Check for updates

Evaluative and enabling infrastructures: supporting the ability of urban co-production processes to contribute to societal change

Henrietta Palmer[1*] , Merritt Polk[2] , David Simon[1,3] and Stina Hansson[2]

* Correspondence: henrietta.
palmer@chalmers.se
[1]Chalmers University of Technology,
Gothenburg, Sweden
Full list of author information is
available at the end of the article

Abstract

As widely attested in the literature, the evaluation of co-production is complex and unsuited to the use of conventional quality, monitoring and evaluation indicators. This reflects the uncertainties, co-contributory factors and time lags involved, particularly when seeking to assess institutional and wider societal effects of multi-stakeholder participatory processes and deliberative fora. The most widely assessed effects include the immediate outputs and outcomes of a project or activity (so-called first order effects) while wider societal or third order effects continue to be the most difficult to capture and, consequently, are the least well studied. Because of this difficulty, the intermediate, second order effects of organisational transformation and policy implementation constitute a growing challenge for evaluation. This is our focus here. After 10 years of transdisciplinary co-productive research practice, Mistra Urban Futures, as an interstitial research space bridging academia and practice working through city-based institutional partnerships called platforms, has reached a phase where some of these effects are becoming distinguishable. Accordingly, we discuss the prerequisites for co-production practitioners, including policy makers, to engage their respective organisations in transitional and incremental experimentation in order to achieve relevant institutional changes. This requires enabling infrastructures that support training, facilitation and the creation of 'safe' spaces to promote trust and legitimacy. These are needed to underpin the long-lasting personal and organisational commitments which are crucial to achieve transformative organisational effects.

Keywords: Transdisciplinarity, Co-production, Deliberative fora, Formative evaluation, Three order effects, Quality monitoring and evaluation, Enabling infrastructure, Mistra Urban Futures

BMC

Science highlights

- Urban transdisciplinary co-production is complex and requires adequate time, space and facilitation, as well as customised evaluation frameworks to ensure appropriateness to a project's objectives.
- By bringing diverse stakeholders together and forging relations of trust, such deliberative processes have important transformative value in themselves.
- Evaluation itself is insufficient for transdisciplinary co-production.
- Supportive infrastructure for transdisciplinary co-production takes several complementary forms that support training, facilitation and the creation of 'safe' spaces that promote trust and legitimacy.
- Measuring second (institutional) and third (societal) order effects is complex and involves time lags and diverse influences, but is important to demonstrate the wider value of such methodologies.

Policy and practice recommendations

- Urban transdisciplinary co-production approaches to service delivery and joint research hold good potential to overcome oppositional and compartmentalised organisational practices that bedevil conventional approaches.
- Nevertheless, investment in supportive infrastructures in the form of appropriate training provision, skilled facilitation and having a 'safe' space to promote trust and legitimacy is vital to realise this potential.
- Appropriate formative evaluation that feeds back into improving the ongoing processes adds important value to conventional quality monitoring and evaluation (QME) processes that focus heavily on quantitative indicators and outcomes.

Introduction

As cities are struggling with urban transitions and transformations to sustainability, both as sites of complex societal problems and as advocates of new ways of planning urban living environments, transdisciplinary co-production of knowledge can have a key role in systems of governance. Transdisciplinary (TD) co-production (CP) is a research approach to problem solving that addresses societal transformation through multi-stakeholder research collaboration. It includes both different academic disciplines and other professionals and experts from outside of academia as well as civil society actors taking part in joint research processes (Pohl 2011). This calls for inclusive knowledge production and evaluation that take into account diverse knowledges, needs and goals (Schipper et al. 2019; Kabisch 2019). Co-production processes focus explicitly on jointly formulating and undertaking the successive stages of the research together, from initial problem identification to analysis and, where possible, even implementation of the results (Lang et al. 2012; Polk 2015). Co-production targets various normative societal goals including empowerment, social learning, public service provision, and adaptive governance (Bremer and Meisch 2017).

The overall premise of transdisciplinary co-production is that inclusive collaborative processes are better able to both capture more relevant problem framings and the breadth of knowledge needed to design sufficient solutions, as well as increase the buy-in and legitimacy to implement such solutions. Within this context, it becomes

important to assess not only the substantive results of such processes, but also to evaluate the characteristics and qualities of the research process that enable this added value (e.g. Polk 2015; Hansson and Polk 2017). In many co-production situations, for instance where a local government organisation or public utility provider engages residents to co-produce public services, evaluation can be a relatively straightforward assessment of the producers' and consumers' respective perceptions of changes in the relevant qualities of the actual services (Watson 2014; Mitlin 2008; Durose and Richardson 2016). However, with increasingly wicked societal problems which engage diverse and numerous stakeholder groups, often with different mandates and power differentials, evaluation becomes far more challenging. In particular, the link between co-productive research processes and societal change is difficult to establish (Hansson and Polk 2018; Lux et al. 2019) on account of process properties, inherent time lags, multiple influences and related problems of the counterfactual (Spaapen and Van Drooge 2011; Wiek et al. 2014; Hellström 2015). This also reflects the specific complexity of such processes and problems areas, as well as the greater unpredictability and use of the intended results (Walter et al. 2007; Bornmann 2013; Koier and Horlings 2015; Belcher et al. 2016). In essence, such complexity is not subsumed within conventional quantitative and qualitative indicators, which are designed for more delineable and predictable processes (Blackstock et al. 2007; Jahn and Keil 2015; Roux et al. 2010). When applied to complex co-production situations, such indicators can even provide an unduly negative picture of underachievement or even failure in relation to stated objectives and timelines. Therefore, evaluation of transdisciplinary research needs to address both the substantive outputs and outcomes in terms of established collaborations, together with impacts such as organisational changes and possible societal transformations. These three types of effects are often framed as first, second and third order effects (Williams 2017).

The evaluation of participatory research processes – of which co-production constitutes a 'deep' subset – is a growing field of study (Walter et al. 2007; Wiek et al. 2014; Hansson and Polk 2017). Diverse evaluation schemes have been proposed that in various ways combine formative and summative evaluation (Harlen and James 2006), i.e. the use of evaluative measures inserted into the research process as a learning function versus evaluation of concluding results. One focus has been the development of frameworks for dealing with contextual, cultural and institutional factors (Bornmann 2013; Belcher et al. 2016). Specific attention has been paid to productive interactions, facilitation, co-reflection that produce shared accountability, and reflexive approaches for continuous adaptation of research processes (Roux et al. 2010; Spaapen and Van Drooge 2011; Popa et al. 2015; Reed and Abernathy 2016; O'Malley et al. 2019; Lux et al. 2019; respectively). Other proposals have concerned transition experiments, a focus on narrative evaluation methods and navigation of power in transitions (Spaapen and Van Drooge 2011; Luederitz et al. 2017; Schipper et al. 2019; Termer and Dewulf 2019).

In this paper, we present experiences with one approach to evaluating the broad societal impact of co-produced TD research, in a multifaceted quality monitoring and evaluation (QME) framework used at the research centre Mistra Urban Futures (Hellström 2015; Williams 2017; Williams and Robinson 2020). We combine a presentation of this framework with results from an impact evaluation done locally, after the first phase of the Centre, with project leaders and Centre partner co-ordinators in Gothenburg. From these experiences, we suggest the need for additional support structures

that can overcome the limitations of the QME framework, and better address the necessary flexibility that allows TD CP complexity to thrive in a beneficial way.

The QME framework at Mistra Urban Futures

The following discussion is based on experiences from Mistra Urban Futures, an international transdisciplinary research centre on urban sustainability (2009–2019). Mistra Urban Futures has utilised diverse forms of TD co-production in undertaking research to realise just cities, which we define as cities that are accessible, green and fair (Simon 2016). This includes a range of research topics framed within socio-spatial, socio-cultural and socio-ecological transformations. Further, it involves research on urban governance, urban knowledge and urban change. The research was undertaken by formal city-based institutional partnerships called Local Interaction Platforms (LIPs). Such platforms existed as different combinations of academic, public, private and civic organisations in Gothenburg, Malmö–Lund and Stockholm (Sweden), Sheffield–Manchester (UK), Cape Town (South Africa) and Kisumu (Kenya). Each city partnership had its own mechanism for identifying and prioritising local issues to be researched by some or all the partners.

Given this context, and in view of the requirement of our core partners[1] and funders, Mistra Urban Futures developed a five-part QME framework in a reflective and interactive process together with the different platform directors and co-ordinators as well as with representatives of funders and board members. It has now been run and refined through three annual cycles. The framework fulfils five complementary functions:

1. *knowledge management and transfer,* highlighting the actions and effects of the Centre's research more clearly to its participants and stakeholders;
2. *performance monitoring,* making it easier to assess the strengths and weaknesses of the Centre in relation to its goals and how they are achieved;
3. *governance*, creating the opportunity to track the effects of the Centre's activities, including tangible and intangible outputs and outcomes on the basis of which the Centre is accountable to its stakeholders;
4. formative *learning,* actively enabling the Centre to develop reflexively and improve in real time, as an integral part of its process management;
5. *legitimacy*, making information available to internal and external stakeholders which enables them to match Centre outputs and outcomes against expectations.[2]

To embrace the difficulties of evaluating co-produced TD research results, the five components of the QME framework capture both summative and formative outputs and outcomes, together with structural and societal impact. The first two components of the framework structure and govern annual reporting of outputs, and assess risks related to the management of each city platform and to the comprehensive endeavour of the Centre. The three latter components evaluate the outputs, outcomes and impacts

---

[1]The programme funders are Mistra, the Swedish Foundation for Strategic Environmental Research; Sida, the Swedish International Development Cooperation Agency; and the Gothenburg Consortium of seven partners, comprising two universities, a research agency and local and regional authority partners which initiated the Centre and undertake the research in Gothenburg.

[2]The scope of the QME framework was based on a pre-study by Tomas Hellström, University of Lund.

of the research activities in different ways. The first of these comprises a summative layout of qualitative and quantitative *performance indicators* organised along the three large themes related to: outputs of transdisciplinary research; outcomes such as capacity building; and impacts such as local, national and international agenda setting. The second constitutes a self-reflexive, *formative evaluation* where individual project members, the platforms, and the international Centre itself, assess the outputs, outcomes and impacts as a collective process of formative learning. The final component seeks to capture wider effects of organisational and societal changes traceable to the work of the Centre.

To link our evaluation approach more tightly to wider societal goals, Centre leadership collapsed the formative evaluation and impact components into one assessment of the Centre's central vision and hypothesis. This focuses on determining the extent to which co-produced transdisciplinary research, as organised in terms of structured comparative research, multi-stakeholder collaborative platforms and an international connectivity and collaboration, is contributing to the realisation of just cities. The evaluative component was consequently re-named the *Realising Just Cities Evaluation* (RJC) and was carried out as a detailed inquiry across relevant projects selected by each platform and across the platforms organisations themselves to capture results in terms of learning and change.[3]

The growing literature on the evaluation of co-production often distinguishes between different types of results in terms of outputs, outcomes and impacts, and different levels or orders of effects (first, second and third) (Wiek et al. 2014; Williams 2017; Williams and Robinson 2020). First order effects are directly related to a research process and comprise the immediate usable products (outputs), enhanced individual and collective capacities together with network effects (outcomes). Second order effects are the ones impacting the system which the process is working within and include decisions, new policies and organisational changes within the participating organisations (outcomes and impacts). Finally, third order effects are impacts that occur in the wider community or society and include new visions or changed behaviours, norms and practices leading to different 'societal imaginations' (Luederitz et al. 2017; Williams 2017; Williams and Robinson 2020).

The documented difficulties in distinguishing the three orders of effects and engaging with them individually through different QME mechanisms, have also proved challenging within Mistra Urban Futures. Accordingly, Williams (2017:4) stresses the importance of processes being "iterative, interactive and reflexive, that provide transparent discourse and collaboration, and embed broad and diverse participation and engagement". This points towards certain qualities of the co-production process that could also lead to first order results in terms of enhanced networks and capacity building. In recommendations to Mistra Urban Futures, Williams emphasised the need to regard these differing results as closely interlinked rather than as separate, and to be able to re-form and redesign the processes themselves to enhance delivery of outputs, outcomes and possible societal impact. The three orders of effects should be understood as "mutual [ly] reinforcing loops of influence" (Williams 2017:7).

---

[3]This evaluation is produced as a book, *Realising Just Cities*, to be published in early 2020.

Through collective deliberations, Centre leadership refined the QME framework to better capture not only first order effects, that are within the mandate of the Centre organisation, but also second and third order effects, that are respectively both partially and completely located outside the organisational structure of the Centre itself (Williams 2017; Institute for Methods Innovation 2019). Nevertheless, the Centre acknowledges that transdisciplinary co-production is not adequately captured by an evaluative mechanism alone, but needs yet another type of infrastructure that can *both* capture *and* support key elements of TD research such as developing skills together with increased learning and understanding.

Capacities to enhance the relationship of processes to impacts

As part of a prior evaluation of the QME work taking place in the Gothenburg platform (GOLIP) in 2015, project leaders and Centre partner co-ordinators from GOLIP's projects were asked to assess the quality of their project processes, their added value (their outputs, outcomes and impact), as well as to identify factors that supported or hindered their project enactment (Hansson and Polk 2017). The interview responses showed great variety in how the different projects had been supported financially, structured, reflected upon and how they had communicated projects' processes and results. GOLIP management assumed that projects that were well structured, had stable funding for long-term engagement, and undertook a conscious reflective process with support for collaboration and communication, would also show relevant and usable results. However, some of the projects with less developed collaboration also achieved results of sometimes comparable relevance and usability. The evaluation thus found no clear-cut correlation between the project processes and results. However, project participants from both subsets of projects identified several crucial issues for successful TD projects. Beyond the anticipated ones concerning time and funding, were the diversity and stability of participation along with process-related issues such as the ability to create space and autonomy for projects to develop organically rather than being prescriptively governed; the ability to change research questions to adapt to a changing context; and an explicit focus on learning through reflection and experimentation (Hansson and Polk 2017:16).

Hansson and Polk (2017:12) identified three capacities that are crucial for both co-production processes and their relationships to societal impact. These include: first, the capacity for learning through the ability to engage with new perspectives and understandings; second, the capacity for new forms of collaboration and ways of working together; and third, the capacity to establish and maintain relationships within and across organisations. These reflections fed into refinements of the Centre's QME framework. In the following text we will elaborate on how an explicit infrastructure could enable these capacities to flourish. As all three capacities challenge current institutional and organisational cultures, they require an infrastructure that is conducive to participation from both academia and practice. However, the results above also suggest that processes of TD co-production need a certain level of informality and openness in order to adapt to the changing context and be flexible in nurturing the process outputs, outcomes and impact. Hence, in addition to a relevant evaluative framework, we argue that co-produced transdisciplinary research needs additional support structures that allow

the complexity of TD CP to thrive in a beneficial way through targeting the capacities identified above.

Creating an enabling infrastructure to support transdisciplinary projects

What type of infrastructure most effectively enables our examples of transdisciplinary co-production processes to impact change? We suggest three components that together strengthen the research and institutional capacity to reach second and third order effects without delimiting the creativity and responsiveness of the processes. These components address capacities for learning, new ways of working together, and maintaining new types of relationships through training, facilitation and spatial support.

The first component focuses on training. The TD literature suggests that capacity building requires continuous learning efforts, since productive interactions between different stakeholders and knowledge cultures do not just happen but need to be developed and practised. This could be done in different ways, such as training in a particular method, in joint analyses of real-life wicked issues and in propositional ways to organise a TD research project around them (Wiek et al. 2014). Training can also include discussions of the theoretical underpinnings of concepts such as conflict, power/knowledge and reflexivity, and in programmes that include both researchers, practitioners and policy makers from the public and private sectors. Mistra Urban Futures' platforms have developed a number of different training alternatives for these purposes. One is the Gothenburg platform's Open Research School,[4] which invites both practitioners and PhD students to participate in a practice-based programme of method training and theoretical analysis. The Open Research School also organises Open Method Seminars addressing a wide local group of practitioners, researchers and professional facilitators to practice particular methods or together engage with the challenges in practising TD co-production.

Another example of training infrastructure is the Cape Town platform's Knowledge Transfer Programme (KTP). This infrastructure consists of an exchange program between the City of Cape Town and the University of Cape Town, where city officials are trained in theoretical considerations of their practice, and PhD students are embedded in City departments to provide critical perspectives on how the departments operate and highlight the value of potential bridges between administrative silos (Patel et al. 2015; Smit et al. 2020). Such targeted initiatives extend beyond training a handful of individuals for a particular project, and seek to build systemic capacity and competencies at the cohort level within and across organisations. The goal is to build up a critical mass of experienced individuals capable of carrying both new perspectives, awareness of wicked issues and tools for how to re-organise around their possible solutions into their home organisations. When they do so, they create a bridge from first to second order effects, ensuring that first order outputs such as articles and reports together with outcomes such as new networks and collaborations between academics and civil servants, are having second order effects in terms of policy- and decision-making in the participating organisations, or contributing to beneficial organisational changes as defined earlier. This illustrates how one particular type of support infrastructure can produce both first and second order effects according to the target and scale of activities.

---

[4]www.mistraurbanfutures.org/en/our-research/research-school

The second component focuses on facilitation. Jordan (2014:51) points out that complex issues require actors possessing "sophisticated capacities for managing different kind of complexities". As few individual actors might fulfil this requirement, skilful facilitation would instead enable groups to accomplish tasks that would otherwise be out of reach of each participating individual. Jordan (2014:50) adds that the knowledge of and interests in facilitation in collaborative processes have emerged primarily from the experiences of co-production among practitioners. Academics engaged in interdisciplinary and TD research tend spontaneously to take on a facilitator's role in these deliberative processes without necessarily having the requisite skills and competences (Stokols 2014). To understand the challenges of facilitation, Jordan identifies different stages in need of different kinds of support, either through guided deliberative methods or structured facilitation. These supports he refers to as a scaffolding (Hmelo-Silver et al. 2007; Stone 1993; Wood et al. 1976, in Jordan 2014), as it comprises deliberative methods and facilitation in terms of a skeleton carrying an ongoing process of construction. Jordan (2014:51) defines six major fields of functions which need an integrative scaffolding: attentional support; relationships; attitudes/ feelings; understanding; empowerment and creativity; and decision-making and co-ordination of action. This range of facilitated functions also points towards a sliding scale of outputs and outcomes of the first order towards possible outcomes of the second order, such as decisions and different kinds of actions.

To establish both awareness of these functions in a co-production process, and skills and capacities for their scaffolding, requires intentional infrastructural support. Consequently, at the Gothenburg platform, a local network of professional facilitators was invited to use the facilities and to 'spill over' into the activities of the platform. The platform also hired a professional facilitator to engage with the platform projects and to build up long-lasting structured support. To embed facilitation in TD processes is also to carry experiences from one TD process to another and, again, to build a critical mass, not of experienced individuals, but of experiences as such, reaching different individuals engaged in different processes respectively. Here, too, the scale and duration of effort required for effectiveness demonstrate a bridge from first to second order effects.

The third component focuses on the spatial dimensions of learning. Learning spaces and spaces for TD research are often discussed in broad and conceptual terms, where 'space' connotes an equal (Polk 2014) trustful and reflexive relationship between the different actors of the knowledge process (Pohl et al. 2010) and their mutual engagement (Perry and May 2010), or the space for action in terms of attention, or setting aside resources such as time, or the legitimacy of participation. Less has been written in the TD CP literature about the physical spatial implications and requirements of TD co-production and learning. Yet, participants in the Gothenburg platform's projects testify to the importance of having access to a 'neutral space', to be able to step out of daily roles and formal alignments to act in a new constellation unburden by these bounds (Hansson and Polk 2017). Others have called this space 'safe' as safe from judgements and pre-set power arrangements (Palmer and Walasek 2016; Patel et al. 2017; Perry et al. 2018). From a spatial discourse perspective, few would argue for space as being neutral, since all spaces are affected by their production and usage, and are loaded with symbolic and representational connotations. However, the essence here is

the importance of having access to a space which is unaligned and not controlled by any of the participating organisations. To explore the role of physical space for TD research and learning further, we suggest three spatial criteria in terms of *location, capacity of allowing* and *dignity*.

Since TD research endeavours often have a university base, the drive to locate the particular spaces at the university campus tends to be unquestioned. However, from our platform experiences, we have seen the significance of a spatial separation of TD spaces from the university and from public agency corridors. In the Kisumu platform (KLIP) the difficulties of bringing a multi-stakeholder group to the university space promoted the renting of a specific KLIP House as starting point for further research collaborations (Palmer and Walasek 2016). With this distinctive and separate space, the different stakeholder groups could, for the first time, come together in a conducive, neutral space fostered by the KLIP Trust as an independent organisation that had its finances ringfenced from those of its member institutions. Participants representing the public organisations in the GOLIP also affirmed the importance of a space located at the edge of the university campus, but separate from institutional facilities, as it also promotes easy access from their everyday engagements in the city. In current discussions regarding how universities need to embrace transdisciplinary education and research, and thereby change pedagogical structures, the matter of spaces for TD learning and their locations should be taken into account, since institutional spaces are always marked by ownership and inherent power. In this way, the *locations* of TD spaces could transform both the physical fringes of the university and of participating organisations, creating new intermediate spaces. Such locations are consequently important for first order effects by enhancing collaborations and the fostering of networks. Furthermore, they facilitate understanding these fringe spaces for collaborative knowledge production as an 'edge-scape' of new societal spaces, possibly having third order effects.

The second aspect of space is its *capacity of allowing* by supporting different kinds of actions and ways of coming together due to its form, size, light and other designed features. This classic architectural knowledge shows how consideration should be given to how spaces should be designed to support different processes of collaborative openness and learning to function as scaffolding. By constructing such spaces, we are not only supporting the TD process itself, but also creating a new physical and institutional space aimed at another type of societal practice, e.g. collaborative learning and investigation. Following Lefebvre's (1991) notion of space as a product of social processes, we can imagine these spaces as new urban contributions, destined for a practice of collaborative urban making. This is clearly an effect of the third order, as it contributes to changing our understanding of the urban and its common and commoned spaces.

Finally, we add yet another dimension to this kind of space, beyond unalignment and allowing features, namely *dignity*. Too often, experimental activities and alternative approaches are directed to leftover spaces. Even so, these spaces need to speak with respect to their users, who then attribute dignity, without being exclusive, and with respect to the research process itself. 'Dignified spaces' is also a concept used to upgrade informal settlements, to make a difference in everyday life, about which spaces are negotiable and which are not, i.e. spaces that cater for everyone's needs and hence belong to everyone (Southworth 2002). For the same inclusive reason, it makes sense to

bring spatial dignity into the TD and CP research. In so doing, perhaps here too we can discern a scalar shift from the specific and particular 'safe' spaces to collective or generic spaces, which open up those spaces to new institutional and urban possibilities as outlined above – and which also target second and potentially third order effects.

This 'system' of different infrastructural components needs to coincide and work jointly in a flexible manner. In total, these three components would ultimately function as an interwoven supportive structure across the different Mistra Urban Futures' platforms, allowing clear entry points for practitioners and academics alike. In any long-term research program, such components would benefit from being promoted as an institutional and structured systemic intervention under constant reflexive development.

Discussion: connecting evaluative and enabling infrastructures

Transdisciplinary co-production needs more than a quality and management system for evaluation. It also needs additional and complementary infrastructure components that can strengthen the capacities for transdisciplinary co-production in terms of learning, collaboration and enhancing relationships across organisations. We believe such additional components contribute to achieving structural changes in the participating organisations, through which they directly promote societal change. Too often, these types of support structures – training, facilitation and learning spaces – are seen as arbitrary, as they are often found on the margin of university institutional programmes. For numerous practical and economic reasons, their importance is questioned and often overlooked because they often don't fit well in either educational institutions or practice-based organisations since they bridge the boundaries of both. Implementing these components, however, not only supports the quality and outputs of TD co-production, but also contributes to new institutional arrangements that, in themselves, produce third order effects. Here, we explore further how the enabling infrastructure contributes to societal impact and then point towards how the QME framework itself enhances second and third order effects.

Training not only increases the participants' practical and theoretical knowledge and awareness of TD research and its challenges and (in our case) concepts of urban sustainability and justice, it also enhances respect and trust among participants who undertake training programmes jointly. The type of training undertaken at the Centre builds on mutual learning across disciplines; across practice and the academy; across silos in organisations; and across students and teachers. Such a training programme itself becomes a 'safe space' for difficult discussions that would otherwise almost certainly not take place at all, and can lift real-life problems and examples of related TD processes and results into a practice of consciously supported reflection. What we have experienced from the Open Research School and the KTP programme is an increased sensibility among the participants of how learning across boundaries takes place, on individual as well as group levels and how this learning affects the behaviours of participants as they return to home organisations, searching for implementation of new practices (Patel et al. 2015). This points towards TD learning as generating second order effects, which contribute to changes in organisations, and new policies and decisions.

One important conclusion raised by project participants and Centre consortium partner co-ordinators was that constructive discussions among participating organisations

on institutional conditions for change as well as their connections to the wider politics of sustainability issues are crucial (Hansson and Polk 2017). Facilitation enables both academic and practitioner TD researchers to create more qualitative and effective processes, where tensions among participants due to different kind of knowledges, previous experience or not of TD co-production, different mandates and institutional norms, can be balanced through the legitimacy inherent in the role of the professional facilitator as an outsider (Jordan 2014). In this sense, facilitation also creates a certain 'safeness', where participants have to worry less about misdirected group dynamics, representation, negotiations and group decisions, and can focus on the creativity of the collaborative research process and how to steer outcomes towards societal change beyond the confines of the specific project – in other words, towards generating and recognising second and third order effects.

Finally, the role of a TD spaces to enable TD co-production processes to enhance their effectiveness, and to create legitimacy and give credibility to the research itself. Our experiences suggest that the characteristics of the physical space and a formative evaluation process are mutually supporting and that spatial conditions of 'certainty' also support the effectiveness of the other infrastructural components of training and facilitation.

These infrastructural components need to be flexible in relation to the processes of knowledge integration and collective decision making and action. They also need to interact with each other. By intersecting and supporting one another, they constitute both an enabling and assuring infrastructure. This infrastructure assures participants in TD co-produced research enough 'safety', at both practical and cognitive levels, to trust the experimental and uncertain characteristics of the process, and thereby also carrying out a formative reflection while engaged in the research processes itself. With its intermediate position, such infrastructure ultimately provides entry points for both academia and practice to enter into the knowledge sphere of the other and to be able to dwell there from a position of entitlement and belonging. This is a step towards making spaces for new and shared imaginations. As such this infrastructure itself facilitates opportunities for second and third order effects.

To capture these elements of learning might be the central and unique feature of evaluation within TD co-produced research. The reflexive component of our QME framework, the Realising Just Cities evaluation, has been our instrument for this purpose. The full results of this work across the Centre are yet to be seen in the forthcoming publication. As learning is also central for transformation to sustainability, we argue that a learning component of an evaluation framework needs an infrastructural support in practice to reach impact of second and third order effects. We have tried to demonstrate some of the steps in distinguishing project-specific first order effects from broader institutional or second order effects in practice, through the support of an enabling infrastructure, which in itself contributes to second and third order impact. Table 1 brings together the qualities of these infrastructures, in terms of supporting learning, with the results from the Mistra Urban Futures' infrastructural examples and their possible effects. However, the ability to distinguish between the second and third order effects of Mistra Urban Futures research requires more time since these effects are inevitably lagged by several years.

Table 1 Qualities and results of an enabling infrastructure for TD projects

	Training	Facilitation	Spatial support
1. Qualities	• Reflexive and open interactions • Iterative and adaptive processes • Theoretical and practical foundations • Relationship based on trust	• Support structures tailored to specific project needs • Open access to 'neutral' and highly skilled facilitators	• Unaligned location • Capacity of allowing through design • Dignity through negotiation of spaces
2. Examples	• Open Research School • Open Method Seminars • Knowledge Transfer Programme	• Facilitation advice • Tailored facilitation of processes • Facilitation workshops	• A specific KLIP House • 'Edge-scapes' of new societal spaces
3. Results	• Critical mass of trained and skilled individuals in TD CP approaches	• Multiple experiences of facilitated processes with high levels of interactions, efficiency and trust • Systematic documentation of facilitation experiences	• Scalar shift from the specific and particular 'safe' spaces to collective or generic spaces, open to new institutional and urban possibilities
4. Effects	• Bridge from first order to second order effects	• Bridge from first to second order effects	• Possible effects of third order

Conclusion

Drawing on the experience of Mistra Urban Futures, we have demonstrated how reflective and reflexive evaluation of TD co-production within sustainable urban development enhances the ability of these methodologies to promote more sustainable societies. One essential requirement is to develop an evaluative framework that is able to capture the various results of co-production research as they appear in different formats, within different timeframes and with different degrees of 'distance' from a project. This requires careful distinction between the internal or first order effects, intermediate or second order effects within the participating organisations, and the wider societal or third order effects. But even after doing so, we continue to ask: What counts as capacity building, when do we need to enhance learning, when do we see new imaginations emerging? Evaluations of co-produced research therefore need to be formative, to respond continuously to the processes and project participants, for them to capture the experiences of capacity building, learning and changed perspectives from current and previous project processes. This formative work needs to be situated in and supported by an enabling infrastructure of training, facilitation and within an accessible, allowing and dignified space which, in turn, invites actors to enter into a 'space of necessity'. It is within such spaces that a formative evaluation can be productive and not only measure, but possibly also contribute to processes of both incremental and more radical societal changes towards sustainability. As the examples from Mistra Urban Futures have demonstrated, this is because formative evaluation processes help build trust and confidence, deepen the sense of shared ownership, and improve the actual practice and outcomes.

Availability of data and materials
The only data are the reports and papers cited, which are published and/or on Mistra Urban Futures website: https://www.mistraurbanfutures.org

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Chalmers University of Technology, Gothenburg, Sweden. [2]University of Gothenburg, Gothenburg, Sweden. [3]Royal Holloway, University of London, Surrey, UK.

References
Belcher BM, et al. Defining and assessing research quality in a transdisciplinary context. Res Eval. 2016;25(1):1–17.
Blackstock KL, Kelly GJ, Horsey BL. Developing and applying a framework to evaluate participatory research for sustainability. Ecol Econ. 2007;60:726–42.
Bornmann L. What is societal impact of research and how can it be assessed? A literature survey. J Am Soc Inf Sci Technol. 2013;64(2):217–33.
Bremer S, Meisch S. Co-production in climate change research: reviewing different perspectives. WIREs Climate Change. 2017; 8:e482. https://doi.org/10.1002/wcc.482.
Durose C, Richardson L, editors. Designing public policy for co-production. Bristol: Policy Press; 2016.
Hansson S, Polk M. Evaluation of knowledge co-production for sustainable urban development. In: Part I: experiences from project leaders and participants at Gothenburg local interaction platform 2012–2015; 2017. www.mistraurbanfutures.org/sites/mistraurbanfutures.org/files/hansson-polk-wp-2017-2.pdf.
Hansson S, Polk M. Assessing the impact of transdisciplinary research: the usefulness of relevance, credibility, and legitimacy for understanding the link between process and impact. Res Eval. 2018;27(2):132–144.
Harlen W, James M. Assessment and learning: differences and relationships between formative and summative assessment. Assessment Educ Principles Policy Pract. 2006;4(3):365–79. https://doi.org/10.1080/0969594970040304.
Hellström T. Formative evaluation at a transdisciplinary research center. In: Polk M, editor. Co-producing knowledge for sustainable cities: joining forces for change. London; New York: Routledge; 2015.
Hmelo-Silver CE, Duncan RG, Chinn CA. Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and Clark (2006). Educational Psychologist. 2007;42(2):99–107.
Institute for Methods Innovation. Mistra urban futures impact analysis report (preliminary findings); 2019. www.mistraurbanfutures.org/en/research-agenda/impact-studies.
Jahn T, Keil F. An actor-specific guideline for quality assurance in transdisciplinary research. Futures. 2015;65:195–208.
Jordan T. Deliberative methods for complex issues: a typology of functions that may need scaffolding. Group Facilitation: Re Appl J. 2014;13:50–71.
Kabisch N. Transformation of urban brownfields through co-creation: the multi-functional Lene-Voigt Park in Leipzig as a case in point. Urban Transformations. 2019;1(2). https://doi.org/10.1186/s42854-019-0002-6.
Koier E, Horlings E. How accurately does output reflect the nature and Design of Transdisciplinary Research Programmes? Res Eval. 2015;24(1):37–40.
Lang DJ, Wiek A, Bergmann M, Stauffacher M, Martens P, Moll P, et al. Transdisciplinary research in sustainability science: practice, principles and challenges. Sustain Sci. 2012;7(Suppl. 1):25–43. https://doi.org/10.1007/s11625-011-0149-x.
Lefebvre H. The production of space. Malden: Blackwell; 1991.
Luederitz C, et al. Learning through evaluation – a tentative evaluative scheme for sustainability transition experiments. J Clean Prod. 2017;169:61–76. https://doi.org/10.1016/j.jclepro.2016.09.005.

Lux A, Schäfer M, Bergmann M, Jahn T, Marg O, Nagy E, et al. Societal effects of transdisciplinary sustainability research – how can they be strengthened during the research process? Environ Sci Policy. 2019;101:183–91.

Mitlin D. With and beyond the state — co-production as a route to political influence, power and transformation for grassroots organizations. Environment and Urbanization. 2008;20:339. https://doi.org/10.1177/0956247808096117.

O'Malley E, Farrell DM, Suiter J. Does talking matter? A quasi-experiment assessing the impact of deliberation and information on opinion change. Int Polit Sci Rev. 2019. https://doi.org/10.1177/0192512118824459.

Palmer H, Walasek H, editors. Co-production in action: towards realising just cities. Gothenburg: Mistra Urban Futures; 2016.

Patel Z, Greyling S, Parnell S, Pirie G. Co-producing urban knowledge: experimenting with alternatives to 'best practice' for Cape Town, South Africa. Int Dev Plan Rev. 2015;37(2):187–203.

Patel Z, Greyling S, Simon D, Arfvidsson H, Moodley N, Primo N, et al. Local responses to global sustainability agendas: learning from experimenting with the urban sustainable development goal in Cape Tow*n*. Sustain Sci. 2017;12(5):785–97. https://doi.org/10.1007/s11625-017-0500-y.

Perry B, May T. Urban knowledge exchange: devilish dichotomies and active intermediation. Int J Knowl Based Dev. 2010;1(1/2):6–24.

Perry B, Patel Z, Norén Bretzer Y, Polk M. Organising for coproduction: local interaction platforms for urban sustainability. Politics Govern. 2018;6(1):189–198. https://doi.org/10.17645/pag.vXiX.XXX.

Pohl C. What is progress in transdisciplinary research? Futures. 2011;43:618–26.

Polk M. Achieving the promise of transdisciplinarity: a critical exploration of the relationship between transdisciplinary research and societal problem solving. New Directions Sustain Sci. 2014;9:439–51. https://doi.org/10.1007/s11625-014-0247-7.

Polk M. Transdisciplinary co-production: designing and testing a transdisciplinary research framework for societal problem solving. Futures. 2015;65:110–22.

Popa F, Guillermin M, Dedeurwaerdere T. Practicing the science of sustainability: the challenges of transdisciplinarity in a developing world context. Futures. 2015;65:45–56.

Reed MG, Abernathy P. Facilitating co-production of transdisciplinary knowledge for sustainability: working with Canadian biosphere reserve practitioner. Soc Nat. 2016;31(1):39–56. https://doi.org/10.1080/08941920.2017.1383545.

Roux D, Stirzaker R, Breen C, Lefroy E, Cresswell H. Framework for participative reflection on the accomplishments of transdisciplinary research programs. Environ Sci Policy. 2010;13:733–41.

Schipper K, Silvestri G, Wittmayer JM, Isoke JB, Kulabako R. Handle with care: navigating the pluriformity of power to enable actionable knowledge for transitions in informal settlements in the global south. Urban Transformations. 2019;1(4):1–11. https://doi.org/10.1186/s42854-019-0004-4.

Simon D, editor. Rethinking sustainable cities: accessible, green and fair. Bristol: Policy Press; 2016.

Smit W, Durakovic E, Sitas R, Johansson M, Haysom G, Dymitrow M, et al. Replicating projects for comparative research: Mistra urban futures' experiences with comparative work on knowledge exchange, food and transport. In: Simon D, Palmer H, Riise J, editors. Comparative urban research from theory to practice. Bristol: Policy Press; 2020.

Southworth B. Urban design in action: the City of Cape Town's dignified places Programme – implementation of new public spaces towards integration and urban regeneration in South Africa. Urban Des Int. 2002;8:119–33.

Spaapen J, Van Drooge L. Introducing "productive interactions" in social impact assessment. Res Eval. 2011;20(3):211–8.

Stokols D. Training the next generation of Transdisciplinarians. In: O'Rourke MO, Crowley S, Eigenbrode SD, Wulfhorst JD, editors. *Enhancing communication & collaboration in interdisciplinary research*. Los Angeles: Sage publications; 2014.

Stone A. What is missing in the metaphor of scaffolding? In: Forman EA, Minick N, Stone CA, editors. *Contexts for learning: sociocultural dynamics in children's development*. Oxford: Oxford University Press; 1993.

Termer CJAM, Dewulf A. A small wins framework to overcome the evaluation paradox of governing wicked problems. Polic Soc. 2019;38(2):298–314.

Walter AI, Helgenberger S, Wiek A, Scholz RW. Measuring societal effects of transdisciplinary research projects: design and application of an evaluation method. Eval Program Plann. 2007;30:325–38.

Watson V. Co-production and collaboration in planning – the difference. Plan Theory Pract. 2014;15(1):62–76.

Wiek A, Talwar S, O'Shea M, Robinson J. Toward a methodological scheme for capturing societal effects of participatory sustainability research. Res Eval. 2014;23:117–32.

Williams S. Evaluating societal effects of transdisciplinary co-production processes. Final report to Mistra urban futures, October; 2017. www.mistraurbanfutures.org/sites/mistraurbanfutures.org/files/Evaluating-societal-effects-Steve-Williams.pdf.

Williams S, Robinson J. Measuring sustainability: an evaluation framework for sustainability transition experiments. Environ Sci Policy. 2020;103:58–66.

Wood D, Bruner JS, Ross G. The role of tutoring in problem solving. J Child Psychol Psychiatr. 1976;17:89–100.

Publisher's Note