secondary effects are routinely ignored, as being too complicated to evaluate. Systematic disregard for secondary impacts might explain why good sectoral (independent, verticalized) policies, based on strong values and even common sense, often lead to disappointing overall results (Chapman, 2004) horizontally.

That, which seems of secondary meaning (namely, influencing the others), validates itself as a key point for elaboration of the primary meaning (Althusser, in Levačić, 2009). This was evident already to Adam Smith, who built it into his key concept of the invisible hand of the market. Spontaneous extension of order (Hayek, 1992) is an achievement of evaluation of a tacit (or dispersed, secondary) knowledge, which is in itself useless to individuals and becomes meaningful only in interaction with others in the evolution of equilibrium price on the market, as its verticalised results. Hayek says our intentions and actions are one thing, but their broader effect is something completely different. The same idea is also relevant to the thought of Popper, who takes the view that the unintended consequences of action are the principal concern of social science and that the existence of such consequences is a precondition for the very possibility of the scientific understanding of a complex society (Vernon, 1976). If a person only did what he thought he is doing, the truth about society would be contained within a simple statement of their intentions. The same stance that secondary meanings are crucial for explaining complex social processes has been taken in the work of evolutionists.

There is an obvious overlap between evolutionary and complex justifications of secondary meanings in the understanding of social issues. However, an essential difference between the two must be highlighted. Gould argued that the evolution drive was not towards complexity, but towards diversification (1996b). For complexity theorist, emphasis on secondary meanings is never forwarded in isolation, but consistently appears in relation to primary concerns. Relying only on secondary meanings would lead to vague presentations of social processes. An evaluation of complex social issue must equally consider a large number of small consistencies (secondary view) and a small number of deep oppositions (primary view), to remain neutral. For Foucault, neither difference nor unity can be seen as primary, but need to be kept in balance (Fisk, 1993 in Olssen, 2002) or at least dealt with on the same plane (Althusser in Levačić, 2009) of inter-paradigmatic standards (Kordig, 1973). The imperative that social issues have to be evaluated in a complex way simply means that they have to be explained with primary meanings which are constitutive for it, so they can only be related horizontally, but in an incommensurable

and dividing way, as well as with its secondary meanings that are the only ones that lead vertically to a synthesis, but merely in contents that are not of primary importance to anyone.

Integration between primary (as tautologous; cf. Luhmann, 1989) and secondary (as paradoxical) meanings enhances evaluation capacities. Where evolutionaries saw some 'vital principle', supraphysical or metaphysical, but itself inexorably simple like invisible hand or evolutionary selection, we see complexity (Morin, 1974). A *deus ex machina* of evolutionary creativity can be better explained with bimodality of meso logic. The theory of complexity seems needed to further explain creativity of evolutionary process.

In a bimodal formula, primary oppositions are no more verticalised as in the standard model but horizontalized, while ephemeral relations in localised meanings are applied vertically with mixing and blending. In meso perspective, horizontality and verticality are inverted compared to the standard micro-macro as well as to micro-meso-macro program. When given social positions are presented vertically, they are evaluated horizontally. And the opposite, although indirect impacts initially emerge in horizontal direction, they are evaluated in vertical direction by their gradual synthesis. In other words, the particularism in meso setting is increasingly expressed holistically, while universalisms are particularized with horizontal pluralisation into multichotomous categories (Bailey, 2006). What we achieved with our evaluation experiment is a procedure for complex simplification where meso decomposes triadic relationship into a set of dual relationships which take place no more between homogenous antagonisms but between oppositions with heterogeneous content. This meso transformation constructs a perspective where every polarity can be translated into multiple reality and every multiplicity can be mesoscopically decomposed to dyadic relations.

According to von Neumann (in Morin, 1974), complexity is not about using existing logic in a modified fashion of ‘logical complexity’, such as micro-meso-macro, but about a new, radically mesoscopic and synthesising ‘logic of complexity’. A mesoscopic way of reasoning about social complexity certainly does not eliminate the linear and simple way of thinking, but only confines it to simple framework. Non-complex perspectives are entirely valid path to researching fragmented facts and partial truths. As Alfred Marshall argued, simplistic models can be enlightening in studying very local contexts (in Dasgupta, 1997). Focusing on one level of analysis is fine as long as one does not make assumptions or inferences about other levels of analysis (McConnell, Moran, 2000). Simplicity belongs to this world, but it is not its fundamental characteristic and can not be attributed to everything (Prigogine, Stengers, 1982). Social processes unravel both on the micro and on the macro levels, but descriptions of the processes on two levels cannot be directly linked. Microscopic issues can only be approached microscopically and macroscopic issues macroscopically. In both cases a uniform way of thinking is used that is simple. Both single level and single scope-based explanations of social

issues are valid within their narrow frame, but they can be coherently explained together only from a third, intermediating or meso level (Dopfer et al., 2004). The inevitability of logic of complexity does not negate traditional logic within the sphere in which it is operational, but sublates it, or retains it while integrating it into a richer logic (Morin, 1974) at the higher level of generality and deeper level of understanding.

Bibliography

Adams R. N. 2009. Energy, Complexity and Strategies of Evolution. Avá n.14 Posadas, July.

Allport F. H. "Group" and "Institution" as Concepts in a Natural Science of Social Phenomena. Publications of the American Sociological Society, 23(1928):83-99.

Arrow K. 1951. Social Choice and Individual Values, (2nd ed., 1963), New Haven: Yale University Press. 124 pp.

Bailey K. D. Systems Theory, in Turner J. H. (ed.), Handbook of Sociological Theory. Springer, 2006.

Baker R. 2012. The Seven Moral Hazards of Measurements. VeraSage institute – Blog, November 2.

Barroso J., C. J. Gollop, M. Sandelowski, J. Meynell, P.F.Pearce, L.J. Collonis. The challenges of searching for and retrieving qualitative studies. Western Journal of Nursing Research, 25(2003):153-178, in Bonds, 2007.

Bar-Yam Y. A Mathematical Theory of Strong Emergence Using Multiscale Variety. Wiley, Complexity, 9/6 (2004):15-24.

Benton T. 1998. Naturalism in social science, in E. Craig (ed.), Routledge Encyclopedia of Philosophy. London: Routledge.

Blundell R., T. M. Stoker. Heterogeneity and Aggregation. Journal of Economic Literature, 43(June 2005):347–391.

Bovens M., P. 't Hart, S. Kuipers. 2008. The politics of policy evaluation. The Oxford Handbook of Public Policy, Chapter 15, pp. 317-33.

Buchanan A., R. O. Keohane. ―The Legitimacy of Global Governance Institutions. Ethics & International Affairs 20/4(2006):405-37.

Burchell G., C. Gordon, P. Miller (eds). 1991. The Foucault Effect: Studies in governmentality. Hertfordshire: Harvester Wheatsheaf, 303 pp.

Carly F. 2004. Information: the currency of the digital age. San Francisco: Oracle OpenWorld, December 6.

Cat J. 2010. The Unity of Science, in Edward N. Zalta (ed.), The Stanford Encyclopedia of Philosophy.

Chapman J. 2004. System failure – Why governments must learn to think differently. Second edition, London: Demos.

Common Assessment Framework. 2006. Brussels: EIPA – European Institute of Public Administration, 51 pp.

Connelly S. T. Richardson. 2004. Exclusion: The necessary difference between ideal and practical consensus. Journal of Environmental Planning and Management 47/1(January 2004):3-17.

Cremaschi S., M. Dascal. Malthus and Ricardo on Economic Methodology. History of Political Economy, 28/3(1996):475-511.

Dasgupta P. 1997. Notes on Slow and Fast Variables in Economics. Notes for the Resilience Network, University of Florida, Department of Zoology, Gainesville.

Demsetz H. Information and Efficiency: Another Viewpoint. Journal of Law and Economics 12(1969):1-22, in Schnellenbach, 2005.

Derrida J. 2002. Positions. London & New York: Continuum, 2nd edition.

Diamond J. Establishing a Performance Management Framework for Government. Instituto de Estudios Fiscales, Presupuesto y Gasto Público, Working Paper 40(2005):159-83.

Donzelot J. 1991. The Mobilisation of Society, in Burchell et al., p. 169-79.

Dopfer K. 2006. The Origins of Meso Economics Schumpeter's Legacy. Jena: Max Planck Institute of Economics, Evolutionary Economics Group, The Papers on Economics and Evolution No. 610, 44 pp.

Dopfer K., J. Foster, J. Potts. Micro–meso–macro. Springer, Journal of Evolutionary Economy 14/3(2004):263–79.

Easterling W.E., K. Kok. Emergent Properties of Scale in Global Environmental Modelling – Are There Any? Integrated Assessment 3/2–3(2002):233–46,

Ekins P. 1992. A Four-Capital Model of Wealth Creation in Real-Life Economics: Understanding Wealth Creation (in Ekins P., M. Max-Neef, eds.). London, New York: Routledge, p. 147-55.

Ekins P., Medhurst J. 2003. Evaluating the Contribution of the European Structural Funds to Sustainable Development. Presented at the 5th European Conference on Evaluation of Structural Funds, Budapest, June 26-27, 48 pp.

Ekins P., Medhurst J. The European Structural Funds and Sustainable Development: A Methodology and Indicator Framework for Evaluation. Sage: Evaluation, 12/4(October 2006):474-95.

Elbers C., J. W. Gunning, K. de Hoop. Assessing Budget Support with Statistical Impact Evaluation: a Methodological Proposal. Amsterdam: Tinbergen Institute, Discussion Paper 2007-075/2, 29 pp.

Elsner W. 2007. Why Meso? On 'Aggregation' and 'Emergence' and Why and How the Meso Level is Essential in Social Economics. Springer, Forum for Social Economics, 36/1(2007):1-16.

Elsner W. The process and a simple logic of 'meso'. Emergence and the co-evolution of institutions and group size. Springer: Journal of Evolutionary Economics, 20/3 (June 2010):445-77.

Encyclopedia of Evaluation (Synthesis). 2004. SAGE Publications.

ESPON - 3.2, Vol. 5. 2006. Spatial Scenarios and Orientations in Relation to the ESDP and Cohesion Policy – Territorial Impact Assessment, Final Report, p. 2-97.

European Commission. 2001. SEA Directive 2001/42/EC on the assessment of the effects of certain plans and programmes on the environment.

European Commission. 2006. Impact Assessment Guidelines. SEC(2005)791.

Faris E. Sociology and Human Welfare. Social Forces 18(1939):1-9.

Foster J. 2004. Why is Economics not a Complex Systems Science? The University of Queensland, School of Economics: Discussion Paper No. 336, 38 pp.

Foster J., Potts J. 2007. A micro-meso-macro perspective on the methodology of evolutionary economics: integrating history, simulation and econometrics. The University of Queensland, School of Economics, Discussion Paper No. 343, 24 pp.

Funtowicz S. O., J. R. Ravetz. The worth of a songbird: ecological economics as a post-normal science. Ecological Economics 10(1994):197–207, in Martinez-Alier et al., 1998.

Geels F. W. Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. Elsevier, Research Policy 31/8-9(2002):1257-74.

Geertz C. 2000. Available Light. Princeton University Press, 271 pp., in Renselle D. 2007. Review of Clifford Geertz' Available Light. Carmel: Quantonics.

GHK, PSI, IEEP, CE, National Evaluators. 2002. The Contribution of the Structural Funds to Sustainable Development: A Synthesis Report to DG Regio, EC. 2002. Volume 1-2. London, Brussels.

Giddens A. 1976. New Rules of Sociological Method: a Positive Critique of interpretative

Sociologies. London: Hutchinson.

Gontier N. Evolutionary epistemology and the origin and evolution of language: Taking symbiogenesis seriously. Springer: Evolutionary Epistemology, Language and Culture (2006), p. 195–226.

Gould S. J. 1996a. The Mismeasure of Man: Revised and Expanded Edition. New York: W. W. Norton & Co.

Gould S. J. 1996b. Full House: The Spread of Excellence From Plato to Darwin. New York: Harmony Books.

Granovetter M. The Strength of Weak Ties: A Network Theory Revisited. Sociological Theory 1(1983):201-33.

Guba E., Y. S. Lincoln. 2000. Competing paradigms in qualitative research, in Denzin, Lincoln (ed.). Handbook of Qualitative Research, 2nd ed., London in New Delhi: Sage, pp. 105–17.

Gutiérrez F., D. Buitrago, A. González. Aggregating Political Dimensions: Of the Feasibility of Political Indicators. Springer Science, Social Indicators Research, An International and Interdisciplinary Journal for Quality-of-Life Measurement, 110/1(2013):305-326.

Hayek F. A. The Fatal Conceit: The Errors of Socialism, The University of Chicago Press. 1991.

Hertin J., A. Jordan, M. Nilsson, B. Nykvist, D. Russel, J. Turnpenny. 2007. The practice of policy assessment in Europe: An institutional and political analysis. EU/FP6 Project MATISSE Working Paper 6, 52 pp.

Heylighen F. 1989. Self-Organization, Emergence and the Architecture of Complexity, in: Proceedings of the 1st European Conference on System Science, (AFCET, Paris), p. 23-32.

Hurley S. L. Rationality, democracy, and leaky boundaries: vertical vs. horizontal modularity. Journal of Political Philosophy, 7/2(1999):126-46.

IAB. 2009. Impact Assessment Board report for 2008. Third strategic review of Better Regulation in the European Union. SEC(2009)55, 17 pp.

Kimberlé C. W. 1994. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color, in Fineman, M.A., R. Mykitiuk (eds.). The Public Nature of Private Violence. New York: Routledge, pp. 93-118.

Kordig C. R. Discussion: Observational Invariance. Philosophy of Science, 40(1973):558-69.

Krossa W. Creating A Horizontal God, retrieved May 2006 from http://home.istar.ca/~wkrossa/articles.html.

Kuhn T. S. 1970. The Structure of Scientific Revolutions. Chicago: The University of Chicago Press, 210 pp.

Leith C., L. von Thadden. 2006. Monetary and fiscal policy interactions in a new Keynesian model with capital accumulation and non-Ricardian consumers. Working Paper Series No. 649, 42 pp.

Leontief W. Environmental Repercussion and the Economic Structure: An Input-Output Approach. Cambridge Mass.: The MIT Press, Review of Economics and Statistics, 52/3(1970):262-71.

Leopold L. B., F. E. Clarke, B. B. Hanshaw, J. R. Balsley. 1971. A procedure for evaluating environmental impact. Washington: Geological Survey Circular 645, 13 pp.

Lipsey M. W. Identifying interesting variables and analysis opportunities, in H. Cooper, L. V. Hedges, & J. C. Valentine (eds). The Handbook of Research Synthesis and Meta-Analysis (2nd edition). NY: Russell Sage Foundation, 2009. Chapter 8, pp. 147-58.

List C., B. Polak. 2010. Introduction to judgment aggregation. Yale University, Cowles Foundation Discussion Paper No. 1753.

Lucas R. E. Jr. Expectations and the neutrality of money. Blackwell, Journal of Economic Theory, 4/2(1972):103-24.

Luhmann N. 1989. Ecological communication. Cambridge, England: Polity Press.

Martinez-Alier J., G. Munda, J. O'Neill. Weak comparability of values as a foundation for ecological economics. Elsevier, Ecological Economics 26(1998):277-86, in Stagl S. 2007. Emerging Methods for Sustainability Valuation and Appraisal – SDRN Rapid Research and Evidence Review. A report to the Sustainable Development Research Network. Final Report (January), 66 pp.

Mason E. 2006. Value Pluralism; entry in Stanford Encyclopedia of Philosophy (in Zalta N.E, ed.), Stanford: Centre for the Study of Language and Information, Stanford University.

McConnell W. J., E. F. Moran (ed.). 2000. Meeting in the Middle: The Challenge of Meso-Level Integration. Ispra, EU Joint Research Centre, An International Workshop (Land Use and Cover Change Project), October 17-20, LUCC Report Series No. 5.

Mertens D. M. Inclusive Evaluation: Implications of Transformative Theory for Evaluation. Sage, American Journal of Evaluation, 20/1(1999):1-14.

Michalos A. C. What Did Stiglitz, Sen and Fitoussi Get Right and What Did They Get Wrong? Social Indicators Research. An International and Interdisciplinary Journal for Quality-of-Life Measurement, Springer, 102/1(2011):117-29.

Morçöl G. 2011. Why Complexity Theory? What Kind of a Theory Is It? Presented at the governance of complex systems conference, Rotterdam, the Netherlands, June 23-25.

Morgan D. L. Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. Sage, Journal of Mixed Methods Research, 1/1(2007):48-76.

Morin E. Complexity. UNESCO: International Social Science Journal, 26/4(1974):555-82.

Mouffe C. 2000. The Democratic Paradox. London & New York: Verso, in Trainor B.T. Politics as the quest for unity. Perspectivism, incommensurable values and agonistic politics. Sage, Philosophy & Social Criticism, 34/8(2008): 905–24.

Mumpower J. L., T. R. Stewart. Expert Judgement and Expert Disagreement. Taylor and Francis, Thinking and Reasoning, 2/2-3(1996):191-211.

Munasinghe M. 1992. Environmental Economics and Sustainable Development. Paper presented at the UN Earth Summit, Rio de Janeiro, Washington: World Bank, Environment Paper No. 3.

Munda G. 2006. A NAIADE based approach for sustainability benchmarking. Inderscience, International Journal of Environmental Technology and Management, 6/1-2(2006):65–78.

Munda G. Choosing Aggregation Rules for Composite Indicators. Social Indicators Research - An International and Interdisciplinary Journal for Quality-of-Life Measurement, Springer, 109/3(2012): 337-354

Munda G. Social multi-criteria evaluation: Methodological foundations and operational consequences. European Journal of Operational Research, 158/3(2004):662-77.

Muradian R. Ecological thresholds: a survey. Elsevier, Ecological Economics, 38/1(2001):7–24.

Nola R., H. Sankey (eds). 2000. After Popper, Kuhn & Feyerabend: Issues in Theories of Scientific Method. Australasian Studies in History and Philosophy of Science, Kluwer.

Olsen R. 2006. Vertical to horizontal: a new workplace reality. Members for Democracy Archive. www.uncharted.ca/content/view/149/21/

Olssen M. Michel Foucault as "Thin" Communitarian: Difference, Community, Democracy. Cultural Studies, 2/4(2002):483-513.

Ooms G., W. Van Damme, B.K. Baker, P. Zeitz, T. Schrecker. The 'diagonal' approach to Global Fund financing: a cure for the broader malaise of health systems? Globalization and Health 4/6(March 2008).

Ostmann A. 2006. The aggregate and the representation of its parts. Bonn: Max Planck Institute for Research on Collective Goods, Preprints of the Max Planck Institute for Research on Collective Goods 2007/11b, 38 pp.

Prigogine I. 1996. La Fin des certitudes: certitudes: Temps, chaos et les lois de la nature. Odile Jaco.

Prigogine I., I. Stengers. 1979. La nouvelle alliance: Métamorphose de la science, Paris: Gallimard.

Prins B. Narrative Accounts of Origins: A Blind Spot in the Intersectional Approach? European Journal of Women's Studies 13/3(2006):277-90.

Radej B. 1995. Onesnaženje naprodaj. Ljubljana: Urad za makroekobnomske analize in razvoj.

Radej B. 2006. Assessment of structural funds effectiveness on the sustainable development: Pomurje regional case study. Final report, 2006, 6th Framework program – Framework project SDRTOOLS: Regional sustainability Appraisal: Developing evaluation methods and sustainable policies (Contract SCS8-CT-2004-502485), 80 pp.

Radej B. 2013. With the exclusion to the Community. Ljubljana: Slovenian Evaluation Society, Working Papers, 6/2(July 2013).

Radej B., M. Golobič. Divided we stand: Social integration in the middle. Ljubljana: Slovenian Evaluation Society, Working papers, 6/1(2013), 24 pp.

Riofrio W. Ubiquity symposium: Evolutionary computation and the processes of life information, biological, and evolutionary computing. Ubiquity, May(2013):1-10.

Ritzer G., B. Smart. 2003. Introduction: Theorists, Theories and Theorizing, in Ritzer G., B. Smart (ed.). Handbook of Social Theory. Sage.

Roig, A. A. 1983. Teoría y crítica pensamiento latinamericano, Mexico: Fondo de Cultura Economica, in E. Craig (ed.), Routledge Encyclopedia of Philosophy. London: Routledge.

Rotmans J. Scaling in Integrated Assessment: Problem or Challenge? Swets & Zeitlinger, Integrated Assessment, 3/2–3(2002):266–79.

Rotmans J. Tools for Integrated Sustainability Assessment: A two-track approach. Vancouver: University of British Columbia, The Integrated Assessment Journal, 6/4(2006):35–57.

Sadler B., R. Verheem. 1996. Strategic environmental assessment: Status, challenges and future directions. Hague: Ministry of Housing, Spatial Planning, the Environment. 188 pp.

Sanderson I. Evaluation in complex policy systems. Sage, Evaluation, 6/4(2000):433–54.

Sankey H. The Problem of Rational Theory-Choice. Epistemologia, 18(1995):299-312.

Schnellenbach J. The Dahrendorf hypothesis and its implications for (the theory of) economic policy-making. Cambridge Journal of Economics, 29/6(2005):997-1009.

Schwandt T. A. Evaluation as Practical Hermeneutics. Evaluation 3/1(1997):69–83.

Scriven M. The Final Synthesis. Sage, American Journal of Evaluation, 15/3(1994):367-82.

Soja E. W. 2000. Postmetroplis: Critical Studies of Cities and Regions. Blackwell Publishing.

Stake R. E. A Problematic Heading. American Journal of Evaluation; 22/3(2001):349–354.

Steward J. 1955. Theory of Culture Change: The Methodology of Multilinear Evolution. Urbana: University of Illinois Press.

Turner J. H. (ed.). 2006. Handbook of Sociological Theory. Springer.

UNCED. 1992. Agenda 21. United Nation Conference on Environment and Development.

Vedung E. Four Waves of Evaluation Diffusion. Sage, Evaluation, 16/3(July 2010):263-27.

Veen A. Van Der, H. S. Otter. Scales in Space. Swets&Zeitlinger, Integrated Assessment, 3/2–3(2002):160–6.

Vernon R. The “Great Society” and the “Open Society”: Liberalism in Hayek and Popper. Canadian Journal of Political Science/Revue canadienne de science politique, 9/2(1976):261-76.

Virtanen P., P. Uusikylä. Exploring the Missing Links between Cause and Effect. A Conceptual Framework for Understanding Micro–Macro Conversions in Programme. Sage, Evaluation, 10/1(2004):77–91.

Wacquant L. 1997. Reading Bourdieu's “Capital”. Foreword to the English-language translation of The State Nobility. Cambridge: Polity Press.

Wallerstein I. 2004. The Uncertainties of Knowledge. Philadelphia: Temple University Press, 224 pp.

WCED. 1987. Our Common Future. Oxford and New York: Oxford University Press, World Commission on Environment and Development – WCED, 400 pp.

Weaver P. M., J. Rotmans. 2006. Integrated Sustainability Assessment: What? Why? How? EU/FP6 Project MATISSE, Working Paper 1 (October), 22 pp.

Wiggins D. 1997. Incommensurability: Four Proposals, in Chang R. (ed.), Incommensurability, Incomparability, and Practical Reason, Cambridge, Mass.: Harvard University Press.

Woese C.R. A New Biology for a New Century. Microbiology and Molecular Biology Reviews, 68/2(June 2004).

COLOPHON

Naslov:	»Apples and Oranges: Synthesis without a common denominator«
Podatki o avtorju:	Radej Bojan
Podatki o izdaji ali natisu:	1. Izdaja
Kraj in založba:	Ljubljana: Slovensko društvo evalvatorjev / Slovenian Evaluation Society
Leto izida:	2014 (Vol. VII, No. 1)
Naslov knjižne zbirke:	Delovni zvezek SDE/ Working paper SES
Podatke o nosilcu avtorskih pravic:	SDE, Ustvarjalna gmajna 2.5/Creative Commons 2.5, Slovenija
Podatek o nakladi (število natisnjenih izvodov):	Elektronska publikacija,
Mednarodne identifikatorje (ISBN, ISMN, ISSN):	ISBN 978-961-92453
Maloprodajno ceno publikacije:	Publikacija je brezplačna/Free.

About the SES's Working papers series

Publishes scientific and technical papers on evaluation of public policies and from related disciplines. Papers are reviewed and bibliographic catalogued (CIP). Freely accessible on the internet. Already published:

Volume 1 (2008), no:

Excercises in Aggregation (In Slovenian; B. Radej, 23 pp) **1**
Synthesis of Territorial Impact Assessment for Slovene Energy Programme (In Slovenian; B. Radej, 43 pp) **2**
Meso-Matrical Synthesis of the Incommensurable (B. Radej, 21 pp) **3**

Volume 2 (2009), no:

Anti-systemic movement in unity and diversity (B. Radej, 12 pp) **1**
Meso-matrical Impact Assessment - peer to peer discussion of the working paper 3/2008(B. Radej, 30 pp) **2**
Turistic regionalisation in Slovenia (in Slovenian, J. Kos Grabar, 29 pp) **3**
Impact assessment of government proposals (In Slovenian; B. Radej, 18 pp) **4**
Performance Based Budgeting (In Slovenian; B. Radej, 33 pp) **5**

Volume 3 (2010), no:

Beyond »New Public Management« doctrine in policy impact evaluation (B.Radej, M.Golobič, M.Istenič) **1**
Basics of impact evaluation for occassional and random users (In Slovenian; B. Radej) **2**

Volume 4 (2011), no:

Intersectional prioretisation of social needs (In Slovenian; B. Radej, Z. Kovač, L. Jurančič Šribar, 45 pp) **1**
Primary and secondary perspective in policy evaluation (In Slovenian; B. Radej, 30 pp) **2**
Aggregation problem in evaluations of complex matters (In Slovenian; B. Radej, 41 pp) **3**
Movement 99%: With exclusion to the community (In Slovenian; B.Radej, 42 pp) **4**

Volume 5 (2012), no:

Excelence squared: self-assessment in public administration with CAF (In Slovenian; B.Radej, M.Macur) **1**
Partial whole: example of terrotirial cohesion (In Slovenian; B.Radej, M.Golobič, 31 pp) **2**

Volume 6 (2013), no:

Divided we stand: Social integration in the middle (B.Radej, M.Golobič, 26 pp) **1**
With Exclusion to the Community (B.Radej, July 2013) **2**

Volume 7 (2014), no:

Apples and Oranges: Synthesis without a common denominator (B.Radej, February 2014, 40 pp) **1**

About the authors

Bojan Radej is a methodologist in social research. Master degree in macroeconomics, University of Ljubljana - Faculty of Economics (1993). Professional Experience Record: Governmental Institute of Macroeconomic Analysis and Development (1987-04; under-secretary to the government), areas of work: sustainable development (1998-04), chief manager of the modelling department (1993-5); initiator and the first editor of Slovenian Economic Mirror (1995-8); Co-Editor of IB journal (2001-04). Founder of Slovenian Evaluation Society.

Supporters

SDE operation is nonfinacially supported by IER - Institut for economic research, Ljubljana,
http://www.ier.si/