Working Papers

SLOVENIAN EVALUATION SOCIETY

Vol 17 no 1 year 2024

Inclusive or Rational Participatory Evaluation?

Bojan RADEJ

Ljubljana, December 2024

info@sdeval.si http://www.sdeval.si

"***Evaluations are still predominantly those conducted for the purpose of donor accountability rather than to address the people's needs***."

Global Evaluation Agenda 2025, forthcoming, with permission IOCE and EvalPartners (XI 2024)

Bojan Radej

Inclusive or Rational Participatory Evaluation?

Abstract: *The problems of ensuring neutrality in participatory evaluation of complex interventions are explored. Four tools applying participatory processes are assessed – Most Significant Change, Causal Mapping, SenseMaker, and Outcome Harvesting – for their neutrality assessed as their contributions to social inclusion and collective rationality. The former relies on effectively managing bias in participatory processes, while the latter depends on the capacity to aggregate participants' inputs at the collective level. This paper finds that the tools try to suppress or bypass bias rather than engage with it productively. Furthermore, they either refuse the synthesis of findings or overgeneralise. A new approach to participatory evaluation is proposed. It is framed within 'the empty middle,' where all identified biases converge. It is inclusive because it evaluates indeterminate (complex) issues as blindsighted (the opposite of enlightened). It is collectively rational because its meta-level aggregation better represents collective concerns in complex conditions than macro-level aggregation.*

Keywords: Participatory evaluation, Social inclusion, Aggregation problem, Empty middle, Blindsight

Vključevalna ali racionalna participatorna evalvacije

Povzetek: *Proučeni so problemi zagotavljanja nevtralnosti v participativnih evalvacijah kompleksnih intervencij. Štiri orodja participativnega vrednotenja so ocenjena glede na nevtralnost s stališča njihovega prispevka k socialni vključenosti in kolektivni racionalnosti – Most Significant Change, Causal Mapping, SenseMaker in Outcome Harvesting. Prvo se nanaša na ustreznost upravljanja pristranskosti v participativnih procesih, drugo pa na verodostojnost agregacije participativnih prispevkov s posamične na kolektivno raven. Članek ugotovi, da so orodja nezadostna pri vrednotenju kompleksnih intervencij: ali zavračajo sintezo rezultatov ali preveč posplošujejo. Pristranskost poskušajo zaobiti, zatreti ali se ji podrediti, namesto jo produktivno uporabiti. Nov pristop je uokvirjen v prazni sredini, kjer se stekajo identificirane pristranskosti. Je vključujoč, ker nedoločene (kompleksne) stvari vrednoti kot slepoviden (nasprotje znanstvene razsvetljenosti), in kolektivno racionalen na meta-ravni kolektivno ki je v nedoločljivih situacijah bolj nevtralno izpovedna kot makro raven.*

Acknowledgements:

(1) The author is grateful to Bojan Keevill for the language editing of the final text and to Dr. Steve Powell, Dr. Rick Davies, and Michael Quinn Patton for their feedback on an earlier version of this text.

(2) AI tools Google Translate, Grammarly, ChatGPT, and Claude were utilised for the literature review, translation from Slovene to English, and text editing.

Ključne besede: Participativna evalvacija, Družbena vključenost, Agregacijski problem, Prazna sredina, Slepovidnost

Acknowledgements:

(1) The author expresses gratitude to Bojan Keevill for the language editing of the final text and to Dr. Steve Powell, Dr. Rick Davies, and Michael Quinn Patton for their feedback on an earlier version of this text.

(2) AI tools Google Translate, Grammarly, ChatGPT, and Claude were utilised for the literature review, translation from Slovene to English, and text editing.

Disclaimer: The statements made in the article are consistent with the spirit of the Founding Charter of the Slovenian Society of Evaluators (https://www.sdeval.si/2008/12/07/ustanovna-izjava/), but do not represent its official views.

Proposal for citation: Radej B. *Inclusive or Rational Participatory Evaluation?* Ljubljana, Slovenian Evaluation Society, Working Papers no. XVII/1(Dec 2024), 41 pp. https://www.sdeval.si/2025/01/08/inclusive-or-rational-participatory-evaluation/

* * *

Vernon Press published a book of the same author in 2021 ("*Complex Society: In the Middle of a Middle World*," with M. Golobič). https://vernonpress.com/book/1083

I. Introduction

Traditional result-based and reason-centred impact evaluations, which compare a policy, program, or project's actual outcomes to its initial benchmarks, prove inadequate in evaluation of complex interventions (Powell et al.; Deprez; Wilson-Grau; Davies, Dart). The impacts of complex interventions, as multifaceted phenomena, develop from a network of intricate interactions, making impact attribution to a specific cause or beneficiary impossible. Therefore, the impacts are incommensurable – lacking a common metric for comparison – so they resist complete and uniform descriptions. This is particularly true for interventions promoting sustainable development, catalysing human and social behaviour change, fostering social transformation, enhancing group interactions, or developing internal group dynamics (Wilson-Grau).

Over the past few decades, evaluation theory has developed beyond the result-based tradition. Theories, grounded in postmodern reasoning, have redirected the focus toward stakeholder-driven and design-based or constructivist approaches, centred on dialogical participatory evaluation (Patton, 2002). Constructivist epistemology – an explanation of how we know things – goes beyond simple rationalist-realist models. Emerging with the ‘fourth generation’ of approaches (Guba, Lincoln) constructivists accentuate collective sensemaking through dialogue, offering deeper contextual insights into complex issues. Stufflebeam defines participatory evaluation as a collaborative assessment process that appraises multifaceted interventions by engaging a wide range of stakeholders with dissimilar perspectives and considering various data types, both quantitative and qualitative.

However, the participatory approach, as subjective, often fail to secure neutrality in evaluation. In the absence of absolute criteria for assessing neutrality, this paper assesses it relative to the contributions of various participatory evaluation approaches to social inclusion and collective rationality about what is best for all. While it strives to open evaluation to multiple viewpoints, the participatory process is not always inclusive. Citizens are not included as equals in unequal societies. Besides, participatory evaluation often lacks a strong theoretical foundation, often sacrificing rationality for inclusivity. Its results are often biased (Patton, 2012) or unrealistic (Radej, Golobič, 2020) and cannot consistently contribute to the collective good. In the assessment of participatory processes, quantitative criteria – such as the notion that more is better – should be abandoned and replaced with more authentic considerations.

The significant limitations of both approaches – objectivist and subjectivist – necessitate a more comprehensive evaluation framework that is simultaneously democratic and collectively rational. This situates evaluative challenges within the scope of collective choice theory, which examines how rationally and inclusively groups navigate conflicting views and derive collective decisions or sensemaking from fragmented, partial, and biased individual requirements. Yet, collective choice faces a fundamental constraint. The famous Arrow's impossibility theorem (Nobel Prize in economics, 1972) asserts that collective choice cannot be fully democratic and rational simultaneously. If we want to increase the sensitivity of collective decisions to collective (and individual) differences, decisions will be less

collectively rational, or vice versa. Policy impact evaluation emerged to somehow overcome the immobilising implications of impossibility theorem in collective choice theory and practice.

The challenge of enhancing both inclusiveness and collective rationality correlates with the broader field of democracy. The concept of democracy intersects with the intrinsic tension between freedom and order – between *demos* and *kratos*. In Greek, *demos* represent the collective body of citizens, who prioritise inclusivity and a diversity of perspectives free from top-down obstruction. *Kratos*, on the other hand, refers to the governance structures that enable decision-making. Every entity or system has its *kratos*, ensuring that collective choices are coherent, just, and functional. To fulfil its mission, *kratos* marginalises certain less typical viewpoints, enforcing one-sided or narrow-minded unification while excluding elements that are irrelevant to its logocentric self (Derrida).

Policy impact evaluation, as an instrument of deliberative democracy, serves as a means for the *demos* to influence *kratos*. Simultaneously, it guides *kratos* to become less exclusionary. Focusing on inclusivity without collective rationality risks devolving participation into a cacophony of voices lacking a coherent narrative – analogous to *demos* without *kratos*. Conversely, overemphasising community needs in collective choice that neglect members' aspirations leads to rational but undemocratic outcomes – *kratos* without *demos*.

Accordingly, the paper begins with two imperatives of participatory evaluation tools: they must be socially inclusive and collectively rational. Inclusivity means everyone contributing to the collective goods can effectively participate in evaluation. Inclusivity as a driver of neutrality depends on the effectiveness of tools in addressing epistemic blindness (Fricker), or simply subjectivity and bias in participative processes. Epistemic blindness results from a situation that is the opposite of epistemic certainty. Blindness is caused by ignorance when excluding anything that does not fit within the logocentric understanding of things (Kahneman). Some biases can be overcome through acquiring new information, learning, or by mitigating prejudices and stereotypes. Others, however, are enduring and cannot be abolished, particularly when dealing with complex phenomena (partly) rooted in uncertainty. Therefore, the most inclusive evaluation tool is the one that can most inclusively deal with the participants' most varied and biased contributions to collective choice.

The imperative of collective rationality requires that the participants in the evaluation recognise which contribution is better for all, such as enhancing sustainable social benefits. This depends on how effectively they aggregate the fragmented and often contradictory contributions gathered through the participatory process. The aggregation framework provides rules for evaluating trade-offs, advantages, and disadvantages, with the aim of safeguarding the legitimacy of collective choices.

Various aggregation approaches exist, ranging from the micro-level (a selective description of the whole based on individual data points) to the macro-level (a top-down perspective addressing broader concerns), the meso-level (explanation derives from the intersection between evaluation domains), and the meta-level (the interpretation of shared concerns). Different approaches produce strikingly dissimilar aggregates, which are not equally rational

at the collective level (Radej, 2021a). Thus, the selection of the aggregation method should be always justified in participatory evaluation by explaining why it is considered superior in identifying the greatest collective good.

Inadequate handling of epistemic blindness and aggregation leads to exclusionary and sense-narrowing participatory evaluations of complex interventions.

To explore this thesis, the paper examines four established evaluation tools as illustrative examples: SenseMaker by Cognitive Edge, Outcome Harvesting by Wilson-Grau, Most Significant Change by Rick Davies, and Causal Mapping (Axelrod, Powell). These tools acknowledge an objective reality (rationalist-realist aspect) whilst emphasising the importance of subjective interpretations (constructivist aspect). They are constructivist-realist – intersectional between the realist material conditions of society and the discursive constructivist practices that shape understanding of these conditions. The intersectional epistemology recognises that while certain social facts or realities e.g., poverty, cohesion, or health may exist independently, understanding these issues and finding solutions requires cross-sectional evaluation. It aims to connect the strengths of both perspectives and seize the best of opposing worlds. The hybrid acknowledges distinctive participants' constructions of reality while emphasising the crucial role of a rational description of the world (Pawson, Tilley).

The paper introduces four realist-constructivist tools and then presents two criteria for their assessment. It then explores if these tools satisfy the two criteria in fulfilling the imperatives of participatory evaluations of complex interventions that ensure its neutrality. Whilst acknowledging the strengths and positive contributions of these four tools, they fail to fulfil the mission in complex circumstances. They largely promote an exclusionary concept of the collective good from the viewpoint of a privileged minority. Prevailing evaluation practices often serve donor accountability rather than addressing people's needs (Global Evaluation Agenda 2025; IOCE, EvalPartners). The assessment is concluded with a call for an anti-postmodern turn (Badiou, Žižek) in evaluation theory that achieves inclusiveness and collective rationality by intersecting their drivers in the empty middle and interpreting the obtained findings as blindsighted.

The concept of the empty middle (Radej, 2021a) is novel in evaluation theory. The core idea is grounded in Eastern philosophy of the middle way (Nagarjuna, Tsongkhapa, The Kyoto School), the notion of nothingness in Western philosophy (Heidegger, Sartre, Lacan, Derrida), and anti-postmodern philosophy. The concept of the empty middle claims that in democratic societies a central space in social discourse must remain ostensibly vacant, rendering it resistant to strategic occupation and manipulation for asserting meaning or authority. Derrida wrote about engaging with the *aporia* or gaps within dominant frameworks. In *The Parallax View*, Žižek elucidates how the same object can be perceived differently from various perspectives, highlighting the importance of engaging with these indeterminacies to achieve a more comprehensive, albeit always incomplete, understanding of complex phenomena. The world can be perceived in complex presentation only through the gap, through the most diverse patterns of void or bias. Similarly, Badiou's concept of the event (in *Being and Event*)

is pertinent here. He asserts that a void represents the unseen, repressed forces. A void in the existing situation reveals the inherent instability and potential of the present moment in the flow of events, where truth emerges only by embracing the radical indeterminacy of the event.

Building participation in the empty middle enables the evaluation to transform blindness and bias into inputs for synthesis, rather than obstacles to be overcome. The decisive factor of the anti-postmodern approach is its ability to reject logocentrism in truth claims as well as within the void. The empty middle transcends the realist inclination towards fixed structures and absolute truths, as well as the constructivist tendency towards a nihilistic or essentialist understanding of the void.

Criterial assessment of the evaluation tools with a strong constructivist component proved a demanding task. Practitioners or consultants, rather than scholars design these tools in a manner that makes formal judgement of them slippery. Design-based approaches are generally not formally defined, as they are fluid and adaptive. Their logic is difficult to falsify, potentially undermining scientific rigour. Their legitimacy often depends on designers reaching consensus, '*rather than the universal accuracy or validity*' (Powell et al., para. #34). The paper acknowledges difficulties and accepts that researching constructed objects must be itself constructed. It limits itself to the four tools presented in selected literature sources, not necessarily the most recent or the most consolidated ones. However, the primary goal is not to describe tools but to pinpoint characteristic obstacles that underscore methodological challenges commonly found in many other established participatory evaluation approaches.

The four tools were presented and discussed at the Fifth Western Balkans Evaluators' Network (WBEN) conference in Ljubljana in late September 2023.[1] The Slovenian Evaluation Society hosted the event on behalf of WBEN.

II. Four Tools

The tools are first presented from the perspectives of their designers. The rationale behind each tool is described, along with their declared contribution to the inclusiveness and collective rationality of participatory evaluation.

Three steps are typically taken when designing such tools. Firstly, evaluators gather diverse narratives from multiple sources, participants and stakeholders. Secondly, they identify the themes, causal factors or patterns, commonalities and differences within the collected material. Thirdly, they facilitate a shared understanding of patterns through a participatory process of constructing collective narrative incorporating diverse stories, respecting their principal differences and evaluating the specifics (sector, theme, region …) of their domains.

II.1 Most Significant Change (MSC)

[1] Conference webpage. Accessed December 2024.

An Australian evaluation consultant, Rick Davies, developed this participatory tool for monitoring and evaluating complex development interventions. MSC provides a simple means of making sense of a large volume of complex information (Davies, Dart). It is best suited to large-scale, open-ended interventions, which would be difficult to evaluate using traditional methods; especially when it is impossible to say with certainty what the outcome will be or how they will vary across beneficiaries.[2]

The MSC is an approach to evaluating intermediate impact.[3] This method collects qualitative information and stories from the intended beneficiaries of an intervention, who are best positioned to explain the outcomes of the intervention. Information is obtained by asking beneficiaries to describe a change resulting from the program intervention they consider the most significant. Next, it seeks an explanation of why respondents think that change is the most significant. Then MSC externally verifies the collected stories to see if the described changes align with what is already known from neutral external sources.

One or more selection panels of beneficiaries follow the verification, each panel representing one intervention domain, to review the collected MSC stories. Participants engage in thorough discussions and then vote on the most significant specific change. At this stage, the evaluator acts as a facilitator who manages the debate, gives the floor to everyone and handles possible setbacks.

The MSC is inclusive since it encourages the active involvement of beneficiaries and stakeholders at different stages of the evaluation process. It ensures that a range of perspectives is considered in order to reduce bias in participatory evaluation while grounding the results in evidence.

The MSC also contributes to objectivity in participatory evaluation by focusing on the changes reported by the actual beneficiaries as the most competent agents.

II.2 Causal Mapping (CM)

A Causal Map is a conceptual framework introduced by the political scientist Robert Axelrod to analyse complex systems by eliciting the causal beliefs and assumptions in a system of relations. Powell et al. applied this theory to develop CM as an evaluation tool. They describe CM as a tool that sheds light on the black-box of qualitative analysis. CM draws logically consistent and inclusive conclusions from qualitative data. It is particularly effective with a

[2] Monitoring and Evaluation NEWS. Accessed December 2024.

[3] Davies asks to pay attention here (in correspondence on 9. June 2024): "*A change in the service being delivered to the respondent could be seen as most significant because of its experienced or anticipated effects, or for other reasons. Any sort of change likely or actually affecting the life of a respondent could be reported as significant. MSC only becomes an* '*intermediate impact*' *when it is located in a conceptual framework of one kind or not, probably those of an evaluator or program manager, but very unlikely in any explicit way by the respondent themselves, if they are beneficiaries for example*". However, also following Davies and Dart (p. 9): "*MSC focuses on … intermediate outcomes and impact*." Furthermore, every change is an intermediate feature between pre- and post-, or between old and new. The article accepts the warning, but the concept of intermediate impact does not need to be understood by the participants, it is a theoretical term relevant only to the evaluator.

graphical representation of causal networks, where arrows symbolise evidence or beliefs about causal influences (GCM) to interconnect factors as nodes or elements.[4]

The CM studies "*information about what people believe*" (Powell et al., p. 1) and organise it in the map. It demonstrates that the success or failure of an intervention is rarely due to a single factor but is instead shaped by a complex web of causal relationships. CM outlines the network of factors driving change through complex intervention, facilitating dialogue among stakeholders and supporting them in making informed decisions and actions. An illustrative example of a causal map in policy impact evaluation is the theory of change, a conceptual framework used by decision-makers and evaluators to articulate the logic underlying their understanding of a complex intervention (Powell et al.).

Powell et al. declare that the objective of CM is inclusive – designed to expound on stakeholder perspectives. CM captures the causal assertions conveyed through narratives rather than deducing them through statistical analysis as pre-structured inquiries like in conventional science. CM offers insights into the cognitive frameworks of stakeholders and represents their reasoning and behaviours (Powell). Participants are not involved merely as subjects of the evaluation or informants but are actively engaged in identifying the internal drivers of intervention-induced change. Mapping fosters causal, and therefore rational, reasoning in participatory evaluation.

The methodology of CM prescribes a rigorous procedure for eliciting causal statements provided by target beneficiaries. Constructing causal maps involves systematically collecting and meticulously analysing narrative descriptions of change. Evaluators encode and analyse these narratives to uncover participants' perspectives on causal relationships (GCM). The encoding classifies causal claims by whether they explicitly link outcomes to specific activities and compares them with the stakeholders' theory of change. This process generates meta-data about the narrative content. The analysis identifies a causal factor network, where elements exert direct and indirect influences on one another along the network's pathways (GCM). Semi-automated generation of summary tables and visualisations guides the interpretation of the evidence (Copestake et al.). Mapping organises diverse connections into a unified global causal map. Using filters and algorithms, the global map can be adjusted to zoom in or out to address different group inquiries (GCM), producing distinct sub-maps tailored to various groups or thematic concerns, helping to identify areas of agreement and disagreement.

A final step, the conclusive interpretation of the implications of these maps, already extends "*beyond the realm of causal mapping per se*" (Powell et al., p. 8). The evaluator's pivotal role in causal mapping is the careful collection and accurate visualisation of causal evidence while leaving it to stakeholders to conclude how to proceed from the revealed maps.

There are different approaches to the CM. Powell does not consider their version of the CM primarily as a participatory evaluation method, arguing that it is participatory only in data collection and is therefore one-sided. This raises the question of whether CM even belongs to

[4] Guide to Causal Mapping. Accessed December 2024.

the set of tools that use a participatory process to evaluate complex interventions. The paper included CM in the list since it does not focus on whether the participatory process in CM is complete, but on whether it involves participants in the evaluation as defined by Stufflebeam. Better Evaluation, one of the credible online sources of knowledge on evaluation, also links causal pathways evaluation with "*a particular focus on participatory processes…, amplifying their voices and narratives.*"[5] As the next two chapters show, there are important lessons to be learned from the experience of using participatory methods in CM regarding the evaluation of complex interventions.

II.3 SenseMaker (SM)

A consultancy firm, Cognitive Edge, developed SM as a narrative-based approach to evaluation. The SM starts from the observation that traditional quantitative evaluation methods cannot fully understand the complexity of communal issues, since it does not draw upon their shared collective wisdom. Voices That Count, a collaborative network of consultants and practitioners, developed a participatory narrative sensemaking (SM) approach.[6] Its mission is to make sense of collective issues by gathering and analysing narratives, stories, and contextual information. The tool is designed to identify similarities and divergences within the collected narratives, providing insights into how groups of people interpret and make sense of a given complex situation. This helps uncover patterns of opposition and overlap (shared understandings) between social groups, themes, or sectors.

SM's methodology prescribes that participants are first presented with carefully selected prompts designed to elicit narratives on specific areas of interest. They are then asked to plot their responses quantitatively, indicating their relative importance to predefined signifiers based on a previously narrated story.

First, participants plot narratives across three evaluative domains, then across 'dyads' (either-or questions on selected issues), and finally, they plot separate multiple-choice questions (Bartels et al.). Plotting enables evaluators to translate qualitative narratives into numerical data, allowing statistical analysis and visual presentation of aggregated results. Additionally, quantification aligns responses with specific semiotic markers (Van der Merwe et al.), such as keywords. Participants assign meaning to their narratives by self-signification, eliminating the need for the evaluator's interpretation of results and thus avoiding researcher bias in evaluation.

From quantification and visualisation, collective meanings emerge (Deprez). SM diverges from conventional tools that focus on the in-depth examination of individual narratives, as seen in MSC or CM. Instead, it emphasises the entire body of narratives, allowing the identification of patterns, deviations, group-specific differences, and positive correlations among distinct signifiers. SM helps evaluators pinpoint key questions, issues, and themes for further exploration in collective sensemaking workshops (Deprez). Workshops include participants with dissimilar views, such as project beneficiaries, program personnel,

[5] Bettter Evaluation, #Causal mapping, Accessed December 2024.

[6] Web page. Accessed December 2024.

managers, and donors (Guijt et al.). Through collective sensemaking, a shared cognitive map is created (Van der Merwe et al.), revealing the underlying order within apparent disorder.

SM is inclusive at both the micro level (gathering stories) and the macro level (sensemaking). It involves participants in data collection by capturing their stories and perspectives, empowering them as co-analysts in the sensemaking process. SM improves collective rationality by facilitating more analytically supported sensemaking in participatory evaluation from heterogeneous narratives.

II.4 Outcome Harvesting (OH)

The author of this participatory evaluation tool is Ricardo Wilson-Grau, a recently deceased but remembered independent evaluator from Brazil. OH recognises that intervention outcomes are complex and multifaceted, making them difficult to assess using traditional quantitative methods. Traditional methods typically assess effectiveness or efficiency but fail to capture the complex reality of the diverse agent changes involved in an intervention. Moreover, they often overlook human and social behaviours.[7]

OH is a particularly appreciated tool for evaluating interventions where outcomes are emergent and not easily predicted in advance, in uncertain environments with many unintended outcomes, including negative ones. It is also useful when the relationships between causes and effects are identified but not fully understood, or when it is impossible to define concretely most of what an intervention aims to achieve (Wilson-Grau). OH can assess interventions aimed at driving social change, or when an intervention has undergone significant modifications – in its design, external conditions, or implementation mechanisms – rendering it ineffective to gauge outcomes against the original plan. Other cases include evaluating community development programs, particularly those aimed at increasing social inclusion (Wilson-Grau).

As a formal procedure, OH invites stakeholders who are closely involved in the intervention to identify relevant behavioural changes, referred to as ‘outcomes’. They encompass actions, activities, or practices by key stakeholders that are necessary for benefiting people or the environment (Wilson-Grau). The tool focuses on the actors who change, rather than exclusively on the ultimate beneficiaries of those changes. The evaluators first collect (harvest) evidence of changes and then, working backwards, determine whether and how the intervention contributed to these changes – rather than working forwards by measuring progress toward predefined goals. The harvested information undergoes several participatory rounds of revision to ensure it is specific and comprehensive. The harvested data is then validated by cross-referencing it with information from independent, knowledgeable, and authoritative sources.

In the synthesis, the evaluator classifies all outcomes according to the objectives and strategies of the key stakeholders. He interprets the results in terms of their ability to address the harvesting questions, primarily ensuring that they are credible and provide a solid

[7] Peroni Fiscarelli A. 2022. Participation, a key focus in outcome harvesting. Accessed December 2024.

foundation for key stakeholders to shape answers to the evaluation questions. The evaluator proposes issues for further discussion among stakeholders on utilising the findings (Wilson-Grau).

OH is an inclusive approach that captures a wide range of perspectives from diverse stakeholders in both identifying outcomes and interpreting their collective implications. Its results contribute to collective rationality by focusing on real-world outcomes observed by involved stakeholders rather than solely relying on measuring predefined indicators. By shifting the interpretation of findings to stakeholders, the evaluator avoids bias.

***

The four tools developed purpose-specific strategies for addressing the selected challenges of collective choice in complex conditions. The tools contribute positively to social inclusion by involving diverse participants and stakeholders in the evaluation process. They also boost collective rationality by focusing participatory evaluation on significant communal issues, identifying real-world changes, providing a causal explanation of uncertain questions, or achieving a shared interpretation of quantified and visualised evidence. Consequently, they enable more robustly informed and inclusive democratic decision-making – at least according to their authors and advocates.

Assessing the four tools against the selected two criteria (Chapter III) will reveal a more intricate picture (Chapter IV). This paper demonstrates that the causal relationships between participation and inclusion, as well as between objective assessment and collective rationality, are not straightforward. The complex situation demands a far more sensitive approach to assessing neutrality of participatory evaluation tools.

III. Inclusiveness and Rationality

This chapter accomplishes two interconnected tasks that build the framework from which the four tools are later assessed. First, it outlines two imperatives of participatory evaluation. Then it identifies two specific assessment criteria that best recognise neutrality – the contributions of evaluation tools to social inclusion and collective rationality.

III.1 Social Inclusion

Assessing social inclusivity is a multifaceted task due to the opposing ways this concept can be approached: from the viewpoint of the minority or majority of the population and the perspective of the included or the excluded.

Mainstream concerns regarding social inclusion emphasise equal opportunities for collective action for all, particularly for deprived minority groups, like disabled, ethnic minority or deprived, whose access to social resources is discriminately constrained (Rawls, Ostrom). Likewise, Arrow's non-dictatorship principle safeguards collective choice against undue influence from privileged minority groups, such as politicians, project leaders, and key stakeholders. The mainstream discussion on social inclusion typically focuses on issues

concerning minority populations, whether excluded (deprived) or included (privileged) members.

A contrasting perspective posits that discourse on social inclusion should encompass the society's majority. Even in democratic cultures, structural and systemic barriers, elitism, corruption, and political opportunism (Laclau, Mouffe) marginalise the majority from shaping society (Bauman). Independent studies support this view: globally, 60% to 80% of people have felt inadequately represented by their governments already for decades (United Nations in Kreisler; Eurobarometer; Henning). People are usually included only as excluded (Wallerstein) when their opportunities to participate in social life and influence social change are not absent, but significantly curtailed and manipulated.

Social inclusion should be then approached from the perspective of the majority of the population. This also includes excluded minorities but they represent only the most acute aspects, not the entire spectrum of social exclusion. Foucault's exploration of power structures (in *Discipline and Punish*) revealed how the marginalised often pursued security within structures rather than questioning them. The minority does not usually strive to reduce systemic exclusion, but only achieve a state in which they are no longer excluded any more than the majority of the population. Often they merely reinforce systemic inequalities.

Similarly, social inclusion can be examined from the standpoint of those included or those excluded, and the findings will diverge again. The mainstream view is typically framed by the included, such as key institutions at national and international levels like The Word Bank[8] and European Union.[9] The socially included perceive the world through a self-referential, logocentric lens (Derrida), shaped by how they understand, organise, and interpret social processes. They take inclusion for granted. Exclusion is for them unwanted side effect of social ordering that can be remedied within the framework of the existing regulation of social relations. They particularly underestimate the pervasive influence of invisible drivers of discrimination embedded within the social structure (Crenshaw). Therefore, the description of social exclusion from the perspective of those included is necessarily inadequate (Badiou).

Postmodern philosophy (Foucault) evoked a contrasting standpoint, emphasising how the excluded perceive society differently. Social exclusion is inherent in any model, system, or theory of truth and any concept, classification, category, code, name, or identity. Derrida argued that exclusion precedes inclusion drawing from Plato and Heidegger's claims that what is not seen or said shapes the horizon of what is seen or said. In constructing meaning, the notion of exclusion finds an analogy in mapmaking: just as a city map omits details like every tree or building to provide a simplified overview, systems of thought similarly exclude ‘insignificant details’ to highlight what is deemed most important. However, the elements omitted from the map – the invisible aspects, the void – do not disappear. Huxley wrote in *Proper Studies* (p. 205) that “*facts do not cease to exist because they are ignored*.” Ignored elements continue to shape the city's life, albeit in repressed and invisible ways. A map would

[8] Web page. Accessed December 2024.

[9] Web page. Accessed December 2024.

be drawn quite differently from the perspective of the excluded, and it would be no less real. It might even address some of the city's attributes more realistically than the official map such as coincidence between city's grey areas and alternative urban communities with their prefigurative, future shaping urban practices.

Those experiencing social exclusion have distinct views of social phenomena (Spivak). As invisible, they can perceive the world only through the void (Radej, 2021b, 2022) imposed on them by ignorance. Viewing the world through the void is particularly relevant when deliberating the indeterminacies of contemporary communal life.

The notion of the void has intrigued philosophers for centuries because it touches on fundamental questions about the nature of reality. Ancient Greek philosopher Parmenides rejected the existence of the void, asserting that ‘what is not cannot be’. In contrast, Democritus argued for the necessity of the void to explain motion and change. In contemporary philosophy, the void is present because our insights are inherently biased in the parallax view of the world in its indeterminacy, due to complex processes in which contemporary societies are embedded.

Two extreme interpretations have been proposed for engaging with the void: nihilist and essentialist (Nagarjuna, Tsongkhapa; in Jinpa). A third, mesoscopic or middle-way explanation has arisen between them.

For postmodern nihilists, everything is nothing. They interpret the void as revealing the absence of stable meaning and the lack of essence at the core of everything, leading to a sense of meaninglessness. Heidegger's philosophy (in *Being and Time*) negated the explanation of the void as meaninglessness or absolute absence. His understanding of the void does not imply the absence of all things but rather the presence of everything invisible. This philosophy parallels quantum mechanics, where the void is not truly empty but full of quantum fluctuations and virtual particles interacting with one another. Sartre, similar to Heidegger emphasised that the void is an open field of possibilities and potential for freedom. It is a human-created absence, inseparable from human consciousness, a dynamic space that points to a relational field of unused potentiality and overlooked co-dependency between coexisting things. Humans enter the world without predetermined essence or nature. By choosing among available alternatives, they create meaning. But when unchecked and left to itself, where no narrative or value is considered intrinsically worthier than any other, the void leads to nihilism, a situation where “*anything goes*” (Feyerabend, in *Against Method*, 1993, p. 28), ultimately destroying a collective sense and replacing it with indifference. In evaluation, nihilist attitude arises when all interpretations of evaluation results are considered equally valid, regardless of their collective rationality.

On the other hand, for essentialists, nothing is everything. They understand the void as an unchanging essence or inherent quality, hanging above everything as something absolute from which everything arises and into which everything eventually dissolves. While it is not be dominant in ‘enlightened’ Western philosophy, this perspective is central in Yogācāra Buddhism. However, the dialectical philosophy of the middle way, found in both Nagarjuna's

Old School (Madhyamaka Buddhism)[10] and the contemporary Kyoto School dismantles the essentialist interpretation of the void. They argue that essentialism merely constructs another form of logocentrism. When overcoming the realist logos, essentialism places the logos of the void at the centre of everything. For instance, widely used constructivist tools often position the evaluator as an invisible guide or virtuous mentor to participants and stakeholders (see Section IV.1) in the evaluation of complex interventions.

Discarding two extreme approaches leaves only one option: elaborating on the void from a non-extreme vantage point, from the middle between polarities. Because it is grounded in the middle, the void cannot be authentically grasped from the extremes. The void in the middle is neither full of nothing nor empty of everything. It signifies the absence of something, a reality that remains unseen yet nonetheless felt (Heidegger). For Heidegger, as well as for Nagarjuna and Sartre, the void denotes the absence of something important. Heidegger refers to the experience of dread in darkness as an analogy. One does not see anything, yet dread arises precisely because of the awareness that certain things are present – somewhere out there, vaguely threatening, even though not revealing any danger in particular. Sartre posits the void as the lack of meaning or purpose, resembling the void left by losing a loved one, or when regaining awareness about things through their absence. Studying the history of the number zero, mathematician Robert Kaplan similarly reminds us that zero was not always defined as a number, as something present. Since it does not represent a quantity, zero was understood merely as a 'placeholder'[11] – a code or symbol signifying something important for the construction of meaning, yet demonstrably absent.

The void is not an absolute or universal entity, precluding the essentialist interpretation. The void does not point to ultimate reality (Sartre). This is justified by various philosophical and logical procedures that first place the void at the starting point of their explanation of things, and then immediately negate it. "*Nothing nothings itself*,"[12] wrote Heidegger, due to the self-nihilating nature of nothingness: it does not exist as a thing or entity with its own substance. Assertion of the void and its immediate negation is also employed in the dialectical method of negation of negation and in double negation in logic. All three approaches engage with the void through the lenses of contradiction, transformation, and the creation of new meaning. The first negation reveals a network of previously unseen relationships. The second negation necessitates a re-evaluation of the original position. This second negation returns us to the

[10] Yogācāra Buddhism teaches reality is shaped by the mind so it might be in a way paralleled with the Western Idealism. Madhyamaka Buddhism would be in certain respects comparable to Phenomenology or Sartrean existentialist philosophy. Both branches of Buddhism, of course, belong to much older philosophical traditions. While it is not possible to equate Eastern and Western philosophical traditions, similarities have nonetheless been recognized and studied for at least a century, most systematically with the emergence of the Kyoto School (Nishida, Nishitani). See the discussion in Krummel J.W.M. 2014. Anontology and the Issue of Being and Nothing in Nishida Kitarō, in JeeLoo Liu, D.L. Berger (eds.). Nothingness in Asian Philosophy, Ch 17. London, Routledge; and Krummel J.W.M. 2019. Nishitani Keiji: Nihilism, Buddhism, Anontology, in Gereon Kopf (ed.), The Dao Companion to Japanese Buddhist Philosophy. New York, Springer., pp. 649-79.

[11] For instance, number 205 would have been written as 2_5, which reads literrary as two Hundreds, no Tens, and five Ones. The symbol '_' denotes an empty space, serving as a placeholder for the absence of something relevant.

[12] "*Das nichts nichtet*", Heidegger, Gesamtausgabe 2, p. 117.

original position, reaffirmed with a deeper understanding derived from viewing it through the void. The second negation always serves distinct purposes of synthesis: existential (Heidegger), transformative (dialectics), and logical affirmation (double negation). By acknowledging the interconnection between different voids, evaluation fosters a more fundamental form of inclusivity. It moves away from an exclusionary approach to the void and instead embraces the void, particularly the overlap between voids, as essential for synthesis.

Our concern cannot be solely on the relationship between what we see and what we overlook (postmodernity). It must also assess the relationships among the overlooked things themselves (anti-postmodernity) – an approach to double negation in evaluation. Attention must focus on the internal coherence of what is absent and on the interdependencies within the invisible, as they generate characteristic patterns with emergent properties that shape understanding of a complex issue. Shadows, for instance, are not merely absences of light; they interact in complex ways, producing dynamic insights that transcend their relationship to visible objects. In analysing images generated by artificial intelligence, inconsistencies in shadowed parts of image expose their inauthenticity. This exposure arises not from the relationship of each shadow to its casting object but from the failure of shadows to maintain consistent relationships with one another in terms of direction, length, and intensity. Likewise, participatory evaluation needs to learn the language of shadows and listen the conversation between the shadows. An evaluation that adopts an empty middle as its foundation would explore the intersections of biases.

Bias of inclusion in intersection with bias of selection would, for instance, reveal the prejudiced inclusion of data in evaluation. Finding that the overlap between them is often considerable would probably not surprise many. Similarly, where confirmation bias intersects with selection bias, evaluation might discover that different groups are merely performing expected roles in evaluation rather than responding truthfully. Certain intervention outcomes are repeatedly reported not because they are most common, but because they are most aligned with official expectations. Likewise, consider the intersection between self-selection and social desirability bias. By noting who is not participating (self-selection) and by meta-analysing the language used by those who do participate (looking for overly positive language indicative of social desirability), evaluation can identify discrepancies between the stated success of the intervention and how it is assessed by less privileged groups in the community.

Traditional inclusionary efforts in participatory evaluation would amplify marginalized voices, addressing the surface-level visibility of the excluded. However, a deeper assessment requires examining how marginalized voices relate to one another – how different excluded groups interact, compete, or align in a meso-matrix of collaboration between the excluded ‘anti-system’ groups (Radej, 2021b). When different excluded groups interact, they create shadow networks, such as food systems involving unofficial exchange mechanisms and local currencies. The intersecting practices of the excluded do not serve as alternatives to the formal system; rather, they form a complex anti-system with their own internal value structures, logic, rules, and mechanisms. They foster new forms of organization, including invisible (infra)structures and invisible governance, which can give rise to novel forms of political

actions. Official society remains unable to recognize these dynamics unless it observes through blindsight – acknowledging the coherence of the unseen and the positive and negative synergies, overlooked by realists and constructivists alike.[13]

The void in participatory evaluation of complex interventions is not homogeneous category but exists as a multiplicity in various contexts. For example, economists, focusing on market dynamics, tend to overlook ecological constraints; ecologists, emphasising environmental limits, often ignore the economic impacts of environmental protection interventions. These two pieces of void belong to a broader void that is not a totality (Badiou), the void of all voids. The void is always incomplete because not all its instances can ever be fully encompassed (an example follows in Section IV.1). In anti-postmodern philosophy as much as in quantum physics, the void is a relative entity. Physics imbues the void with different characteristics and potentials, which, if shifted, "*create a different definition of empty space*."[14] When quantum physics speaks of the void, it does not mean a unitary, unchanging, and absolute void.

Then the truth and the void can be equated in a decisive respect: both are relative. Their relativity enables a non-logocentric critique of logocentrism: as ordered but not oppressive, as rational but not exclusive, acknowledging radical diversity but not enforcing nihilism...

Mainstream understanding of the relativity concept falls into two already discussed extremes. On the macro level, general relativity (e.g., Einstein, Structuralism) provides a model for understanding large-scale ideological structures, like liberal or conservative democracy, which shape political choices similar to how mass bends space-time. This is a universalist framework of relativity. It is critiqued by complexity theory arguing that macroscopic relativity is logocentric and too rigid to account for the fluid interaction between order and disorder (Prigogine in *The End of Certainty*).

On the other hand, the microscopic concept of relativity is subjectivist, relating to everything ephemeral and particular, as seen in postmodernist or constructivist thought. It focuses on how individuals shape perspectives of reality. Habermas criticises postmodern relativism in The Philosophical Discourse of Modernity, arguing that it overemphasises subjective constructs of reality. Particulars never exist as isolated entities outside of foundational concepts, for instance, a shared conception of the good rooted in communal life. Microscopic relativity is responsible for fragmenting broader structural or historical understandings of social processes undermining universal norms necessary for social cohesion.

Relativity operates at all three levels of analysis (Luhmann in *Social Systems*): micro and macro levels are extremes, while meso-level relativity occurs between them. A non-extreme concept posits that the truth about the relativity of truth is itself relative. Things cannot be relative in an absolute or extreme way, such as absolute relativity, because of the inherent contradiction in claiming absoluteness for relativity. At the meso level, the character of

[13] Radej B. Sounds of Silence in Evaluation. Accessed December 2024.

[14] Hippel von M. 2024. Vacuum of Space to Decay Sooner Than Expected, Quanta Magazine. 22. July 2024. Accessed December 2024.

relativity is reflexive (the application of relativity to itself; Derrida) and intermediate (Luhmann).

Thinking about (complex) things through the void eludes the comprehension of an enlightened scientist and requires instead a blindsighted evaluator who understands things as indeterminate without ever falling into extremes of totalisation or infinite regress. The blindsighted evade two key pitfalls in evaluation: neither overlooking unimportant things to participants nor being blinded by what they consider true. He acknowledges that every piece of knowledge comes with a degree of uncertainty and that the truth lies somewhere in between many biased claims. The blindsighted perspective transcends mere factual judgment while maintaining respect for scientific arguments. Facts do not lose validity in uncertain conditions, something else becomes problematic, their points of reference and meaning become fluid and susceptible to dissimilar interpretations. Navigating uncertain situations is akin to travelling through a foggy landscape, where the limited visible information can be useful or misleading, depending on how observations are interpreted. Data obscure as much as they reveal (Žižek). This bears an analogy to a participatory evaluation where data previously seen as merely descriptive are reinterpreted structurally which can change the previous understanding of the evaluated object.

The blindsighted recognise the value of contrasting viewpoints without deciding one or the other. He remains epistemically indeterminate (Wittgenstein in *Philosophical Investigation*), non-judgmental and inconclusive. This does not imply indecision, only indeterminacy. It provides a non-logocentric synthesis by uncovering unused opportunities for synergy and inclusion without prescribing a specific course of action. Whilst it does not resolve the collective choice dilemma, blindsighted evaluation puts forward a more inclusive, non-logocentric but nevertheless synthesising deliberation about collective choice alternatives in complex conditions. Evaluation, starting from the standpoint of the empty middle, is inclusive because it approaches the dilemmas of collective choice in terms of what is excluded. It does not guarantee a stronger majority, but it addresses dilemmas of collective choice as indeterminate.

While participation in evaluation promotes openness to diverse perspectives, it does not ensure genuine inclusivity (Chambers). Merely inviting people to participate does not warrant their voices will be heard, appreciated, and meaningfully integrated. The relationship between involvement and inclusivity is far more intricate, depending on how successfully the tools neutralise bias in participatory evaluations. Complete bias mitigation remains elusive. The most inclusive evaluative tool is one that transforms the epistemic blindness of participants, stakeholders, and evaluators at the outset of the process into epistemic blindsightedness at its conclusion.

III.2 Collective Rationality

The second imperative of neutrality in participatory evaluation demands advancing collective rationality. The collective rationality of tools depends on their aggregative capacity (Scriven, 2003), ability to identify the most representative expression of the collective will (output) from given participatory contributions (inputs). The aggregation problem in participatory

evaluation arises from the multifaceted nature of collective choices involving rational and legitimate yet opposing alternatives.

Presently dominant aggregation methods in evaluation are restrictive. They produce suboptimal aggregates that validate inferior policy decisions, which disadvantage the collective good and constrain collective rationality. It erodes trust, breeding community disengagement and reluctance to cooperate in communal efforts. Selecting the appropriate aggregation procedure is then a decisive task in participatory evaluation.

This section assesses the collective rationality of some of the common aggregation approaches. Various approaches to constructing wholes exist, such as the hierarchical structure of the MSC and OH, the horizontal network pattern of the CM, or a combined vertical-horizontal structure in the SM. Each construction is appropriate for different conditions, modern, postmodern, or complex (anti-postmodern). Scriven (2003) emphasised that the aggregation logic must always align with the pre-existing conceptualisation of the whole, the evaluation object, the evaluation scope, and the underlying theory of change (Creswell). The choice of aggregation method in the evaluation must always submit its logic to its preformatted dispositions!

Evaluation of simple policy or program interventions requires a simple approach to aggregation. It directly aggregates from the individual to the community level, from micro to macro. This method is descriptive and quantitative; it cumulates precise, definite, linearly additive, and commensurable inputs. Its scope is accumulating similar elements, averaging, or identifying the frequency of specific values or data points within a dataset. Conversely, evaluating complex interventions necessitates complex methods of synthesis that connect micro and macro through the meso level. They involve incommensurable and non-linear inputs. The goal is to identify characteristic patterns, contradictions, or novel meanings hidden in the aggregate. Though aggregation methods are complex, they need not be computationally demanding, quite the opposite. Complex approaches can remain procedurally easy if addressed evaluatively as synthesising mesoscopic phenomena.

Evaluation theory has historically demonstrated a perplexing relationship with the aggregation. A central point of contention is the absence of a unified perspective on the role of aggregation in the evaluation process. This lack has instigated deep divisions in the field, splitting it into antagonist traditions – non-aggregative and aggregative, each holding distinct philosophies and methodologies. This divide represents more than a methodological choice – it established itself as a question of epistemological commitment.

Non-aggregative proponents claim that aggregation obscures diversity involved in input data, reducing rich, contextual information to a homogenised form that cannot represent them. Proponents of non-aggregative approaches emphasise the importance of understanding diverse contexts at the elementary level between individual particulars and prioritising deep contextual analysis often absent from aggregation.

A prominent advocate of the non-aggregative approach is Luna Leopold. Leopold et al. developed a detailed matrix method for impact assessment at the micro level, revealing how various policy actions impact specified assessment criteria. They asserted that aggregating

fragmented findings presented in a matrix into policy-relevant conclusions at the macro level – identifying the aggregate impact of all actions on a given criterion or all actions on all criteria – requires value judgments, which neutral evaluators should always avoid. According to Leopold, the evaluator's primary task is to inform and comment on specifics rather than generalise. He argues that refusing to aggregate findings erects a safety boundary between the evaluator and policy-maker, protecting the former from political interference.

Since Leopold, many generations of evaluators have refused aggregative thinking. Patton (2010) claims that in complex conditions "*Evaluation informs continuous adjustment and navigational reorientation as systemic realities unfold…, but no summative synthesis is ever possible and, if done, is meaningless. It is instantly merely a historical account of what has been but is already outdated when produced because the complex system dynamics have changed. In developmental evaluation, there can be no meaningful summative synthesis*."[15] The Impact Assessment Board (IAB),[16] advisory to the European Commission, found that most evaluation studies did not provide policymakers with useful information at the macro level. Hageboeck et al. observed the same in their meta-evaluation of 340 evaluation reports prepared for the United States Agency for International Development. Huitema et al., who assessed the synthesis quality in 259 evaluation studies commissioned for the European Union's climate policy, drew a similar conclusion.

Detailed impact assessments certainly aid in grasping policy interventions at lower levels of the institutional or program hierarchy. However, fragmental findings cannot assist decision-making at the highest levels, such as key stakeholders who sought policy advice. If evaluation findings remain disaggregated, they lead to information overload and fail to capture the complex reality within which policymakers operate, providing only trivial answers to complex policy questions (Virtanen, Uusikylä). Scriven (1994, p. 378) accordingly noted that rejecting synthesis in evaluation is akin to "*letting the client down at exactly the moment they need you most*." It is precisely through synthesis that evaluation fully realises its purpose providing judgments that are both informed and fair, especially in complex conditions.

The rejection of synthesis and shifting this task to policymakers, programme managers, or other participants assumes they can perform the task neutrally and in line with the complexity of the challenge. This assumption is difficult to justify (Stiglitz et al.) due to participants' epistemic blindness. If they were truly neutral, the evaluation would be unnecessary. Scriven (1994) warned that delegating aggregation to non-evaluators exposes evaluation results to manipulation by key stakeholders.

The second tradition in the evaluation field is aggregative. Despite acknowledging aggregation as a requisite in evaluation, these methods do not resolve the aggregation problem; instead, they merely restate it from the opposite extreme by employing approaches that seriously impede collective rationality.

[15] Michael Quinn Patton, in e-mail correspondence on 6. June 2024.

[16] Reorganised in the Regulatory Scrutiny Board in 2015. Accessed December 2024.

One of the most customary methods aggregates commensurable assessment findings into a composite indicator of the intervention's overall success. Input data are linked to the aggregate via a common metric, such as number of votes. Aggregation is reduced to mechanically combining already-aligned elements, simply relating like with like while excluding everything else. When using the common denominator in aggregation, the primary goal is simplicity of calculation rather than enhancing collective outcomes. Commensurable aggregation does not generate new meaning at the collective level so it cannot significantly contribute to the collective intelligence.

The dissimilar constructivist tools use different approaches to evaluation and synthesis as illustrated in Chapter II. They nevertheless share several key features. Most notably the rootedness of their designs at the meso-level. The four tools all share a lineage within the meso-level tradition. The authors of the CM and OH apply meso-level categories, such as participant-related structures, routines, or sub-maps. MSC evaluates intermediate impacts situated between policy, program, or project actions and their outcomes. Unlike traditional outcome-driven approaches, the OH, MSC, and CM do not identify the ultimate impacts of interventions. Instead, they focus on changes during the intervention process, such as behavioural shifts, process drivers, emerging patterns, or contextual influences. The SM aligns with meso-level tradition by sub-aggregating assessed data in a triadic structure, involving three distinct context-dependent evaluation domains.

Various aggregation approaches are available in meso-level evaluation, each with distinct collective rationalities and approaches to synthesis. These include network synthesis, causal models, meta-synthesis, mixed-methods research, mid-range theories, and multicriteria methods. Mid-range methods, for instance, achieve aggregation only from the micro to meso level – clusters, sectors or themes – while explicitly rejecting macro-level aggregation. Alternatively, multicriteria methods synthesise their results in multiple parallel results.

Meso level concept integrates multi-, mid-, and mixed-logic.

Many meso-level approaches are not authentically mesoscopic. While they integrate diverse data sources, their frameworks often operate across micro or macro scales, lacking meso-level reasoning. Multicriteria methods and mid-range approaches for instance typically build on what philosopher Charles Sanders Peirce, the founder of semiotics (theory of sign), referred to as ‘absolute thirdness’ between three independent and opposing participatory evaluation domains (A, B, C). In contrast, authentic meso-level evaluation is grounded in his concept of ‘relative thirdness’ (A, *ab*, B). This relational concept starts by acknowledging two opposing poles, like Option A versus Option B, and then introduces a third, hybrid category (*ab*) that draws from both as an overlap that bridges the gap between them. For example, socio-economic development (*ab*) signifies the intersection or overlap between the economic (A) and social domain (B) of sustainable development (analogous to its socio-environmental, *ac*, and economic-environmental, *bc*, overlap; Radej, 2021a). The relative organisation of triadic arrangement emphasises interconnection – when entities of a complex system co-constitute each other through interactions. Deleuze and Guattari argued in *A Thousand Plateaus* for the

fluidity of relations over fixed (structural) entities, suggesting that ab signifies the 'becoming' or emergence of something new, or transformation occurring between A and B.

Relative thirdness bridges contradictory claims intersectionally, akin to overlapping segments in a Venn diagram (Radej, 2021a). Relative thirdness is an original synthesising concept. It enables the building of collective structures from collective contents, in the meta-overlap between overlapped domains such as ab, ac, and bc in a Venn diagram of sustainable development. This contrasts with individualist constructions of the collective, as in Arrow's model, and generally in macroeconomics, where aggregates arise from summing individual preferences at a micro level. The distinction between individualist and collectivist constructions of wholes is crucial, as the latter encompasses a much broader and deeper scope of collective rationality.

Relative thirdness as a concept is related to Peirce's three universal categories for understanding phenomena: firstness (independent quality), secondness (opposition), and thirdness (mediation with synthesis between polarities). Firstness represents absolute qualities existing without reference to anything else, governing areas such as counting, classification, logic, mathematics, and religion. Secondness is about discovering the truth, relative forces, interactions, and relationships between different qualities, overseeing division, causality, dialectics, and correlation. Thirdness mediates between firstness and secondness evaluatively, as blindsighted, from the middle-ground.

When working with a triadic framework, Peirce developed the 'the secondness of thirdness' concept, an advanced form of relative thirdness, which formally frames synthesis from the meso (group) to the meta-level (collective). It bridges from two-part to three-part reasoning. The bridging occurs between many pairs of opposite domains organised in a square meso-matrix. A matrix includes at least three evaluation domains (e.g., economic, social, and environmental domains of sustainable development), structured by three rows representing complex policy, program or project intervention domains in intersection with three columns representing impact assessment criteria (economic, social, and environmental). Secondness addresses dyadic relationships between pairs of domains ('Meso 2 sublevel'; Radej, 2021a), while thirdness interprets the meta-overlapping between these correlated domains, aligning with the concept of the 'Meso 3 sublevel' that shapes the authentic collective perspective of a given complex matter.

The secondness of thirdness frames trialectics – a triaxial model of dialectics. It can be graphically illustrated using a Venn diagram, which depicts three (or more) partially overlapped circles (evaluation domains) – first between two and then between all three (or more) domains (collective, meta-level; see also Radej, 2010, Figures 4 and 5 illustrate formal development of a meta-level concept from meso-level).

The purpose of synthesis in the secondness of thirdness is not to construct macro-level aggregates, as this would contradict the mesoscopic nature of the challenge. The difference between macro and meta-level explanations of collective issues is crucial in the participatory evaluation of complex interventions. The macro perspective is logocentric and structural, aiming for synthesis through unification and hierarchy. It explains collective structures and

processes from a bird's-eye view. In contrast, the meta-level perspective bridges divergent macro explanations, presenting collective issues in multiple colours rather than black and white. Changing from macro to meta explanation modifies or even inverts the understanding of an evaluated object. For example, the phrase 'break a leg' means 'good luck' in the theatre; it should be understood at a meta-level, beyond the literal meaning of the words (macro-level). The meta-level combines the literal sense with tone, context, or cultural factors that invert the original meaning. Things are expressed at the meta-level through irony (Vermeulen, van den Akker), metaphors, paradoxes, or overlaps, evaluatively…; claims are not untrue, situations not unchangeable, messages not incomprehensible, domains not incompatible…

The collective rationality of participatory evaluation is therefore most appropriate when describing complex social phenomena from a meta-level perspective. This is the highest level of synthesis that due to its authentic mesoscopic foundation avoids trade-offs between inclusivity and collective rationality.

IV. Obstacles and Gaps

This chapter intersects the discussions from the preceding two chapters.

While each of the four tools effectively addresses specific evaluation questions, they all share a critical flaw from the same methodological obstacle: the inauthentic application of the meso level in participatory evaluation. This occurs when accomplishing meso-level-based evaluation with methods typical of the macro level (MSC, OH); when evaluation refuses the synthesis of its findings from groups to the collective level (CM, SM), or when middle-ground logic is pursued as if it was logocentric – when it captures things in an in-between state but not off-centre (Chödrön).

The following two sections explore examples of common obstacles to achieving neutral and authentic participatory evaluation, which are also found in many other currently dominant evaluation approaches. Sections conclude by examining new ways to resolve these obstacles through the empty middle.

IV.1 Epistemic Blindness

The four tools employ multiple techniques to mitigate biases in participatory evaluation and constrain implications of epistemic blindness in evaluation of complex interventions. For instance, SM, CM, and OH work to reduce confirmation bias – the cognitive tendency to interpret results in a way that reinforces existing beliefs and biases. Similarly, selection bias involves the prejudiced inclusion of data. When participants provide responses they believe are socially acceptable rather than honest, social desirability bias emerges. Davies and Dart observed that within the MSC, only 5-10% of narrative ratings typically resulted in a critical assessment of intervention outcomes.

While efforts to mitigate bias in evaluation are commendable, they can also fall victim to epistemic blindness. For example, CM tackles confirmation bias using robustness checks, including cross-validation of biased participatory contributions (Copestake et al.) with

evidence-based arguments and expert knowledge. The problem is that these checks undermine the democratic function of a participatory process. The primary goal of democratic processes is not to compete with or even supersede empirically derived knowledge, but to legitimise rational arguments. Similarly, evidence-based knowledge prioritises truth-seeking, with its role being to inform participatory processes ex-ante, rather than validate their outcomes ex-post. If the verification process is inappropriate, its outcomes – bias mitigation – will necessarily be itself biased.

A further example addresses the mitigation of social desirability bias. Davies and Dart propose creating a dedicated section in the evaluation report to collect negative narratives. Instead of presenting them in dynamic dialectical opposition to success stories, they merely place negative stories in an appendix, isolating them rather than allowing a productive tension between contrasting accounts to emerge. For evaluation to be free from bias, it requires, following Patton (2012), an integrated view where successes and failures inform each other.

The following case is also typical. To mitigate confirmation bias, the CM employs blindfolding techniques, with double-blinding. Recognised as the gold standard in experimental research (Duflo et al.), double-blind evaluation demands that both the interviewer and respondent are unaware of key information, such as the evaluation object or scope, or the definition of participant groups. This method helps "*avoid overly narrow agenda-setting, asking prompting or leading questions, … and explicitly or implicitly encouraging respondents to emphasise specific causal factors*" (Copestake et al., p. 8).

Double-blinding is an effective technique for eliminating specific biases in specialised evaluation studies, particularly those requiring binary answers about causality e.g., the effectiveness of new drugs in curing illness. Yet, it presents significant difficulties when evaluating complex interventions, where clear-cut cause-effect relationships are uncommon. Double-blinding also raises concerns about the inclusivity of participatory evaluation. Copestake et al. (p. 34) acknowledge that the approach requires a "*difficult trade-off*." While "*blindfolding may increase the credibility of respondents' voices from the perspective of the … primary audience, this must be offset against the potentially disempowering effect of not immediately revealing to respondents everything that could be revealed about the intervention being evaluated.*" Accepting trade-offs at the expense of inclusivity is seriously problematic for tools claiming to promote social inclusion.

In some other cases, constructivist evaluations tackle bias using a 'working backwards' technique based on reverse engineering. In high uncertainty, 'working forwards', a linear progression from premise to conclusion, as used in scientific inquiry, becomes unfeasible. The working backwards technique inverts conventional logic. First evaluators identify what has been achieved, what has emerged, or which view prevails among participants. They then work backwards, employing retrospective analysis to explain or justify these achievements or views. The MSC begins by collecting qualitative data in the form of narratives from beneficiaries, who describe the change they view as most significant. The evaluator then explores the reasons beneficiaries consider this change most important. Similarly, the OH first

identifies relevant outcomes, followed by a backward analysis to determine how, and to what extent, the intervention contributed to those outcomes.

While working backwards can mitigate bias in evaluations with well-defined goals and established theories of change (Bibri), it presents significant challenges when applied in complex conditions. The main disadvantage is its inability to provide a deeper level of analysis,[17] as it asks participants to engage with an idealised conception of reality. It assumes that individuals have perfect knowledge, enabling them to describe complex problems logically – as has already been established, this assumption is untenable from a collective choice perspective. Despite reasoning backwards (from effect to cause), draw forward conclusions, asserting that a caused b. They particularly neglect factors beyond their control and overlook alternative, unexpected, or creative pathways for change (Kahneman).

One of the most drastically biased methods of alleviating bias in participatory evaluations involves imposing ignorance on the evaluator. This method follows from the assumption that existing knowledge and beliefs act as blinders, obstructing alternative perspectives. To overcome this bias, the evaluator adopts a stance of unknowability and invisibility (Guijt et al.) to engage in evaluation in a restrained role solely as a facilitator, coach, or mentor, while leaving the evaluation task to participants and stakeholders.

Some authors suggest that inaction and lack of opinion enhance ethical behaviour and relieve evaluators of professional and societal responsibilities. Yet, an evaluator who achieves neutrality by remaining ignorant about the bias in collective choice commits a fundamental error, the logical fallacy of false neutrality. Insisting on moral neutrality when a decision is needed is a form of cowardice, wrote the medieval poet Dante Alighieri in *Divina Commedia*. He placed those who remained neutral during moral crises outside the gates of hell, in the *ante-inferno*,[18] rendering them unworthy of both heaven and hell, left in a state of perpetual, aimless wandering, doomed for eternity to chase a blank banner, symbolising their futile and directionless existence, a life without meaning or purpose.

The demand for the evaluator's invisibility in synthesis also contradicts the core evaluation principles, necessitating a solid understanding of the subject matter. Patton (2002) aptly asks evaluators to strive for epistemic modesty when forming conclusions in complex conditions. This entails acknowledging biases and limitations, although not allowing them to impede the pursuit of shared understanding and inclusive sensemaking.

Bias mitigation methods that employ blinding techniques cannot eliminate bias (Pierson et al.). They merely repeat it in modified form. The assumption that ‘good bias’ can nullify ‘bad bias’ is also mistaken. Blinding methods obscure bias more effectively, layering a black-box of qualitative evaluation within a black-box of bias mitigation techniques. This creates the illusion of impartiality in evaluation, as though social exclusion is no longer a concern.

[17] Babich N. Working backward: Why it is a must-have approach for product designers. Accessed December 2024.

[18] The term ‘*ante-inferno*’ does not appear in *Divine Comedy* itself but emerged through later literary interpretations.

Besides, each mitigation technique addresses only specific types of bias while leaving many others unaddressed.

Furthermore, evaluator must consider that participatory processes themselves are inherently exclusionary, since they are not carried out in a neutral domain, unbounded by the institutionalised forces of social domination. Social inclusion can never be pursued in participatory processes from a neutral space using neutral instruments (Mouffe). Participatory processes cannot be designed as islands of neutrality in the see of injustice and power imbalances. Participatory processes exclude individuals or groups prioritising autonomy over collective concerns, such as when aiming for self-determination, or egoism – both legitimate aspirations in complex situations. Besides, participation is not truly inclusive when involving uninformed participants, if they cannot influence the collective outcome. Or when the excluded can participate using only the mental frame and language of those who have excluded them. When the excluded are denied intrinsic and autochthonous epistemic agency, they must also give up the freedom to express themselves authentically through the voids and gaps of the dominant logocentric realm (Tanaka-Ishii). In all these situations, participatory evaluation invokes 'double exclusion' (Radej, 2021b), where legitimate concerns are ignored first in conventional (representative) democratic processes, and again in direct participatory processes.

When confronted with epistemic blindness, mainstream approaches to evaluation many times respond inauthentically with polarised methods. The first extreme stems from traditional rationalist result-based evaluation, which aims to abolish, neutralise, or at least minimise epistemic blindness. Bias is considered an error that distorts objective truth (Wilson-Grau).

The second and third extreme approaches align with the constructivist framework. The second approach draws on the nihilist concept of the void, suggesting that diverse policy or program outcomes – as different qualities – hold equal importance, with none being the most significant or valuable. Tools that follow this path risk steering the participation process towards uncritical acceptance of any democratically agreed-upon synthesis, regardless of what would be best for the community. An example is a facilitated synthesis (next section).

The third extreme approach to mitigating epistemic blindness aligns with the essentialist philosophy of the void such as when evaluators rise above the inherent contradictions in collective sensemaking. By establishing themselves as invisible guides or virtuous mentors, they imply a superior position, benevolently overseeing the middle world of incompleteness from above. Evaluators detach themselves from social oppositions, mirroring the portrayal of god or king in medieval frescoes, elevated above societal castes while mediating between them (Vignjević). They perform meso-level evaluation from the macro level, which immediately undermines its authenticity. A similar extreme example is the idea of the evaluator as a virtuous agent. Wilson-Grau outlined 'basic pointers' for the interpretive synthesis of findings asking the evaluator to avoid undue focus on quantitative data, discerning facts from opinions, and ensuring comparisons are made "*carefully and appropriately*" (p. 119). He seems to consider the evaluator the virtuous agent elevated above contradictions of collective choice. This idea may stem from the post-liberal communitarian

notion of virtuous politics – a mesocentric participatory democracy guided from the macro-level by the most virtuous agents (Pabst).

A different approach is required to address biases in participatory evaluation, moving away from extreme framings. The only viable alternative involves understanding epistemic blindness in a non-extreme, mesoscopic manner. This approach neither seeks to replace bias with truth nor obscures truth due to the inevitable presence of bias. Bias is neither entirely devoid of truth nor completely void, but rather filled with the invisible and overlooked aspects of a complex world, which can only be perceived from the empty middle.

The mesoscopic standpoint assumes both the indeterminacy of truth and the indeterminacy of blindness. The latter is constructed through a process of double negation (Section III.1): initially, the void is identified within truth, only to be nullified in a subsequent negation. The truth about epistemic blindness itself remains elusive, perpetually subject to its own contradictions, thus becoming self-negating or self-nihilating. Epistemic blindness deconstructs logocentric assertions but carries no intrinsic (essentialist, nihilist) meaning in the evaluation of complex issues.

The second negation unfolds through the exposure of the void against itself in the meso-matrix (or in a Venn diagram; Radej, 2021a). Rather than aiming for neutrality through a double-blind approach, participatory evaluation employs a double exposure of blindness, positioning biases in opposition to one another.

Schrödinger argued that, in complex conditions, the overlapping region shared by different scientific perspectives constitutes a truer representation of the world than those derived from any singular viewpoint. A grain of shared meaning arises in the overlap of many partially valid (biased) statements. This paper extends that view, claiming that overlapped presentations are truer because they absorb more epistemic indeterminacy. The deeper the overlap – such as in the double exposure of many double exposures (a Venn diagram of Venn diagrams) – the more the indeterminacy is integrated into evaluative reasoning. The empty middle is constructed through self-reinforcing logic, spiralling comprehension downward into an increasingly indeterminate understanding of complex issues, previously conceived as blindsighted evaluation. This process resists finality, never reaching an absolute void but expanding our awareness of uncertainty.

Double negation enhances the neutrality of the evaluation process, thereby uncovering more credible truths about indeterminate phenomena. In mesoscopic framing, bias is no more a cognitive flaw to be corrected but a generative space of indeterminacy – a valuable epistemic resource containing hidden knowledge and unexplored perspectives. From this generative space, blindsighted understanding of indeterminacy emerges.

The overlap of voids not only reveals deeper truths but also exposes deeper voids. The void is relative, first to other manifestations of the void and, more critically, to the void of voids. There is no singular void; rather, there are as many voids as there are explanations of complex phenomena. For example, in the context of sustainable development, the economic biases of ecologists and the ecological biases of economists, when intersected, reveal a further void – a void of voids – that neither perspective alone can address, the social dimension of

sustainability. It remains absent in economic-ecological relations, highlighting the cyclical incompleteness of such assessments (similarly, the economic void of void is absent in socio-ecological relations, and analogously for the ecological void of void). The concept of the void of voids underscores that no assessment can fully encompass everything absent, overlooked, or indeterminate, as there is always a void beyond void and no absolute void.

In simpler terms, the empty middle consists of relative voids and the void of voids. Framed by the empty middle, participatory evaluation thrives in this uncertainty, 'sailing the void' of complexity (Radej, 2021b).

By treating biases as epistemic resources, mesoscopic approach enhances participatory evaluation's transformative potential. The empty middle creates a shared epistemic space, rendering the contributions of excluded groups visible and opening pathways for their contribution in evaluations without requiring conformity. In this way, authentic participatory evaluation redistributes epistemic power from the included minority to the excluded majority, inherently fostering social inclusion. This has profound implications for the participatory evaluation of complex policies. Instead of trying to overcome biases, evaluators should systematically map what different biases consistently overlook, researching the patterns in these oversights with bias intersection analysis as a formal evaluation tool.

The four tools lack authentic inclusivity due to their failure to produce their results in the empty middle with blindsighted participatory evaluation. Although the tools declare themselves inclusive, their inclusivity is often an illusion. They shield the evaluation from epistemic blindness by attempting to suppress, outsmart, or submit to it, but never engage dialectically with it from the empty middle. Often, unintentionally and unknowingly, they themselves actively contribute to renewing the existing pattern of social exclusion. Unless the tools overcome their largely self-imposed obstacles, they will add to the status quo and reify the existing patterns of social exclusion.

IV.2 Aggregation Problem

As long as synthesis is inauthentic or even absent from evaluation ('*per se*'; Powell) and left to be accomplished by non-evaluators, its results remain narrow in scope, leading to impoverished collective rationality. This section finds the four tools are aggregative in a restrictive, inauthentic manner that is not categorically more collectively rational than traditional result-based evaluation. Often they are aggregative only in the first phase, from the individual to the group level (from micro to meso-level), like CM and SM: their commitment to aggregation diminishes significantly from the inter-group to the community level, aggregating from meso to macro. MSC and OH perform the synthesis directly from the micro to macro level without applying meso-level synthesis, even though they are declared mesoscopic tools. Inauthenticity emerges also when OH narrows the scope of synthesis on concerns of the included minority (key stakeholders). Others, such as MSC and CM, achieve this indirectly by choosing evaluation techniques rooted in the mindset of the included minority.

Several examples demonstrate how various aggregation approaches diminish collective rationality in participatory evaluation. Take for instance vote counting, by grouping similar preferences. The CM summarises collected narratives by counting the respondents who independently identify distinct causal connections (Powell et al.). The causal links mentioned with the highest frequency are considered most influential in explaining causality. The MSC also employs majority voting to differentiate more significant from less significant narratives. The SM adheres to the 'One Person – One Voice' principle. It assigns equal weight to each participant's voice (Hilhorst, Guijt), ensuring that dominant voices do not silence others. Deprez (p. 2) explains, "*Every voice, every story is used equally, it gives equal voice to those people who are often not heard*." The underlying assumption is that by valuing each individual's contribution equally, the resulting policies will more accurately reflect the broader needs of society and will be more collectively rational.

From the aspect of enhancing collective rationality, the demand for equal weighting of voices is misleading. Equal weights would be appropriate if contributions of individuals were collectively valued equally – in such a case, the problem of collective choice would dissolve. Equal weighting of voices is a foundational principle in political choices between ideological alternatives, not collective choice. As an instrument of collective rationality, participatory evaluation should enhance collective wisdom: after ensuring inclusivity for all, it should strive to assign higher weights to collectively more rational contributions (Sen, 2011).[19]

Furthermore, by framing collective choices as binary (Yes/No) among antagonist options (A vs. B), voting fosters polarisation rather than inclusivity, emphasising individualism over collective values. Even when a community deliberates on more than two options but the choice mechanism remains binary, voting secures the relative (included) majority[20] to prevail. The same results from the opposite technique, quantitative plotting in SM is accomplished on the interval between complete acceptance and rejection of predefined options with an infinite number of intermediate options. This approach fragments principal binary opposition to an infinite number of equally irrelevant choices, where 'anything goes', where the best rational choice offers no advantage over alternatives and, in particular, is no better than the status quo.

Selective synthesis is a further example of a constructivist methodology with pretentiously narrowed collective rationality. Any aggregation requiring a common denominator – meaning it only aggregates similar items, excluding everything else, is selective. A selective method may be well-suited for identifying the most important outcomes or extracting the most significant changes of an intervention, when a consensus is reached on well-defined and equally understood collective concerns. Regrettably, this describes the exact opposite situation of complexity.

[19] Sen nevertheless argues that, while some inputs are less rational in the collective sense, they should not be disregarded. Excluded contributions should be acknowledged and possibly redirected to policy considerations with the narrower collective scope of choice, such as those without overlapping domains, like sector-specific or thematic policies (Radej, 2022). If even this redirection proves infeasible, other options may be open in progressively narrower frames (sub sectoral, etc…).

[20] A relative majority means the largest single group has more votes than any other individual group, but still constitutes less than half of the total votes.

A representative example is the MSC.[21] It aggregates narratives of significant changes through a hierarchical filtering process, gradually progressing from the micro level (individual participants) to the macro level (key stakeholders). Participants discuss and submit the most noteworthy stories within their domain to the next higher level. The subsequent level selects the most significant changes from all submissions before forwarding them upwards. This aggregation effectively condenses a large volume of locally significant stories into a smaller collection at the macro level, representing the most widely endorsed narratives. However, majority voting accommodates synthesis within existing asymmetrical power dynamics so it risks becoming a lever of bias.[22] Besides, this aggregation logic raises concerns regarding the scope of the MSC as a tool for evaluating complex interventions, where no single 'most significant' change holds a superior epistemic status. Lastly, the same selective synthesis logic is applied in the MSC from the lowest to the highest level of aggregation, despite the incommensurability of lower and higher levels. Complex systems exhibit different behaviours at different levels, and the relationships between these levels are often not well understood (aggregation problem in collective choice!) and rarely applied appropriately in methodologies of evaluation synthesis.

Selective aggregation shares implications with partial aggregation methods. The latter aggregates data from the micro to meso level, producing sub-aggregates or sub-maps differentiated by groups, sectors, themes, or domains. While partial aggregation effectively highlights discrepancies and conflicting viewpoints between domains at the group level, it leaves participants and stakeholders with the most challenging task in evaluation: interpreting tensions and synergies between domains in the second phase of synthesis, from the meso to macro level.

Davies and Dart note appropriately that partial aggregation methods cannot be applied as standalone methods for generating summative judgments. They can merely complement result-based evaluation within a mixed-methods evaluation strategy. Understanding participatory evaluation not as autonomous but only a supportive tool has very negative consequences, reducing participatory evaluation from a leading tool of the *demos* to merely an auxiliary device of kratos, again mostly aiding collective rationality of the included minority.

'Facilitated synthesis' is another prominent constructivist approach that narrows the collective rationality of participatory evaluation. OH and CM employ ¸it for synthesis from meso to macro level.[23] It requires the evaluator to renounce the role of social aggregator and instead become solely a facilitator, wisely guiding stakeholders through a sensemaking or decision-making synthesis (Wilson-Grau). The evaluator-facilitator is asked to act as an 'invisible

[21] In e-mail correspondence from 9. June, Davies outlined a distinction between '*summary-by-selection*' and '*summary-by-inclusion*' (reffered to in *An Evolutionary Approach to Organisational Learning: An Experiment by an NGO in Bangladesh*, in Mosse, D., Farrington, J., and Rew, A., eds. 1998. Development as Process: Concepts and Methods for Working with Complexity. London. Routledge). '*Summary-by-inclusion*' concept is defined in commensurable terms (see here. Accessed December 2024), which is a form of selective synthesis (Radej, 2021a).

[22] Significant changes that are not submitted for a discussion at the higher level of evaluation are not disregarded in MSC but relegated in evaluation conclusions to a subordinate position within the hierarchical structure.

[23] Facilitation in collecting narratives, as in MSC, is not problematised until directed by the evaluator.

agent' (Guijt et al.), a discreet mentor, gently guiding participants through a discovery process (Guijt et al.). Patton proposed in a blog published on April Fool's Day, The evaluator act a jester[24] who 'speaks truths to power' through humour – not through connecting facts with values. Evaluator's task is to "*train people to understand evaluation and how the participatory process works, as well as teaching them basic research skills*" (Wilson-Grau). The evaluator should merely assess the quality of evidence, identify crucial questions for consideration, and propose topics for discussion while strictly refraining from making qualitative judgments about aggregate findings and their decision-making implications. The facilitated synthesis approach aligns with guidelines published by the World Bank asking that stakeholders evaluate and evaluators facilitate (Rietbergen-McCracken, Naray).[25]

However, the roles of the evaluator and facilitator are completely different (Aubel). An evaluation provides neutral judgments, while facilitation enhances communication effectiveness and quality. Attaining facilitation and coaching skills is valuable in evaluation, but these skills can complement, and never replace evaluative thinking (Vo, Archibald), since it radically surpasses the collective horizon of a facilitator, coach, and even a jester.

A recurring theme within constructivist evaluation is that the evaluator's original role should diminish in meso-to-macro level synthesis. Following this guideline, an evaluator enters the empty middle and stays there immobilised and disoriented about what is best for all, while presenting this state as neutral. This approach echoes a pre-complex mentality with a Leopoldian aversion to evaluative synthesis, this time obscured by buoyant constructivist justification. The efforts to divert evaluators from their distinctive role in bridging gaps between opposites present a paradox: the tools declared as interpretive find their mission in transferring this sacred task of the evaluator – 'where he is needed most' – to participants and, generally, non-evaluators.

Wilson-Grau indeed observed a reluctance among many key stakeholders at the macro level to engage in interpreting the meaning and relevance of evaluation findings obtained at the meso-level. This reluctance likely stems from the belief that interpreting significance of participatory evaluation findings to decision-makers is the evaluator's core responsibility. Merely organising and moderating participants' discussions is insufficient. Imagine an evaluator as an auto mechanic who, instead of fixing your car, hands you the tools and teaches you how to repair it professionally while offering encouragement and moral support. In doing so, the evaluator only shirks responsibility for the common good, violating one of the core ethical standards in policy evaluation.[26]

Recent decades have witnessed an epistemic turn in evaluation theory away from social constructivism to more realist approaches again. However, while intersectional constructivist-realist epistemology appears on the surface integrative, as previously illustrated by the four

[24] Youtube. Accessed December 2024.

[25] But not endorsed by the Bank since the publication contains the standard Word Bank's disclaimer.

[26] For instance the fifth Guiding Principle for Evaluators of the American Evaluation Society, asking that "*Evaluators strive to contribute to the common good and advancement of an equitable and just society.*" See also epigraph at the beginning of this Paper.

tools, it has critical pitfalls, raising crucial concerns about their methodological coherence and credibility of their results.

A prominent example is the asymmetrical treatment of subjective and objective methods or their results. Standard constructivist-realist evaluations often intersect these aspects asymmetrically, habitually privileging the objective over the subjective or vice versa, rarely achieving balance and seldom integrating them. The request that evaluators present stakeholders with objectively verified participatory contributions (as in CM or MSC) is biased in favour of a realist view. Similarly, participatory contributions are sometimes interpreted only when translated from their original qualitative to quantitative expressions through transcription, statistical, and analytical methods (as in SM and CM). Sometimes, the intersection is inauthentic, with one axis skewed or dominant, and even undesirable from the perspective of the collective good (Radej, 2022) because they are undemocratic or intolerant.

Constructivist-realist methods furthermore tend to enhance methodological eclecticism. It manifests in the uncritical combination of disparate theoretical frameworks within participatory evaluation. While methodological pluralism enriches evaluations, eclecticism risks superficial integration rather than deep synthesis. For example, SM combines a partial aggregation approach with a facilitated synthesis – the former belongs to mid-range evaluation, while the latter grounds in macro-level evaluation, which is explicitly refused in the former. Similarly, the OH draws on two distinct evaluation concepts: utilization-focused evaluation and responsive evaluation. The former embraces a goal as state-oriented stance, while the latter focuses on the process.

Disparate epistemic frames diverge in describing reality as complex (partly ordered, partly disordered), chaotic (disordered), or systemic (ordered; Stacey), reflecting distinct basic assumptions about the nature of reality. Some view reality as objective, others as subjective, experienced, negotiated, or co-constructed, dynamically evolutionary or context-dependent. Different assumptions invoke disparate explanatory theories, casting the evaluator in varied roles: analyst-expert, aggregator, facilitator, coach, interpreter, or social agent. The uncritical intersection confronts methodological pluralism with the aspiration for synthesis. It risks epistemic distortion with reinforcing narratives, which results in reductive synthesis and collective rationality of participatory evaluation.

Intersectional evaluations are not inherently problematic, provided their logic is coherent. Their coherence depends on the robustness of the theoretical framework that explains how disparate methods support one another in contributing to the enhancement of collective rationality and social inclusion. Interdisciplinary models should offer meta-theoretical justification, built upon meta-methodological insights grounded in meta-data. Constructivist-realist methodologies often lack deep meta-level understanding, either because they are underdeveloped – do not aggregate beyond the meso-level – or because their elaboration is overstretched to the macro-level. A lack of meta-level methodological justification traps intersectional methods in meta-blindness (Tremain): instead of resolving epistemic blindness, it enhances it further.

Intersectional evaluation methods are constructed with middle-level thinking. However, to be authentic, the intersection must occur precisely in the empty middle. Without fulfilling this basic condition, the intersectional methods impulsively crisscross through the problem space generating the illusion of intermediation, superficial and self-indulgent logocentric plurality in absolute thirdness, where different groups and institutions become entrenched in self-referential modes of thinking focused on multi-polar centrality rather than higher collective rationality. Logocentric pluralism between evaluation domains or between collective choice alternatives is immature. It is anchored to itself, where differences are amplified or showcased without integration at collective level (Mouffe in *The Democratic Paradox*).

Participatory evaluation is authentic only when the middle ground (intersectionality) and the void (bias, blindness) coincide. This coincidence forms an indeterminate space with a new epistemology, articulating complex issues in the language of shadows, through relative thirdness and relative void, using triaxial dialectics in the secondness of thirdness, and employing a double exposition of bias and blindsighted reasoning. As participants oscillate in between opposing valuations of complex interventions, they repeatedly traverse the empty middle, absorbing the indeterminacy of the evaluated concern. They first acknowledge biases in their valuations and recognize the exclusions these biases produce. Then, they accept the void as a common denominator—a shared epistemic reference point—securing epistemic indeterminacy in the evaluation and ensuring their valuations contribute to an evaluative synthesis. This synthesis is assured through shared indeterminacy rather than forced integration founded on uniformity. Engagement with the empty middle is iterative (Radej, 2022), with participants repeatedly crossing it during evaluation: first by organizing a dialogical space akin to a round table, then by collecting and analyzing indeterminate data, followed by the non-totalizing synthesis of opposing insights, and finally by interpreting evaluation results at a meta-level (see Annex 1).

New approach to participatory evaluation from the empty middle achieves neutrality, enabling evaluation to transcend participants' logocentric valuation of alternatives in collective choice without becoming logocentric itself.

The idea of the empty middle aligns with the prominent Kyoto School philosopher Watsuji Tetsurō's assertion: "*betweenness is emptiness, and emptiness is betweenness*" (Carter, McCarthy, p. 10). It describes the world's continuous emergence through the interdependence of intersected opposites that contain the void at their core. Watsuji drew from Nishida, the founder of the Kyoto School, and his thinking about a relational space situated in the 'space of nothing.' Both thinkers built on the eight centuries-old wisdom of Tsongkhapa, the Buddhist dialectic philosopher of the Middle Way. He admitted that he was deeply intellectually moved by the realisation of the simultaneity of the continuous birth of the world through the interdependence of all things and the emptiness at their core (Jinpa).

This ancient wisdom offers a remarkably insightful frame for describing anti-postmodern phenomena.

V. Participatory Evaluation from the Empty Middle

The effectiveness of the four tools in achieving specific evaluation tasks has been proven many times. They can protect participation and its collective rationality from unsubstantiated, unstructured, and non-communitarian contributions. They also foster more democratic decision-making by incorporating the perspectives of participants, particularly those from disregarded groups.

The four tools are nonetheless recognised for their inefficiency (Van der Merwe et al.), poor methodological consistency, and weak policy, program, or project relevance. Their contributions to furthering two imperatives of collective choice are constrained and not neutral due to oversimplified approach and unresolved methodological obstacles. As illustrated, the dominant approaches decisively fall short of a promise to achieve the "*radical and ground-breaking*" transformation (Klugman in Wilson-Grau, p. 7) of evaluation. On the contrary, they are most effective in simple, complicated (systemic), or chaotic situations but truly struggle in evaluating complex interventions where order (collective rationality) and chaos (bias) work together. Cooke and Kothari realised in their book (in *The New Tyranny*) that public participation is usually a bureaucratic ritual that legitimises existing power structures and masks continued power imbalances. Chambers similarly argued that despite good intentions, participatory methods often reflect the reality of the powerful minority rather than the powerless majority.

The paper explored the challenges of ensuring neutrality in participatory evaluation of complex interventions beyond democratic procedures and scientific evidence. The four tools achieve neutrality by mitigating participants' biased, logocentric valuations of collective issues. Conversely, the novel evaluation shifts the focus from truth-based neutrality to neutrality based on indeterminate judgments. It evaluates participants' biased valuations as relative truths. This is logically sufficient to ensure the neutrality. Biased evaluations of collective concerns are unavoidable and legitimate, provided participants acknowledge the empty middle as the framework for their presentation. This is an anti-postmodern extension of conventional democratic logic, legitimising a broad spectrum of oppositions, as long as that these oppositions live their antagonistic lives within the framework of democratic rules.

Participatory evaluation affirms blindness as an epistemic resource rather than attempting to eliminate, outwit, or submit to it (Radej, 2021b). Ensuring neutrality requires shifting the focus from participants' biases to the evaluator's mesoscopic authenticity; this is from logo-centre to the empty middle. The empty middle, as the intersection between different voids, is more inclusive because it recognises the multi-dimensional nature of exclusion. It is also more collectively rational. By focusing on the interplay of invisible voids, the evaluation offers a truer, more actionable understanding of complex interventions, better aligned with the needs of all stakeholders, while revealing new strategies and mechanisms for their provision.

The paper concludes that neutrality ultimately depends on how participatory evaluation approaches complex situations, avoiding extremes while adopting the language the shadows in the empty middle. The demand for neutrality requires evaluation not to exclude the evaluators but the excluders (Mouffe) – those who reject the kratos of democracy by

disregarding principles of dialogue, such as mutual respect, sincerity, and openness. Exclusion is also warranted when participants resist organisational, institutional, and structural learning, reject collective sensemaking, or refuse to share responsibilities in collective action. Certain groups or individuals are excluded because they fail to meet dialogical criteria (e.g., coherence, reflexivity; Schwandt), such as hate groups, or they reject science-based policies or perpetuate exclusionary ideologies. The exclusion likewise applies to those who fail the test of the empty middle.

The empty middle is indeed exclusive but not indiscriminately or definitively. By excluding of excluders, participatory processes align with a Rawlsian ethos of fairness (in *A Theory of Justice*), while fostering the emancipatory potential discussed by Derrida and Mouffe. Excluders are not excluded for holding differing opinions or pursuing their core values. They are not excluded from democratic deliberation nor from the comparative evaluation of diverse perspectives. They are excluded, most strictly, from evaluation synthesis during the final phase of deliberation, because their contributions could limit the scope of synthesis or undermine collective sensemaking. In this way, the evaluator safeguards collective concerns against incoherence, while preserving the possibility for inclusion of excluders should they accommodate themselves in the empty middle. Democratic societies do not ignore excluders, just the opposite, they allocate substantial communal resources to critically examine and mitigate exclusionary assertions while incentivising actors to engage constructively.

A compelling example involves excluding the participation of actors who spread misinformation. Governments and civil society organisations may collaboratively enforce measures such as regulating media platforms or establishing fact-checking mechanisms to counter misinformation, thereby fostering a more informed public sphere. Over time, as societal norms evolve and misinformation loses its influence, individuals or groups previously excluded can re-engage with democratic processes.

The concept of the empty middle is not new to democratic societies, as they are built upon it. Collective choice and blindsided evaluation are examples discussed earlier and there are many more to consider.

Take, for instance, the separation of political powers into the executive, judicial, and legislative branches. Separation prevents any one branch from holding absolute authority. The tripartite division organises political power around an empty middle space open to negotiation, checks and balances, and participation from various societal actors, ensuring no single entity dominates.

The court is similarly organised around the empty middle, positioned between arguments *pro et contra* before an impartial judge or jury. Derrida further posits that legal decisions often involve a moment of undecidability. From this indeterminate space, the judge exercises discretionary judgment, guided by the wisdom of blindsight, ensuring that legal decisions are not only legally sound but also just and fair in their application to real-life situations.

Where solutions require a collective effort, encountering a void in the middle is common. For instance, in interfaith or intercultural dialogue, cross-border cooperation, or the governance of common resources. Groups of people gather around a round table with an empty centre.

A further example is scientific inquiry, which must traverse the empty middle to navigate complexities in data interpretation, hypothesis testing, and theory formulation (Popper in *The Logic of Scientific Discovery*). The void in the middle is also emphasised during scientific revolutions (Kuhn in *The Structure of Scientific Revolutions*) when paradigms shift: the old paradigm no longer provides satisfactory explanations, and a new one has yet to emerge. Transdisciplinarity also operates from the empty middle; its findings emerge from the empty middle between multiple single-discipline perspectives.

The empty middle gives various art forms meaning, structure, and functionality. Poetry is the space between the words (May). Music is the silence between the notes (Debussy). In architecture, the emptiness between walls defines the essence of space, echoing Lao Tzu's wisdom. In short, for the blindsighted, the empty middle is present everywhere.

Especially in the field of evaluation, from which the empty middle originates!

Evaluation has undergone a three-stage transformation in recent decades in response to the increasing complexity of public affairs. The first stage saw a postmodern shift from rationalist to design-based constructivist participatory evaluation (the fourth generation; Guba, Lincoln). Yet, its characteristic detachment from external constraints and unrealistic constructs again revived more realist approaches in evaluation.

The rise of constructivist-realist evaluation marks the metamodern turn (the fifth wave; Vedung) in mainstream doctrine. It created space for new forms of discourse and governance that are less oppressive and more receptive to diverse perspectives. Metamodernity established itself as a dominant culture of continuous revision and re-examination of discourse, emphasising intersectionality, mutuality, context-dependence, and co-dependence between opposing evaluations. In participatory evaluation, the focus has been placed on pluralism, enabling shared processes without requiring uniform conclusions among participants.

Yet, constructivist-realist approaches are not sufficiently synthesising, susceptible to asymmetries, methodological eclecticism, a lack of rigour, and a polar attitude to indeterminacy and bias. They do not meet the requirements for neutrality superiorly better than traditional result-based or design-based evaluations. While they address old problems to some extent successfully, they also create new ones that ultimately lessen their overall impact.

Moving beyond the metamodern stage constitutes a concluding third step, marking an anti-postmodern turn (Radej, 2022). It embraces the metamodern intersectionality but also embraces the void in the intersected middle. This void represents ‘a generative absence’ (Žižek, 2006, p. 236), harbouring the potential for creating something profoundly new that is neither apparent from a constructivist-realist perspective nor merely in opposition to it.

Participatory evaluation arises now as a tool for cultivating generative absence. Its core function is no more to resolve contradictions but to nurture authentic deliberation. Participatory evaluation cultivates an empty middle where biased views transform into blindsighted ones, binary thinking shifts toward trialectics, and dualist antagonisms evolve into complex relationships from secondness to the secondness of thirdness.

When the democratic process develops such profound transformative capacity, all community members benefit as demos and kratos. On the one hand, pursuing concerns of the demos enhances a new form of kratos that is more collectively rational. Conversely, the blindsighted participatory evaluation model enables kratos to operate without enforcing what occupies the centre, making it less restrictive and repressive. This corresponds to Foucault's idea of power (in *The History of Sexuality*) operating not only through direct domination but also through subtle guidance and normalisation, shaping behaviour without overt coercion in a way that is not authoritative.

Blindsighted participatory evaluation presents a mental model for reasoning about the anti-postmodern society. It transforms how we approach collective issues, envision the future, and understand the nature of knowledge by absorbing the void in the middle. The third shift triggers a radical transformation or ‘a radical rupture’ (Badiou) in the status quo, opening new possibilities for understanding complex issues that were previously unthinkable. In the new age, evaluation becomes a tool for navigating social complexity, a democratising force, and a catalyst for social transformation.

Bibliography

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

Aubel J. 1999. Participatory Program Evaluation Manual Involving Program Stakeholders in the Evaluation Process. 2nd ed. Child Survival Technical Support Project and Catholic Relief Services.

Axelrod R. 1976. Structure of decision: The cognitive maps of political elites. Princeton University Press.

Bartels S.A., S. Michael, L. Vahedi, A. Collier, J. Kelly, C. Davison, J. Scott, P. Parmar, P. Geara. SenseMaker. Evaluation and Program Planning, Volume 77(2019).

Bauman Z. 2000. Liquid Modernity. Polity Press.

Bibri S.E. Backcasting in futures studies: a synthesized scholarly and planning approach to strategic smart sustainable city development. Eur J Futures Res 6/13(2018).

Carter R., E. McCarthy. "Watsuji Tetsurō", in The Stanford Encyclopedia of Philosophy (Winter 2019 Edition), Edward N. Zalta (ed.).

Chambers R. 1997. Whose Reality Counts? Putting the First Last. Practical Action Publishing.

Chödrön P. 2002. When Things Fall Apart: Heart Advice for Difficult Times. Boulder, Shambhala Publications.

Copestake J., M. Morsink, F. Remnant. 2019. Attributing Development Impact: The Qualitative Impact Protocol (QuIP). Rugby, Case Book. Practical Action Publishing.

Crenshaw K. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. Stanford Law Review 43/6 (July, 1991):1241-99.

Creswell J. W., J.D. Creswell. 2022. Research design: Qualitative, quantitative, and mixed methods approaches (6th ed). Thousand Oaks, CA: Sage Publications.

Davies R., J. Dart. 2005. The Most Significant Change (MSC) Technique: A Guide to Its Use.

Deprez S. Using a Narrative Inquiry Approach to Guide Project Interventions: SenseMaker in Practice. UK aid - Smart Peace, MEL Practice Paper 3, August 2021.

Derrida J. 1978. Writing and Difference, trans. Alan Bass. London: Routledge

Duflo E., R. Glennerster, M. Kremer. Using randomization in development economics research: A toolkit. Cambridge, MA, NBER Technical Working Paper Series, Technical Working Paper 333(Dec. 2006).

Dumézil G. 1987. Tridelna ideologija Indoevropcev. Transl. and foreword by D. B. Rotar. Ljubljana, ŠKUC, Znanstveni inštitut Filozofske fakultete. Zbirka Studia Humanitatis.

Eurobarometer. 2005. Special Eurobarometer 215 / Wave 62.1 – TNS Opinion & Social. European Commission, DG Press and Communication.

Feyerabend P. 1975. Against Method. Outline of an anarchistic theory of knowledge. London, Humanities Press.

Fricker M. Epistemic Injustice: Power and the Ethics of Knowing. Oxford. Online edition, 1 Sept. 2007.

Guba E.G., Y.S. Lincoln. 1989. Fourth Generation Evaluation. Sage.

Guijt I., M.A. Gottret, A. Hanchar, S. Deprez, R. Muckenhirn. 2022. The Learning Power of Listening, Rugby, Practical Action Publishing.

Hageboeck M., M. Frumkin, S. Monschein. 2013. Meta-Evaluation of Quality and Coverage of USAID Evaluations 2009-2012. Management Systems International – DevTech Systems, USAID.

Heidegger M. 1999. What is Metaphysics? In Pathmarks. Trans. W. McNeill, Cambridge University Press.

Henning D. 2007. Poll shows 82 percent of Germans feel politically disenfranchised. The International Committee of the Fourth International, World Socialist Web Site, Jan. 16.

Hilhorst T., I. Guijt. Participatory monitoring and evaluation: a process to support governance and empowerment at the local level. World Bank, A guidance paper 558 129(2006).

Huitema D., A. Jordan, E. Massey, T. Rayner, H. van Asselt, C. Haug, R. Hildingsson, S. Monni, J. Stripple. The evaluation of climate policy. Springer, Policy Science, 44/2(2011):179-98.

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

Jinpa T. 2002. Self, Reality and Reason in Tibetan Philosophy. Tsongkhapa's Quest for the Middle Way. New York, Routledge Curzon.

Kahneman D. 2011. Thinking, Fast and Slow. New York, Farrar, Straus and Giroux.

Kaplan R. 2000. The Nothing that Is: A Natural History of Zero. New York, Oxford University Press.

Kreisler H. 2001. Conversation with Manuel Castells. Identity and Change in the Network Society. Berkeley, Institute of International Studies, Conversations with History series.

Laclau E., C. Mouffe. 1985. Hegemony and Socialist Strategy: Towards a Radical Democratic Politics. London: Verso Books.

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

May R. 1982. Freedom and Destiny. New York, W. W. Norton & Company.

Pabst A. Political Economy of Virtue: Antonio Genovesi's ‘civil economy’ alternative to modern economic thought. European Journal of the History of Economic Thought 25/1(Julij 2018):1-23.

Patton Q.M. 2002. Qualitative Research and Evaluation Methods. 3rd ed. Sage.

Patton Q.M. 2010. Developmental evaluation applying complexity concepts to enhance innovation and use. New York, Guilford Press.

Patton Q.M. 2012. Utilization-Focused Evaluation. Sage.

Pawson R., N. Tilley. 1997 Realistic Evaluation. London: Sage.

Peirce C. S. 1966. Collected Writings. Edited By C. Hartshorne, P. Weiss, A. W. Burks. Cambridge, Harvard University Press.

Pierson E., A. Feller, S.Goel, A. Huq. 2017. Algorithmic decision making and the cost of fairness. ArXiv.

Powell S. Visualizing theories of change: How not to confuse causes and definitions. EVal., December 2017.

Powell S., J. Copestake, F. Remnant. Causal Mapping for evaluators. Advance. Preprint, 2023.

Radej B. 2021a. Aggregation Problem, in Radej B., M. Golobič, 2021.

Radej B. 2021b. Organisation Problem, in Radej B., M. Golobič, 2021.

Radej B. Je anti-postmoderna družba še lahko liberalno demokratična? Odgovor iz avtentične sredine. Ljubljana, Slovensko društvo evalvatorjev, Delovni zvezki SDE št. XV/1(Dec 2022), 40 str.

Radej B. Več-nivojski vidik družbene kompleksnosti in njen pomen za metodologijo družbenega raziskovanja - primer vrednotenja učinkov politik. Ljubljana: Slovensko društvo evalvatorjev, Delovni zvezki, let. 3, št. 3 (september 2010), 57 str.

Radej B., M. Golobič. 2021. Complex Society. In the Middle of a Middle World. Vernon Press.

Radej B., M. Golobič. Vrednotenje učinkov osnutka Strategije prostorskega razvoja Slovenije 2050 na teritorialno kohezijo. Ljubljana, Slovensko društvo evalvatorjev, Delovni zvezki XIII/1(Okt. 2020).

Rietbergen-McCracken J., D. Naray. 1998. Participation and Social Assessment. The World Bank.

Sartre J.-P. 1943. Being and Nothingness. Hazel E. Barnes.

Schrödinger E. 1996. Nature and the Greeks, and Science and Humanism. Foreword by R. Penrose. Cambridge University Press, Canto.

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

Scriven M. Evaluation synthesis: The new frontier. In: Donaldson S.I., Scriven M., eds. Evaluating social programs and problems: Visions for the new millennium. Mahwah: Lawrence Erlbaum Associates; (2003):129-44.

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

Sen A. 2011 (1983). Rational Choice, Collective Decisions, and Social Welfare. Cambridge University Press.

Spivak G.C. Can the Subaltern Speak?, In Marxism and the Interpretation of Culture, edited by Cary Nelson and Lawrence Grossberg, 271-313. London: Macmillan, 1988.

Stacey R. D. 1996. Complexity and Creativity in Organizations. San Francisco, CA Berrett-Koehler Publishers.

Stiglitz J. E, A. Sen, J.-P. Fitoussi. 2009. Report by the Commission on the Measurement of Economic Performance and Social Progress. Paris, CMEPSP.

Stufflebeam D.L. 2003. The CIPP Model for Evaluation. In: Kellaghan, T., Stufflebeam, D.L. (eds) International Handbook of Educational Evaluation. Kluwer International Handbooks of Education, vol 9. Springer, Dordrecht.

Tanaka-Ishii K. Void of sign. Semiotica, 218(2017):119-35.

Tremain S. 2017. “Knowing Disability, Differently” in The Routledge Handbook of Epistemic Injustice, ed Kidd I.J. , J: Medina, G. Pohlhaus Jr., Routledge.

Van der Merwe S.E., R. Biggs, R. Preiser, C. Cunningham, D.J. Snowden, K. O'Brien, M. Jenal, M. Vosloo, S. Blignaut; Z. Goh. Making Sense of Complexity: Using SenseMaker as a Research Tool. Systems, 7/2(may 2019):25.

Vedung E. Four Waves of Evaluation Diffusion. Sage, Evaluation, 16/3(July 2010):263-77.

Vermeulen T., R. van den Akker. Notes on meta-modernism. Journal of Aesthetics & Culture, 2/1(2010):56-77.

Vignjević T. 2004. Trije redovi sveta. Upodobitve treh redov in stanov na prehodu v Novi vek (Three orders of the world. Depictions of three orders at the transition to Early Modernity). Ljubljana, Založba /cf*, Oranžna zbirka.

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.

Vo A. T., T. Archibald. New directions for evaluative thinking, in A. T. Vo, T. Archibald (Eds.), Evaluative Thinking. New Directions for Evaluation. 158(2018):139–47.

Wallerstein I. 1989. The Modern World-System. III*:* The Second Era of Great Expansion of the Capitalist World-Economy, 1730-1840s. San Diego, Academic Press.

Wilson-Grau R. 2019. Outcome Harvesting Principles, Steps, and Evaluation Applications. Charlotte, Information Age Publishing.

ANNEX 1: Crossing the empty middle: An example of participatory evaluation-supported preparation of a national strategic development document

How does crossing the empty middle look in practice? Let's take the example of preparing a strategic development document step by step, from the initial definition of the situation to weighing options, forming solutions within coalitions for change, and finally evaluating the impacts of efforts to solve a strategic problem. The preparation of the document oscillates between participatory and systemic phases. The systemic phase encompasses the stages carried out by official institutions, while the participatory phases are open to public participation. Each step crosses the empty middle at least once, allowing the preparers to adopt an authentic middle-ground mindset.

First, the strategic development challenge is defined as a systemic problem. Planning foundations are prepared, including an assessment of past trends in key indicators across different areas. Synergies and conflicts among strategic domains are analysed and problematic areas identified (crossing the empty middle). Initial study also considers findings from independent sources like civil society (crossing the empty middle), to verify definition of strategic challenges or supplement them. Priority issues are defined and alternative solutions are identified.

In the second step of the preparation process, the public is involved through a participatory process. Participants agree rules of the proces together. The goal is to develop dialogue in the empty middle (crossing the empty middle). Participatory process also gives a voice to certain groups, overlooked in the previous step (crossing the empty middle). The aim is to substantiate the necessary changes in the definition and approaches to adressing the strategic problem, together with key constraints and milestones determined in the previous preparation phase.

A systemic step follows again. A multi-criteria (horizontal) and multi-level (micro, meso, and macro) analysis of the alternative solutions identified in the previous phase is conducted. The comparative analysis should demonstrate the impact of the alternative proposals on resolving the strategic problem and refine the proposed alternatives. The rationale for disregarding well-founded but disregared proposals is established, jusifying the exclusion of certain issues from consideration in the next step (crossing the empty middle).

The process continues with a participatory phase. The goal is to form the broadest achievable coalitions for change. This phase can be carried out through voting by public representatives or even through direct public participation (crossing the empty middle). The aim is not necessarily to achieve a coalition 'in favour,' but perhaps a coalition of minimal resistance—where no group exercises a 'veto' (crossing the empty middle).

The next step is again systemic, the preparation of the final document. Some well funded but excluded proposals might warrant further attention, for example, through the targeted research or experimental projects in selected narrow problem areas (crossing the empty middle).

If the preparation of the document begins with a systemic phase, it would conclude with a participatory phase. For instance with participatory evaluation of resolving the strategic

problem and impact on public welfare. Attention is also given to evaluating previously overlooked aspects, side and unforeseen effects, key risk factors, and outcomes that intersect with the findings of studies or pilot tests of excluded proposals (crossing the empty middle).

COLOFON

Title	Inclusive or Rational Participatory Evaluation?
Author	Radej Bojan
Edition	1st edition
Location and Publisher	Ljubljana, Slovenian Evaluation Society
Publishing year	2024 (Volume XVII, Number 1)
Series Title	Working Paper SDE
Property rights holder	Slovenian Evaluation Society, Creative Commons 2.5, Slovenija
Circulation	Electronic publication
Publication's International Identificator	ISBN 978-961-92453
Retail Price	Free

Published so far

Year	Title	No.
I/2008	Vaje v seštevanju neseštevljivega (B. Radej, 23 str.)	1
	Sinteza vplivov nacionalnega energetskega programa na prostorsko kohezijo Slovenije (B. Radej, 43 str.)	2
	Meso-Matrical Synthesis of the Incommensurable (B. Radej, 21 str.)	3
II/2009	Anti-systemic movement in unity and diversity (B. Radej, 12 str.)	1
	Meso-matrical Impact Assessment - peer to peer discussion of the Working paper I/3(2008) (B. Radej, 30 str.)	2
	Turistična regionalizacija Slovenije (J. Kos Grabar, 29 str.)	3
	Presoje javnih učinkov vladnih politik (B. Radej, 18 str.)	4
	Ciljno usmerjen državni proračun: med obeti in možnostmi (B. Radej, 33 str.)	5
III/2010	Beyond »New Public Management« doctrine in policy impact evaluation (B.Radej, M.Golobič, M.Istenič,25 str.)	1
	Osnove vrednotenja vplivov javnih politik za priložnostne uporabnice/ke (B.Radej, 43 str.)	2
	Več-nivojski vidik družbene kompleksnosti - primer vrednotenja učinkov politik (B. Radej, 57 str.)	3
IV/2011	Presečno določanje razvojnih prioritet (B. Radej, Z. Kovač, L. Jurančič Šribar, 45 str.)	1
	Primarna in sekundarna perspektiva vrednotenja politik (B. Radej, 30 str.)	2
	Agregacijski problem kompleksnih družbenih vrednotenj (B. Radej, 41 str.)	3
	Gibanje 99%: Z izključevanjem do skupnosti (B.Radej, 42 str.)	4
V/2012	Odličnost na kvadrat: samopresoja v javni upravi (B.Radej, M.Macur, 41 str.)	1
	Delna celota: primer prostorske kohezije (B.Radej, M.Golobič, 31 str.)	2
VI/2013	Divided we stand: Social integration in the middle (B.Radej, M.Golobič, 26 str.)	1
	With Exclusion to the Community (B.Radej, 23 str.)	2
	Jabolka in hruške: Sinteza brez skupnega imenovalca (B. Radej, 31 str.)	3
VII/2014	Apples and Oranges: Synthesis without a common denominator (B.Radej, 40 str.)	1
	Social Complexity: Operational definition (B.Radej, 65 str.)	2
VIII/2015	Measuring Smartness of Innovation Policy (B.Radej, K. Ž. Jazbinšek, M. Dolinšek, June 2015, 22 str.)	1
IX/2016	Interventna logika prostorskega razvoja v Sloveniji (B.Radej, M.Golobič, 38 str.)	1
	'Pametna' interventna logika – primer dediščinske politike (B.Radej, J.Pirkovič, 19 str.)	2
X/2017	Smart Heritage Policy (B.Radej, J.Pirkovič, P. Paquet, 17 pp.)	1
XI/2018	Sistemski vidiki vključevanja kulturne dediščine v razvoj (mag. Zvezda Delak Koželj, 19 str.)	1
XII/2019	Ob 10-letnici delovanja SDE (Bojan Radej, ur., 19 str.)	1
XIII/2020	Vrednotenje učinkov Strategije prostorskega razvoja Slovenije 2050 na teritorialno kohezijo (B. Radej, M. Golobič, 32 str.)	1
XIV/2021	Strateške izbire v času temeljne negotovosti (Bojan Radej, 29 str.)	1
XV/2022	Je anti-postmoderna družba še lahko liberalno demokratična? Odgovor iz avtentične sredine (B.Radej, 41 str.)	1
XVI/2023	Knjiga povzetkov prispevkov na 5. konferenci evalvatorjev WBEN (B.Radej, ur.)	1
XVII/2024	Inclusive or Rational Participatory Evaluation? (Bojan Radej, 41 pp.)	1