

D 8.1 Framework for comparative assessment

N **NEW** **HORIZON**

Grant Agreement No.	741402
Project Start Date	May 1 st , 2017
Duration of project	48 months
Deliverable Number	D.8.1
Deliverable Leader	UvA
Version	2.0
Submission Date	December 21 st , 2018
Authors	Anne Loeber, Joshua Cohen



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 741402. The opinions expressed in this document reflect only the author's view and in no way reflect the European Commission's opinions. The European Commission is not responsible for any use that may be made of the information it contains.

1. Executive summary	2
2. Introduction.....	3
2.1 Responsibility in research and innovation	3
2.2 RRI as an incentive for transformative change on system level.....	4
2.3 NewHoRRizon as an effort at inducing system transformative change.....	6
2.4 Evaluating NewHoRRizon: WP8.....	8
3. Evaluating NewHoRRizon	11
3.1 Towards an evaluation approach that fits the object of research	11
3.1.1 An apt approach to evaluating NewHoRRizon is <i>responsive</i>	13
3.1.2 An apt approach to evaluation <i>captures moments of disruption</i>	14
3.2 Narrative approach to evaluation	14
3.2.1 The ‘why’ of narrative evaluation	15
3.2.2 The ‘how’ of narrative evaluation	17
4. A framework for comparative assessment and narrative analysis	19
4.1 Framework components in a relational perspective.....	19
4.2 Data collection.....	20
4.3 Analysis and comparative assessment	23
4.4 Inventory of data production requirement and time line.....	28
5. References	30
6. Annexes	36

1. Executive summary

The European Union's (EU) 8th framework programme, Horizon 2020 (H2020), which runs from 2013-2020, seeks to support R&I to expand the scientific and technological basis of the European economy and industry, while fostering broader benefits for society and tackling pressing societal challenges. Among the cross-cutting themes characterising the framework programme is the attention for responsibility in research and innovation, including a focus on strengthening the dialogue among researchers, innovators and the public at large about the (un)desirability of possible implications of R&I, and of gearing research and innovation efforts to addressing societal challenges. This aspiration for upholding European values of inclusiveness and democratic politics in the realm of research and innovation has been dubbed Responsible Research and Innovation (RRI, EC 2013a). The NewHoRRizon project is set up as a collaborative effort to investigate the extent to which and the way in which RRI is presently observable throughout H2020, and to help promote the uptake of RRI in practices related to the funding scheme. To achieve the latter, NewHoRRizon will organize 19 Social Labs, one per programme line in all sections of H2020. The Social Labs will present a forum for stakeholders to meet and discuss the meaning of 'responsibility' in the context of their programme line and to design pilot projects ('Pilot Actions') and implement these with the support of the Social Lab management team. The Social Labs seek to incite a learning-by-doing approach to elaborating the meaning of RRI in practice and to implement RRI-oriented strategic efforts, and at the same time present a range of opportunities for third parties to become acquainted with RRI.

This deliverable presents a framework for comparatively assessing and evaluating the Social Labs set up in NewHoRRizon, in terms of input, process and output, with the aim of generating narratives of action that promote the acceptance and use of RRI in R&I. Informing Work package (WP) 8 in the NewHoRRizon project ("Evaluation and overall narratives and storylines") are the questions:

- how do Social Labs spur reflection on responsibility in research and innovation within H2020 to inform and enable a context-specific concretization of the RRI concept in dedicated Pilot Actions?
- in which ways do Social Labs help create the conditions under which the RRI-oriented dynamics set in motion by NewHoRRizon and the Pilot Actions are plausibly deemed perpetuable?
- and how can these experiences and resulting insights in promoting the uptake of RRI be communicated in narrative form?

Given WP8's objective to 'generate and consolidate storylines and overall narratives on the NewHoRRizon RRI practices', the data to answer this question will be collected in narrative form. By combining the analysis of narrative reflections with the construction of narratives, WP8 is designed to contribute to achieving NewHoRRizon's expected impact as specified in the SwafS-09-2016 original call (EC, 2016, p.24), namely to develop "narratives' ... in relation to the various parts of Horizon 2020 [which] will allow RRI to be an integral part of a more coherent Work Programme in Horizon 2020. They will impact as well on the relevant stakeholder communities as well as in the European Research Area and beyond." The framework presented in this Deliverable consists of an elaboration of the narrative approach developed to evaluate NewHoRRizon's Social Labs and associated Pilot Actions, including its rationale and operationalization to guide empirical data gathering. This framework serves as a basis for subsequent WP8 evaluation and dissemination efforts, as well as method developments.

2. Introduction

NewHoRRizon's Work package (WP) 8 seeks to comparatively assess and evaluate the project's key activities – the Social Labs and Pilot Actions – in such a way that the evaluation process and its findings contribute to NewHoRRizon's (NH) objectives and aspired impact, that is, to promoting the uptake of elements of RRI in the EU's research funding scheme H2020 and beyond. In order to do so, the evaluation approach needs to be geared to the project's characteristics and objectives. This chapter describes the way in which the NewHoRRizon project and the Social Labs and associated Pilot Actions are conceptualized as objects of research in WP8 (namely as elements of concerted action aspiring to induce system transformative change) and how this conceptualisation presents a rationale for developing an evaluation approach set on stimulating learning and narrative construction.

Below, paragraph 2.1 briefly introduces the topic of Responsible Research and innovation (RRI). Paragraph 2.2 outlines how RRI presents an incentive for system transformative change. In 2.3, it is argued that NewHoRRizon is an effort at inducing such system transformative change, which serves as a rationale for the evaluation approach developed in WP8, which is outlined briefly in paragraph 2.4.

2.1 Responsibility in research and innovation

Research and innovation (R&I) serve as a cornerstone in the European strategy for smart, sustainable, and inclusive growth (EC 2010). Yet, research and innovation and resulting scientific and technological developments may have undesirable implications or unfavourable impacts on the environment and society. In order to direct scientific and technological developments towards addressing societal needs, and to help avoid or mitigate undesirable effects, research and innovation need to unfold 'responsibly.' As a major sponsor of research and innovation, the EU has developed an interest in exploring options for exerting influence on research and innovation from the perspective of responsibility via conditions for funding and associated incentive schemes. To capture this line of reasoning, the epithet 'Responsible Research and Innovation' (RRI) was embraced in the Science with and for Society (SwafS) Unit on EC research policy level.¹ The SwafS programme aspires to influence the research system towards becoming increasingly responsive to societal actors and towards specifically addressing societal challenges. To work towards this ambition, dedicated projects are funded that focus on (the assessment and development of) aspects of RRI.

One of these efforts is the NewHoRRizon project (European Commission Grant Agreement No 741402), a Coordination and Support Action (CSA) that seeks to promote the integration of RRI and Open Agenda approaches into national and international R&I funding. NewHoRRizon is among the most encompassing RRI-oriented projects. It aims at x-raying the entire H2020 funding scheme to observe current manifestations of RRI-uptake, and to promote a further integration of RRI in the framework and in the R&I efforts it funds. To that end, it takes the way that the EC operationalizes RRI as a point of departure.

¹ In addition to the EC, also on the national level, research councils seek to operationalize RRI by making public funding conditional on the elaboration of social and ethical considerations in research proposals (among them the Research Council of Norway and the Netherlands Organisation for Scientific Research). In addition, RRI is actively promoted by researchers associations of EC funding grantees, such as the Marie Skłodowska-Curie actions (MSCA) alumni association, as well as by NGOs such as EuroScience, a non-profit grassroots association of researchers in Europe, which initiates discussions on RRI (<https://www.euroscience.org>).

The EC defines RRI as structural (institutionalised) attention for the following aspects (“keys”) and dynamics in research and innovation:²

- **Gender equality**, including a gender balance in R&I teams, and accounting for gender dimensions of R&I projects;
- **Public engagement**, envisioned as a two-way communication and learning process to include in R&I industry and SME, policymakers, non-governmental organisations (NGOs), civil society organisations (CSOs), and citizens who would not normally interact with each other on matters of science and technology;
- **Science education and science literacy**, to nurture modes of scientific inquiry, curiosity, and creativity;
- **Open access and open science**, to make data and results of research more publicly accessible to improve R&I;
- **Ethics**, going beyond legal compliance and researcher integrity to include also reflection on questions of how R&I do and do not relate, or respond to, societal interests;
- **Governance**, to ensure effective, inclusive, and sustainable ways of co-designing agendas and activities to achieve the above and broader objectives of European R&I.
- More recently, the Commission has made additional commitments to Open Science, Open Innovation, and Open to the World (EC 2016a) as part of its continued prioritization of fostering alignment among science and society in R&I. The EC Open Agenda includes three dimensions:
 - **Open Innovation** — “co-creation” that unfolds across innovation ecosystems and requires knowledge exchange and innovation capacity of all actors involved, be they financial institutions, public authorities or citizens, businesses, or academia (EC 2016a, p.12).
 - **Open Science** — a concept of transformed scientific practice, wherein the foci of researcher activity shifts from “publishing as fast as possible” to “sharing knowledge as early as possible,” in manners that are accessible to as many parts of the innovation ecosystem as possible (EC 2016a, p. 34).
 - **Open to the World** — “Fostering international cooperation in research and innovation” to enable access to “the latest knowledge and the best talent worldwide, tackle global societal challenges more effectively, create business opportunities in new and emerging markets, and use science diplomacy as an influential instrument of external policy” (EC 2016a, p. 59).

2.2 RRI as an incentive for transformative change on system level

This operationalization adorns the RRI concept with a level of concreteness that fits European science policy-making, and that helps connect various dynamics evolving on the interface between science and society. In addition to this reading of what RRI entails, yet there are other ways in which RRI is elaborated and interpreted, and made operational. In academia, RRI is explored extensively in a dedicated body of work that feeds into, and reflects, RRI-related societal developments and discourses, also those outside of the EU policy circles.

The RRI concept was originally coined in the context of academic work, in view of the notion of Technology Assessment (TA) and related concepts regarding the ethical, legal and social implications

² The following formulation is based on Bernstein et al., 2018.

(ELSI) of research and innovation (Von Schomberg, 2011, 2013; Owen, Macnaghten and Stilgoe, 2012). RRI in academic circles is considered to differ from its TA predecessors in terms of its ambitions, which are more encompassing than those of TA. The latter has been formalised as a range of (policy-oriented and/or technology-oriented) design instruments applied in concrete projects (Klüver et al., 2000; Joss and Bellucci, 2002; Loeber, 2004). RRI in contrast is defined as indicating a broader change, namely a change in R&I towards *a transparent and interactive process by which societal actors and innovators become mutually responsive to each other with a view to the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society)* (Zwart, Landeweerd, Van Rooij, 2014). This encompassing In addition to scale, the ambition captured in the RRI notion also differs from its TA / ELSI-predecessors as it “flips the main problem framing of earlier policies: the question is not what society wants to avoid, but which future we want to create” (Stilgoe et. al. 2013).³ This pushes the agenda of science governance beyond project level, to raise the question of how to organise a responsible development of science and technology that is responsive to societal needs and concerns *on a systemic level*. Therefore, in WP8, NewHoRRizon’s efforts are understood as, and assessed in terms of, efforts aiming at a *transformation* of the research and innovation system.

This tallies with the way in which RRI informs the SwafS Work Programme that underlies the funding of NewHoRRizon. There, RRI is characterised as a process of “better aligning research and innovation with the values, needs and expectations of society,” which requires “a regime shift” in the evolution of science and of its interactions with society (EC, 2016, p.6). This reading implies that the implementation of RRI amounts to a ‘transition’ of science and innovation as a socio-technical system, since “transitions are defined as shifts from one regime to another regime” (Geels, 2011, p.26). The word *regime* here denotes the semi-coherent set of rules and conventions that orient and coordinate the activities of the social groups via whose actions a system is reproduced.

The system transformative quality of the RRI notion has been elaborated as a challenge in its own terms in various ways. Owen et al. (2012, p. 756) for instance point at the concept’s implications for fundamentally reconsidering how responsibility is currently encapsulated in legal frameworks that assess responsibility in hindsight. This consequentialist perspective on responsibility is deeply problematic, the authors argue, for future-oriented innovation. RRI not only entails a reframing of what ‘responsibility’ implies in view of innovation but also requires the institutionalization of such a redefined understanding of the notion, which is value-based rather than rule-based, into standing legal frameworks. In a similar vein, Stilgoe et al. (2013, p.xiii) draw attention to the problematic relation between innovation and regulation: “If Responsible Innovation is to be viewed as a constructive endeavor, it must escape a predominant perception that [it] is about regulation – saying “no” to things. Indeed, if Responsible Innovation is to make a difference, it will be through questioning the separation between innovation and regulation.”

Both elaborations serve to illustrate that **RRI presents a challenge of institutional transformation**, as is acknowledged in the SwafS funding program. This understanding of what RRI entails – which runs the risk of being easily overlooked when taking the “keys and O’s” operationalization as the sole

³Furthermore, RRI presents a more encompassing label as it combines societal and ethical considerations with economic ambitions: “...the ambition of the European Union [is] to ensure that research and innovative ideas can be turned into products and services that create jobs and prosperity, as well as help preserve the environment and meet the societal needs of Europe and the world.” (Von Schomberg, 2013 p. 11)

focus in thinking about implementing RRI – puts NewHoRRizon’s objective of stimulating the uptake of RRI in relevant funding schemes in a broader perspective. It projects NewHoRRizon’s focus on the uptake of RRI as a collective action effort dedicated to supporting and inciting system transformation.

2.3 NewHoRRizon as an effort at inducing system transformative change

The way in which NewHoRRizon’s general objective is detailed speaks to this broader projection of what the project aspires. The idea of promoting ‘an increased use and acceptance of Responsible Research and Innovation (RRI) across H2020’ is elaborated to involve the “conceptual and operational foundation for integrating RRI into the next Framework Programme and in national R&I funding programmes across Europe and beyond” (objective §1.1.1; NH 2016: 3⁴). In addition to this substantive objective, several procedural objectives and guiding principles are elaborated that are considered instrumental in achieving the overall objective. Three process aspects can be identified that characterize these elaborations: an emphasis on context, a focus on learning, and attention for the relation between short-term project activities and longer-term impacts.

The **emphasis on context**, and on the context-specific elaboration of what RRI entails, first and foremost comes to the fore in the elaboration of NewHoRRizon’s core activity: “bringing together different stakeholders to co-create tailor-made actions within each section of H2020 that foster the uptake of RRI” (§1.1.2). This aim takes shape concretely in the organisation of so-called Social Labs, NewHoRRizon’s main activity. These Social Labs (of which there are 19, one for each ‘Programme Line’ in H2020) bring together relevant actors to jointly develop the ‘tailor-made’ actions mentioned (*i.e.* the Pilot Actions). While no limited set of criteria was specified at the onset of the NewHoRRizon project, it is clear from the Social Lab Manual (Griessler et al., 2019) that participants are not supposed to be selected as to their representativeness in numerical terms. Rather, their involvement is based on having ‘a stake’ in the issue under scrutiny: “NewHoRRizon social lab participants come from different stakeholder groups from different parts of Horizon2020 and beyond” (p.22) whereby stakeholders are defined as “societal actors that have an impact on the R&I process at hand or are impacted by the same process” (p.12). This specification suggests that Social Lab participants are selected on the basis of their “proximity to the issue” (Fox & Miller 1996, p.126), that is, of their practical understanding of what is at stake, an understanding that is acquired through a continuous and serious involvement in the relevant field.

It is this kind of *practical knowledge* and expertise that is valued in NewHoRRizon. The project wishes to combine state of the art knowledge on RRI via its consortium members responsible for the Social Labs, of whom some were engaged in previous RRI-oriented EC projects (see guiding principle §1.1.9.1), with insights in *de facto* RRI as perceived and implemented by Social Lab participants from the perspective of their own professional context, “recognizing the need for context specificity, variety, and concreteness” (§1.1.9.3). It is a relevant empirical question in view of evaluation *how in the design of Pilot Actions their ‘context-specificity’ is ensured, resulting in actions that are relevant and suitable for the professional practices and institutional settings in which (potential) stakeholders operate (“tailor-made”)*.

Attention for the situatedness of Pilot Actions and the practical knowledge that may inform their design is of great relevance, in other words, yet may also play out as a mixed blessing. If not

⁴ All mentions in this section of objectives and guidelines specified in NewHoRRizon share this reference.

problematized and subjected to critical review, it carries the risk that RRI actions inadvertently reproduce the very system that they are intended to help change. It is in this light that the second aspect of NewHoRRizon's process design comes into view: its **focus on learning**. 'Learning' as a verb has many meanings (Grin & Loeber, 2007). The most general meaning of learning as acquiring new knowledge certainly plays a role in NewHoRRizon: one of the aims is to share knowledge (§1.1.9.2) in order to familiarize H2020 R&I actors with the concept of RRI, and to help develop clarity in the terminology. Yet learning in terms of 'reflection-in-action' (Schön, 1983; Argyris & Schön, 1994; alternatively dubbed 'frame reflection'), that is, the capacity to reflect on and reframe the considerations that triggered someone to opt for a particular line of action over alternatives (Laws & Loeber, 2011), is arguably even more relevant in view of RRI. Such learning ("reflection in and on action", Schön, 1983) is crucial for preventing the risk indicated above. This is not a trivial issue. It is via recurrent daily practices incorporating and unobservantly expressing prevailing rules and conventions, that socio-technical systems tend to be relatively stable over time (Geels, 2004; Grin et al., 2010).

This insight from transition studies suggests that the envisaged (institutional) change towards an RRI-informed regime in the research and innovation system requires a conscious effort to explicate tacitly held conventions and make them the object of reflection. A *de-* routinization of standing practices enables critical reflection on what is considered feasible, desirable and appropriate (the rules of the game) under given circumstances, which then allows for a conscious *re-*routinization of new ideas into standing practices (Wilk, 2009; Loeber, 2015). The word 'effort' is in place here: it is often a tough call to become aware of what is usually taken for granted (Schön & Argyris, 1994). The downside of the aforementioned 'proximity to the issue' of participants is the lack of critical distance. Actors are by default not fully aware of their own practices and the rules and conventions that rein these and inform their actions. They operate from what Bourdieu calls "learned ignorance" (1977:19), aptly defined by Risseuw (1988:170) as "a mode of practical knowledge that excludes knowledge of its own principles."

The ambition to promote the uptake of RRI puts centre stage the *need to incite reflection on tacit principles and conventions* from the perspective of normative RRI-related concerns. Because of the afore-argued encompassing, transformative nature of the RRI, addressing it effectively implies a (re)negotiation of the 'rules of the game' (cp. Hajer, 2003; cp. Clark, 1985), and their subsequent institutionalization. This is imperative in NewHoRRizon as the project explicitly aspires to take the current situation as a point of departure, yet to "nudg[e] existing practices towards RRI" (§1.1.9.4), and to "embed [...] the promotion of RRI in existing institutions and routines." The literature suggests that the specific composition of the participant group, for instance, may help overcome this hurdle, as others can direct the attention to "what [we] ha[ve] worked to avoid seeing" (Schön, 1983, p. 283), but there may be many other techniques and methods, specifically relevant in view of RRI. Furthermore, the literature on learning by and large ignores the strategic aspects that come into play when seeking to re-routinize and institutionalize the newly developed 'rules of the game'. The reflections may well bring to light which structural barriers prevent well-intended actors to 'act otherwise' lest transaction costs in terms of career opportunities, reputation and the like are too high. It is therefore a relevant empirical question to investigate *how the Social Labs seek to incite processes of learning in terms of frame reflection on what is considered feasible and appropriate in relation to what is articulated as desirable in the light of RRI, and how they help participants to act practically and strategically on the resulting insights.*

Strategic choices are crucial for ensuring that what is discussed and designed in the Social Labs has an impact beyond the sphere of influence of those immediately involved in the Social Labs and in NewHoRRizon as such. This directs the attention to the third process aspect, the **attention for the relation between short-term project activities and longer-term impact**. The idea of a long-term perspective on change is inevitably at odds with the short life of concrete projects. Envisioned changes are not bound to take place overnight and will require protracted efforts. Yet the project specifies as an explicit objective that it will seek to “promote the integration of RRI in national R&I funding programs” (§1.1.7) and to “promote acceptance of RRI across H2020 and generate long-term impact” (§1.1.8).

These are challenging objectives both in view of practical time constraints implied in project-based work, and in view of the persistence of deeply rooted structures, like the kind that RRI-ambitions seek to change. In the learning-oriented literature the issue of impact is addressed as a question of how to expand the conditions under which learning takes place to contexts *outside* the project setting, to enable non-participants to “learn vicariously” (Guba & Lincoln, 1989) along similar lines as the participants in a project did. In view of system transformative aspirations, which requires strategic action in relation to learning as we have seen above, the question furthermore is phrased as one of how to “dynamically anchor” the output of a project (Loeber, 2003; Elzen et al., 2012). Anchoring refers to the process of actively embedding concrete results and other outputs of ambitious change projects in incumbent institutions, and of actively linking relevant actors and networks to ensure dedicated protracted interaction to help the results and institutional ‘novelties’ progress into becoming a ‘given’ in standing practices. On project level, the objective of NewHoRRizon to cast the results of the project in narratives (§1.1.2: “Transform the activities into narratives and widely disseminate them for use”) and disseminate them as such (in a “Guide to Good Practices for RRI”) presents a practical way forward in view of inciting vicarious learning on RRI among non-participants. This will be further elaborated in the next chapter. On the level of the Social Labs, it is a matter of empirical inquiry to observe *how in practice the Social Labs seek to ‘anchor’ Pilot Action designs by helping to create the conditions (embedding them in standing institutional settings, connecting networks) under which these can be protractedly pursued and/or implemented after the NH project has come to an end.*

2.4 Evaluating NewHoRRizon: WP8

Included in NewHoRRizon’s project design is a work package (WP8) to comparatively assess and evaluate the Social Labs in a way that will generate narratives of actions that promote the acceptance and use of RRI in R&I. More specifically the **objectives** are:

- a) To develop a framework for comparatively assessing and evaluating input, process, and outcome of the Social Labs
- b) To evaluate the Social Labs
- c) To generate and consolidate storylines and overall narratives on the NewHoRRizon RRI practices.
- d) To develop a Guide to Good Practices for RRI

These objectives have been elaborated in NewHoRRizon’s Description of Work (DoW; EC, 2017) in a way that specifies the evaluation’s role in helping to achieve NewHoRRizon’s objectives, both in terms of its substantive aim – the promotion of the uptake of RRI in research and innovation funding schemes, in particular H2020 – and of its procedural aims in terms of learning and communication.

WP8's objectives c) and d) are specified in the context of NewHoRRizon's substantive aim, to contribute to the promotion of the uptake of RRI via the construction and dissemination of narratives (§1.3.4.5 and §2.1): "NewHoRRizon will turn the successes and failures experienced in the pilot actions into 'storylines' and 'narratives' (Call H2020 SWAFS9-2016 [EC, 2016]) and will create a Guide to Good Practices for RRI (WP8)." Storylines are understood as condensed statements by which stakeholders convey complex narratives and make sense together (Hajer, 1995), so as to be able to discuss (RRI in) research and innovation practices that are bound to be widely diverse. Narrative reflections on past courses of events in WP8 will be made the object of joint reflection, on the basis of which the WP8 team will construct so-called "emplotted" narratives (Ricoeur, 1992), that is to say, narratives that synthesize multiple events into a meaningful story, weaving a thread between planned and unplanned actions and encounters, placing them in a particular temporal order to convey reflections on and conclusions from these incidents in a way that readers are able to follow and to make sense of (cp. Halsema, 2011).

As concerns NewHoRRizon's procedural aims as specified above, WP8 is assigned a role in enabling learning and continuous communication (§1.3.4.4): "Tools designed in ... WP8 – i.e. ... evaluation forms – would help to stimulate exchange and learning...". This role, WP8 takes up in the elaboration of its objectives a) and b), that is, in the development of a framework for comparatively assessing and evaluating input, process, and outcome of the Social Labs, and in the actual process of evaluation. The framework for assessment and evaluation presented in this Deliverable is the result of, and benefited greatly from prolonged and intensive exchanges with Consortium partners about RRI, Social Lab experiences and evaluation approaches. The deliberations with Consortium partners also informed the conceptualization of the NewHoRRizon as an interventionist project and helped delineate WP8's object of research.

As is elaborated above, WP8 perceives NewHoRRizon as an interventionist project seeking to incite system transformative dynamics informed by RRI. This conceptualisation informs how the Social Labs and the Pilot Actions are understood as **WP8's object of research**, in relation to NewHoRRizon's object of change. The Social Labs are conceived of as the project settings in which NewHoRRizon's core interventions, the Pilot Actions, are designed and developed, in order to achieve NewHoRRizon's objective of promoting the uptake of RRI in the H2020 research and innovation system.

Given WP8's objectives and considering the above elaboration of its object of evaluative research, the central questions guiding the assessment and evaluation effort in NewHoRRizon read:

- *How do Social Labs spur reflection on responsibility in research and innovation within H2020 to inform and enable a context-specific concretization of the RRI concept in dedicated Pilot Actions?*
- *In which ways do Social Labs help create the conditions under which the RRI-oriented dynamics set in motion by NewHoRRizon and the Pilot Actions are plausibly deemed perpetuable?*
- *And how can these experiences and resulting insights in promoting the uptake of RRI be communicated in narrative form?*

In order to be able to answer these questions, an approach to assessing and evaluating NewHoRRizon's Social Labs and Pilot Actions has been elaborated that fits the project's objectives

and WP8's specified role in achieving these. This learning-oriented approach to narrative evaluation is introduced in chapter 3. Chapter 4 presents the operationalization of the approach into a framework for comparative assessment and evaluation of NewHorRizon.

3. Evaluating NewHoRRizon

In this chapter the approach to evaluating NewHoRRizon as developed in WP8 is elaborated. There is a myriad of evaluation research strategies possible, and the question is which is the most apt to achieve the objectives of WP8 as specified in the previous chapter. WP8 heralds the basic insight from evaluation studies (see e.g. Hellstern, 1991) that an evaluation approach should fit the characteristics, objectives and context of the evaluand, *i.e.*, the project or policy that is being evaluated. This implies in the case of NewHoRRizon that its conceptualisation as a project aspiring at contributing to system transformative change, and the three questions that follow from it, inform the approach to evaluation to accommodate a context-specific perspective, to be oriented on learning and to be capable of including a systems perspective in view of longer-term impacts. Furthermore, the approach should enable a comparison between widely diverse Social Lab and Pilot Action efforts and their contexts.

Below, paragraph 3.1 discusses how the system transformative nature of NewHoRRizon presents a set of design criteria for developing a fitting evaluation approach. On this basis, in paragraph 3.2, WP8's narrative approach for evaluating NewHoRRizon is outlined.

3.1 Towards an evaluation approach that fits the object of research

Since NewHoRRizon's focus is on *developing* interventions (the Pilot Actions) in an *experimental setting* (the Social Labs) with long-term goals, its concrete interventionist goals are emergent. Therefore, WP8 focuses on process evaluation rather than on goal-attainment (impact) evaluation, and features a formative rather than a summative approach. Formative evaluation is explicitly learning-oriented, and as such fit NewHoRRizon's and WP8's objectives. There are many approaches possible by which to conduct a formative evaluation, and decisions have to be made about how to capture the way in which the project's implementation takes place in relation to its contexts, how to examine the range of activities and their plausible effects, and how to transfer insights from this monitoring back to those involved.

Given NewHoRRizon's central objective of contributing to a furthering of the uptake of RRI in the science and innovation system, understood as system change, a way forward in developing a suitable approach is to turn towards those approaches to evaluation that embrace system thinking (see e.g. Turnheim et al.'s (2015) approach to evaluating "sustainability pathways"; Anzoise & Sardo (2015) "Dynamic Evaluation", Urban, Hargreaves, Trochim's (2014) "Evolutionary Evaluation", and Taanman's (2012) "Transition Monitoring"). What binds these widely varying approaches is their **embracing of complexity theory** as a constitutive factor in developing evaluation practices, embracing and developing "complexity consistent methods" (Walton, 2014). This implies, first of all, a turning away from the traditional, in most policy circles still dominant linear approach to evaluation, which assumes that clear cause-effect relations can be observed between interventionist actions and societal effects, resolving possible attribution issues ('what effect is sorted by which interventionist (policy) action?') on the basis of programme theory: "The interest in 'complexity' seems to have centred on a growing recognition of the disjunction between the non-linearity and unpredictability of change processes and the protocols and procedures that govern [policy] ... interventions that assume otherwise" (Guijt et al., 2011, p.16-17). In contrast, complexity theory-inspired approaches to evaluation assume that processes of societal change evolve neither linear nor predetermined (Urban, Hargreaves, & Trochim, 2014), and that the conceptualization, design and

implementation of interventionist projects and plans themselves also should be regarded as on-going experiments that develop along, act upon and feed into their dynamic context.

A second feature of such non-linear approaches is furthermore their abandonment of the rational choice paradigm that dominates traditional policy-making and associated evaluation practices. This rationalistic behaviour logic assumes, as e.g. outlined in Fishbein and Ajzen's (1975) 'theory of reasoned action' (cp. Ajzen 1991), that behavioural intention is rational and *precedes* actual behaviour, informing options for action in the light of self-interested utility maximization, the formulation of which results from "interactions between an individual's attitude towards the behaviour in question, their beliefs about what others think about the behaviour – the subjective norm – and their perceived level of control over the behaviour, or perceived behavioural control" (Hargreaves, 2011, p.81). As Elisabeth Shove (2010) argues convincingly, the emphasis on choice in such line of reasoning a priori systematically and fundamentally overlooks the (limiting and enabling) conditions under which actors operate in practice, by externalising context. Shove argues (2010, p.1276) that the "invocation of complexity is theoretically bewildering" if it is not translated into a redefinition of the relation between the evaluand actors and context. To do so one has to let go, Shove posits, of the idea that it is possible to combine causal models of change with a systemic perspective. Instead, she embraces the 'practice turn' in the social sciences (Schatzki et al., 2001; Reckwitz, 2002) to develop an alternative to what she calls the ABC-model for understanding social change (in which the 'A' stands for attitude, the 'B' for behaviour, and the 'C' for choice) and the role of dedicated policy efforts therein. Since externalising context is of no use in thinking about action aiming at inspiring system transformative change, as we have argued in the previous chapter, we follow Shove in arguing that NewHoRRizon is in need of a **practice-oriented approach to evaluation**. Such an orientation on practice helps shed a light on what was defined previously in terms of a "regime", namely the agglomerate of (codified or tacit) rules and conventions, that rein in specific chains of settings and networks (a "nexus of practices", Shove, 2003; Shove & Walker, 2007) and that not only informs action from the perspective of what is considered desirable and appropriate in a certain context, but that also co-evolves with those actions (cp. Hoffman & Loeber, 2016). These structures do not predetermine action. Rather, in their actions, actors creatively interpret these structures, to a certain extent drawing on them consciously if they are considered instrumental or unavoidable. In doing so, they may well re-interpret them and through their actions create a possibility of changing them. Such structures hence are both the medium and the outcome of action (cp. Giddens, 1984).

Not only need the contexts in which RRI-oriented actions are designed, and on which they are bound to sort an effect, be put centre stage in an evaluation of NewHoRRizon. Ideally, as was argued in chapter 3, such an evaluation is also helpful in problematizing those context-specific 'regimes'. The evaluation could play a role in stimulating Social Lab teams and participants to avoid taking the very structures that are the object of change as an un-reflected boundary to their planning and designing of Pilot Actions. So-called '**reflexive evaluation**' (or 'reflexive monitoring', Van Mierlo et al., 2010; cp. Loeber, 2003; 2007) as a specific type of formative evaluation, seeks to do so explicitly: "A reflexive perspective in evaluation ... means evaluating programmes or initiatives that seek to contribute to system change in order to deal with complex problems, while [itself] supporting the change process ... Basically, it ... encourage[s] groups of actors to reflect on the rules and relations underlying current practices in order to induce institutional change" (Arkesteijn et al., 2015, p.108). The quintessence of reflexive evaluation, in other words, is that it not only seeks to support the learning processes that

take place in a project or program, but also acknowledges that those who are at the wheel – the change agents; here Social lab management teams and participants, that is, Pilot Action protagonists – are themselves embedded professionally in the system they seek to challenge. By helping to explicit tacit assumptions in a world where usually ‘the fish don’t talk about the water’ (Risseuw, 1988), the evaluation of NewHoRRizon may contribute to Pilot Action design that challenges those structures that may a priori be considered ‘self-evident’, so as to move beyond picking the low-hanging fruit to promote the uptake of RRI in terms of furthering its institutionalisation⁵ (that is, its uptake in the regimes of rules and conventions dictating what is ‘normally’ desirable, feasible and appropriate).

These three considerations present the rationale in the elaboration of an evaluation approach for NewHoRRizon.

3.1.1 An apt approach to evaluating NewHoRRizon is responsive

To accommodate NewHoRRizon’s emergent character and to ensure sufficient focus on context from the perspective of Social Lab teams and participants, WP8 builds on the principles of responsive evaluation (Stake, 1974; Guba and Lincoln 1989). Key in responsive evaluation is that themes and issues that emerge from those involved in the evaluand serve as an organising principle in the organisation of an evaluation (Loeber 2010).⁶ Such an evaluation includes in its focus those issues (that is, ‘claims’ when positively assessed by the person who raises it; ‘concerns’ when negatively assess by the person who raises it; Guba & Lincoln, 1989) that the participants in the project, that is being evaluated, themselves bring to the table. That does however not preclude the possibility to outline the evaluation up-front, yet sufficient flexibility should be incorporated to ensure that it does justice to the experimental character of the evaluated project.

A responsive evaluation is, in other words, always to some extent a form of negotiated evaluation in which the viewpoints of various parties involved form the starting point in organizing the knowledge production process. In the evaluation approach developed in WP8, this comes to the fore in two ways. First, the evaluation activities are temporally geared to the project’s dynamics. The **efforts at gathering, analysing and feeding back evaluative data are made contingent on the timing of Social Lab events** (see this document, §4.2). Secondly, evaluation efforts are developed and implemented in close consultation with those involved. This brings along that the evaluation approach and design are developed in close interaction with NewHoRRizon’s Consortium members, reflective of and relevant for their experiences in shaping and executing the project, and that the evaluation of the Social Labs and Pilot Actions is made responsive to the lived experiences of Social Lab teams and their respective participants. The reason behind this choice is that “[e]valuation is the more relevant and actionable when it is actually engaged with those who are being evaluated and is able to generate processes of development and learning related to real problems, ... [which] allows participants to reflect on their experience, shared information and relational processes,” as Ivaldi et al. (2015, p.499, 502) put it. Therefore, **data gathered about the Social Lab and Pilot Action experiences is analysed inductively** (to ensure claims, concerns and issues of those involved get due

⁵ Institutions are here considered to refer to relatively coherent sets of rules and resources, whether or not grouped together as an organisational configuration.

⁶ This tallies with WP8’s elaboration in the NH DoW which specifies (EC 2017, p.31) that “[t]he framework for comparative assessment and evaluation of the Social Labs, their input, process, and output will be developed that builds on, but does not limit itself to, the way in which the RRI package in FP7 and H2020 has been made operational in terms of identified keys, [yet will be] complemented with ... the themes that emerged from stakeholders during the project.”

attention) and **findings are fed back to Social Lab teams and Pilot Action protagonists respectively** (see §4.3).

3.1.2 An apt approach to evaluation captures moments of disruption

RRI-oriented Pilot Actions are ideally ‘disruptive’ in regard to standing, by and large non-RRI-informed, research and innovation and associated funding practices. Disruption here refers to an instance of surprise or friction that helps foreground what usually remains tacit. Such a disruption, however gentle, may raise awareness of the ‘silent structures’ that influence what is considered desirable and appropriate to say or do. On such critical moments, some of the tacit understandings of the “logic of practicality” (Pouliot, 2008) that is, of what is shared between a person and context, become explicit and are made the object of reflection. Therefore, in regard to data gathering, **WP8 focuses specifically on** such (positively or negatively appreciated) moments of surprise or friction, that is, on those ‘critical moments’ when ‘things don’t go as expected’ **as occasions for reflection and learning** (see §4.3). Such moments that urge one to adjust plans and actions to new found circumstances, firstly, present occasions for *learning* (cp. Schön, 1983), that is, of a revision of earlier held (often tacit) beliefs and understandings. In an experimental setting such as NewHoRRizon’s Social Labs, these present relevant findings, namely (on an aggregated level) they amount to ‘lessons learned’. Secondly, those moments of friction or surprise are an indicator of where ‘the system works against’ the project’s objective of promoting a transformative change such as mainstreaming RRI. After all, if things go smoothly, it might mean that the objective is achieved already. But it might also mean that at first glance we cannot see where and how that is not the case and not likely to be the case, because we look at it ourselves from within that system. Only when we ‘bump into’ some obstacle (like a wall in the dark) we become aware of it being there, and the structuring elements that inform our (here: research and innovation) practices become objects of reflection (cp. Knorr Cetina, 2001). Disruptions, or ‘critical moments’ therefore can be highly instrumental as clues indicating where to direct efforts at system change, and deserve attention in the evaluation.

3.2 Narrative approach to evaluation

In order to meet these requirements of responsive (context-specific) and reflective (learning-oriented) evaluation, WP8 will adopt a narrative approach to evaluating NewHoRRizon, that is, the Social Labs and Pilot Actions. A narrative approach to evaluation offers the research the possibility, methodically, to bring what makes sense in a particular situation in relation to contextual dynamics in moments when an actor disengages from the force of habit.

The methodical choice for narrative evaluation builds on, and relates to, the attention for narratives that is outlined both in the Work programme underlying the NewHoRRizon project and in its objectives. The aforementioned Horizon 2020 Work Programme 2016 - 2017 regarding Science with and for Society (SwafS; EC, 2016) accords a central role to the notion of ‘narrative’ in relation to the concept of RRI. The approach to evaluation developed in WP8 is informed by a conceptualisation of this relation between narratives and the ambition of promoting RRI.

The Commission’s focus on producing narratives is understandable. Narrative knowledge is “the primary form by which human experience is made meaningful” (Gold et al., 2003, p. 53, cp. Fisher, 1987; Polkinghorne, 1997). According to Souto-Manning “Narrative is one of the most broadly employed ways of systematizing human experience. As human beings, we experience our worlds and live our lives by telling stories. It is through narratives that experiences are ordered and permeated with meaning” (Souto-Manning, 2014, p.162 cp. Bruner, 1990). A choice for narratives to

communicate findings from a CSA project such as NewHoRRizon is a choice for a context-rich presentation of insights gained, in a way that can itself, *because of* the narrative form, contribute to the attainment of the project's objectives, in this case, to a furthering of RRI-oriented change in research and innovation. A narrated form in disseminating insights contributes to the transferability of the experiences and to their relevance for interested third parties.

In addition, a narrative form of making sense of the development and achievements of the Social Lab and Pilot Actions has merit beyond the (ex-post) communication about the project. Narrative, and narrative evaluation, contributes - it can be argued on the basis of the above elaboration of (hurdles to) system transformation - to the process of inducing RRI-oriented change that is aspired in the Social Labs and with the Pilot Actions as such. It is through narrative that, as Wagenaar and Cook posit, actors produce and negotiate practices: "by telling *stories* about their and other people's actions within the various elements of their community ... actors make sense, to themselves and others, of their actions" (2003: 156, drawing on Cook and Brown, 1999; Wagenaar, 1997). This is through this performative dimension of narrative (Souto-Manning, 2014, p.162) that the dialectical relation between actors and their surrounding world, as discussed above, takes shape (cp. Lave, 1988):

Stories [or: narratives] are not merely representations of actions and consequences; they are also performative and generative. As a form of discourse, by telling stories, actors simultaneously shape, grasp and legitimate both their actions and the situations that give rise to them. ... [T]he meaning of a story in a given context emerges from an interplay between text and context that is actively managed by the actor / teller in interaction with the 'listener' in a given context. (Wagenaar and Noam Cook, 2003, p.156)

With the focus on narratives in WP8, we position the research in the burgeoning field of academic work on language-oriented evaluation. In contrast to a numerical orientation in evaluation, language-focus evaluations help transform an enumeration of heterogeneous (planned and unplanned) events and components into a story, connecting them through the introduction of a plot (resulting in "emplotted" narratives, Ricoeur, 1992). Emplotment can be understood as "the synthesis between heterogeneous elements. That means in the first place that the plot synthesizes multiple events or incidents into one completed story" (Halsema, 2011, p.114). The relevance of these events or incidents roots in their role in the narrative, as is the case of the other heterogeneous elements, such as "planned and unplanned actions, coincidental or intended encounters, interactions ranging from conflict to collaboration, persons that suffer from actions, and persons that commit them, etc." (ibid., p.114).

3.2.1 The 'why' of narrative evaluation

There are several reasons to choose for the collection and presentation of data as narratives. Narratives are capable of capturing the uniqueness and complexity of particular events, incidents and encounters in a specific time and place (Abma, 1999). Because they help bring a variety of heterogeneous elements together in a way that makes sense, narratives have the "potential to produce nuanced and in-depth understanding of the impact of initiatives which quantitative or universal/standardized indicators lack" (Constant and Roberts, 2017, p.1, drawing on Bornmann, 2013). Since adoption of the RRI package in R&I and funding in specific sections, countries, and sectors is bound to vary, an evaluation of RRI-oriented efforts has to be sensitive to context (national and cultural variety), disciplinary differences, and societal challenges addressed. A narrative approach to evaluation implies that the contextual character of the experimentation is included in

the data that is gathered, analysed and conveyed: “Narratives offer the possibility of unique, context-based evaluations through time-oriented structures ... revealing how changes occur and evolve” (Constant & Roberts, 2017, p.4).

The latter is, moreover, of particular relevance in the light of communicating the insights from projects that focus on complexity in system transformation. Communication in, and about, such projects present a specific challenge because of the interconnectedness of the various aspects and elements at issue, and the emergent character of both project and its unfolding object of research (McGowan et al., 2014). Narratives in such project form a way not only to communicate about project activities in their context-in-flux, but also help *connect* such initiatives in spite of their very diverse forms and sizes, which are tailored to their particular context and issue (Wittmayer et al., 2015).

In other words, a choice for a narrative approach to evaluation potentially contributes to NewHoRRizon’s objectives in three ways. First, it contributes to the **transferability** of NewHoRRizon’s findings and insights as narratives enable others, who were not present, to not only read about the findings but to experience ‘what it’s like’ (Hermans, 2011; see below) to live through what is related in the narrative. Thus, a focus on narratives enables vicarious learning (Guba & Lincoln, 1989; see above), a form of **learning beyond NewHoRRizon’s project setting**, enabling non-participants to learn about the project, and how to set up Pilot Actions and their effects. Because of their contextually empirically rich nature, they enable the reader to *de*-contextualize lessons learned so as to *re*-contextualize them to fit the particularities of her own professional context, enabling the reader “to create their own meaning and find ... the intersection between the familiar and the unknown” (Bedford, 2001, p. 33 in Zhai and Dillon, 2014, p. 423; cp. Gherardi and Nicolini, 2000). By including sufficient contextual detail about what happened in a Social Lab and/or Pilot Action, the narratives that WP8 will reconstruct, which will be disseminated in a *Guide to Good Practices for RRI*, will thus enable non-participants to ‘learn vicariously’ about NewHoRRizon experiences and insights.

This mechanism is, secondly, an asset for NewHoRRizon as such as well. The aforementioned possibility of producing ‘unique, context-based evaluations through time-oriented structures’ (Constant & Roberts, 2017, p.4) enables **communication** and **learning between the various Social Labs**. The Social Labs are working in a total of 19 very different contexts, from the excellence-oriented bottom up Marie Skłodowska-Curie Actions (MSCA), to a programme aimed at solutions for sustainable food production (FOOD) to risk finance that is available for promising start-up businesses (RISK), to a programme line focused on nuclear fission and fusion research (EURATOM), working with all different kinds of people and Pilot Actions. Therefore, the evaluation should do right to this plurality and diversity, and simultaneously contribute to a mutual understanding of what binds these efforts together, so as to be able to draw inferences about ‘what works’ on an aggregate level. In this way the evaluation can contribute to achieving NewHoRRizon objective to: “Create, based on mutual experiential learning in the Social Labs, viable narratives of how to implement RRI in the programme lines.”

A third asset of narrative evaluation is its promise to spur explication and thus **reflection** and **learning within each of the Social Labs**, among and between a project’s participants and Social Lab team. A narrative evaluation asks project collaborators to relate in narrative form the events that they consider of relevance, and to focus thereby on some sort of disruption or disequilibrium into the actor’s usual business (cp. Herman, 2011). This focus is related to the attention for moments of

surprise and friction discussed above, as well as to the broader focus of NewHoRRizon to develop Pilot Actions that are disruptive to an extent that they trigger a de-routinization of standing R&I practices. If evaluative questions are formulated in a way that prompts reflection and explication, they can enable a process of learning-in-context on the part of the actor answering the evaluation questions as such.

3.2.2 The 'how' of narrative evaluation

For a narrative evaluation to function and perform in these ways, it needs to combine the *analysis of narratives* with *narrative analysis*. Narrative analysis is the process of **constructing emplotted narratives** from data. In NewHoRRizon this **data is collected in narrative form**, and subsequently analysed.⁷

Data collection is organised via a 'Template for Reporting and Reflection' (see §4.2), which is issued out to Social Lab teams six times during the NewHoRRizon project, and asks for information on Social Lab experiences in narrative form. To enable the eventual construction of well-substantiated narratives, WP8 asks Social Lab Teams that report on their experiences via the Template, to pay attention to what constitutes good narrative data. Four necessary elements of good narrative data can be distinguished (cf. Herman, 2011, p.2, p.7):

- (i) situatedness: narrative data are situated in, and must be interpreted in light of, a specific discursive context or occasion for story-telling [such as here, a Social Lab and the context in which it is situated];
- (ii) events sequencing: narrative data presents coherence to situations in flux as the focus is on a structured time-course of particular events;
- (iii) disruptive quality: the events that are presented speak of some kind of disruption of the usual, on-going, meaning making processes;
- (iv) transferability: the data convey to others who 'were not there' what *it is like* to live in the world-in-flux about which is reported.

The analysis of such narrative data results in "descriptions of themes that hold across ... stories or in taxonomies of types of stories, characters, or settings" (Polkinghorne, 1995, p.12) and thus produces generalized knowledge about the studied phenomena. As the analysis focuses on commonalities between ideas, actions and developments reported, this kind of knowledge production is illuminating yet "abstract and formal, and by necessity underplays the unique and particular aspects of each story" (Polkinghorne, 1995, p.15).

The focus on commonalities is in specific contradiction to the other element in the narrative evaluation mentioned above, namely the *construction of narratives*. Such narratives present each of the described phenomena (Social Labs and Pilot Actions) in a unique and context-specific story. This kind of narratives are produced by the researcher / evaluator (here: the WP8 team) in consultation with those involved, by tying descriptions of events and happenings together to "... synthesize or configure these by means of a *plot* into a story" (idem, p.12, italics added). Put differently, instead of

⁷ To understand the difference between the two, the concepts of paradigmatic and narrative cognition (Bruner, 1985, p.11; Polkinghorne, 1995, p.8-11) are of help. Polkinghorne (1995) uses paradigmatic cognition for the capacity of classification, identifying a particular instance as belonging to a specific category or concept. This capacity helps create order in experiences "by seeing individual things as belonging to a category" (1995, p.10; cp. Schön's 1971 elaboration of creating meaning via comparison, in terms of 'seeing-as'). This order can be created deductively, by checking experience against pre-existing categories or concepts, or inductively, by examining "data items for common themes and ideas" (Polkinghorne, 1995). Narrative cognition, on the other hand, "is specifically directed to understanding human action [...] Human action is the outcome of the interaction of a person's previous learning and experiences, present-situated presses [*sic!*], and proposed goals and purposes" (idem, p.11).

a collection of abstract themes, the resulting insights are presented in a coherent account of a lived experience because a plot is added. The evaluator should “develop or discover a plot that displays linkage[s] among the data elements as parts of an unfolding temporal development” (idem). **A plot** thus relates events and actions to context in a temporal and causal manner: “[it links] a prior choice or happening to a later effect. The significance and contribution of particular happenings and actions are not finally evident until the *denouement*⁸ of the episode. Events which might have appeared insignificant at the time may turn out to have been a crucial occurrence affecting the outcome” (Polkinghorne, 1995, p.8). The plot, in other words, is the lifeblood of a narrative and is something that sets it apart from other pieces of text such as a regular report.

Information for plot construction can come from various sources – in addition to the narrative data provided by those involved in the evaluand, also from document analysis etc. – that need to be “integrated and interpreted by an emplotted narrative” (idem, p.12). This cannot simply be any configuration; it “must fit with the data while at the same time bring [...] an order and meaningfulness [to it] that is not apparent in the data themselves” (idem, p.16-18). Elements that are redundant to the ‘*denouement*’ are not included in the resulting narrative (a process called *narrative smoothing*; Spence, 1986). The result of such a narrative construction process is “an explanation that is retrospective, having linked past events together to account for how a final outcome might have come about” (idem, p.16).

Two quality criteria should be applied in the process of narrative construction and ‘smoothing’. First, “the explanation needs to satisfy the subjective needs of the reader ... to understand how the occurrence could have come about. [...]. The evaluation of the story has a pragmatic dimension in the sense that its value depends on its capacity to provide the reader with insight and understanding” (idem, p.19-20). Second, the process needs to be attentive to the “accuracy of the data and the plausibility of the plot” (idem, p.20). The latter implies that the construction conveys coherence among the situated, contextual, and particular elements and thus has explanatory power (cp. Connelly & Clandinin, 1990). Polkinghorne specifies further that in order to enhance the communicative capacity of the constructed narratives, they should be rich in contextual detail yet followed by a commentary analysis that discusses differences and similarities on the basis of a comparison between the narratives.

It is this rich contextual detail that allows the narratives to serve as a means to enable parties who were not part of NewHoRRizon to engage with its resulting insights, and to learn – as it was formulated above – vicariously. Because of the contextually empirically rich nature, they provide the reader with the opportunity to *de*-contextualize lessons learned so as to *re*-contextualize these to fit the particularities of their own professional context.

The narrative approach to evaluating NewHoRRizon that is elaborated in this chapter fits NewHoRRizon’s objectives and characteristics, and illuminates how WP8 can answer the central research questions phrased in chapter 2. It does not as such give substantive directions as to how to navigate the ‘sea of stories’ conceivable in the context of NewHoRRizon. In chapter 4 a framework for comparative assessment is presented that directs data collection and comparison in WP8.

⁸ ‘Denouement’ stems from the French verb for ‘untying’, and refers to the drawing of a conclusion to the narrative.

4. A framework for comparative assessment and narrative analysis

Transformative change is facilitated not only by action, but also by the development of modes of knowledge production that help create the conditions for the co-evolution of new ways of thinking and doing (Loorbach, 2014; Switzer, 2019). Chapter 3 argued that the responsive, and practice-oriented narrative evaluation is an appropriate approach for evaluating NewHoRRizon. In addition to the “appropriateness of methods,” NewHoRRizon’s DoW directs attention EC, 2017, p.106) to the relevance of the “comparability of results.” Considering the size and diversity of H2020 and associated Social Labs, this is a challenging remit. This chapter present a framework for comparative assessment that picks up this challenge in a manner that fits the approach to evaluation discussed in the previous chapter.

In paragraph 4.1, NewHoRRizon is described as the object of evaluation from a relational perspective to fit the evaluation’s rationale outlined in chapter 2. Thereupon, the process of data collection (paragraph 4.2) and data analysis and comparative assessment (paragraph 4.3) developed in WP8 are described. In the final paragraph, 4.4, the implications for data production and collaboration of fellow Consortium members (Social lab teams and Work package leaders) are listed.

4.1 Framework components in a relational perspective

As was elaborated in chapter 2, WP8 views NewHoRRizon as an interventionist project that aims at supporting and inciting system transformative dynamics informed by RRI. To reach this objective, Social Labs are developed as ‘reflexive arrangements’ (see below), that is, as project settings in which NewHoRRizon’s core interventions, the Pilot Actions, are designed and developed. The Pilot Actions are interventionist actions aimed at achieving NewHoRRizon’s objective of promoting the uptake of RRI in the H2020 research and innovation system. The three-tiered construction of NewHoRRizon can be depicted as a dynamic process in which input, process and output can be distinguished as follows:

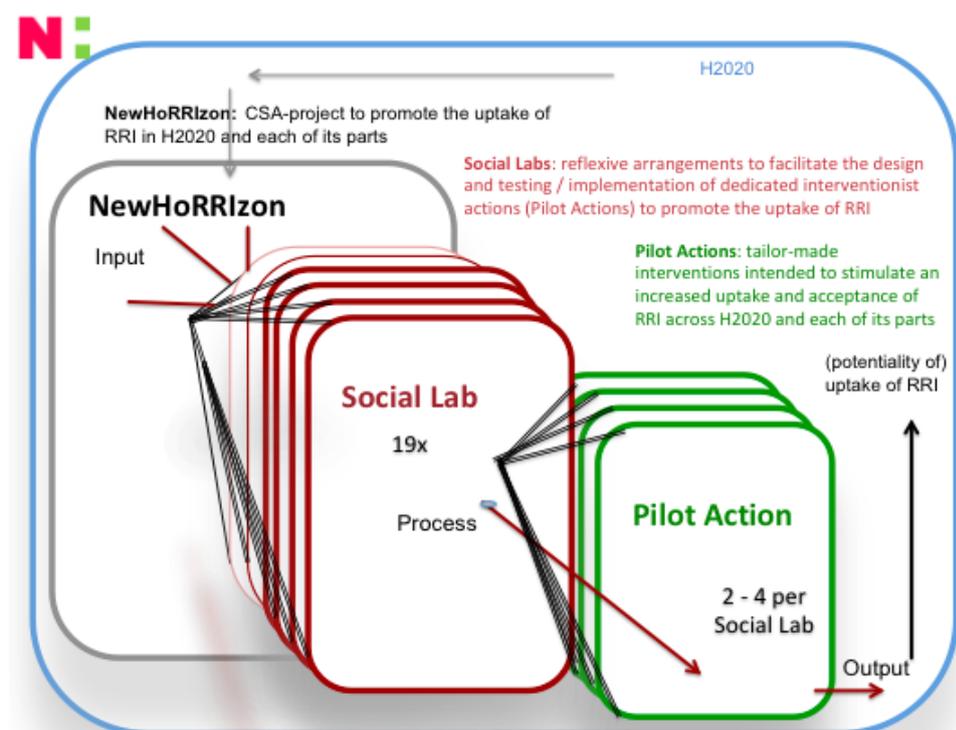


Fig. 1 NewHoRRizon as an object of evaluation

A depiction of an interventionist process in terms of input, process and output inevitably brings along the connotation of inserting dedicated action in an otherwise static system. That image, it may be clear from the previous chapters, is not what is intended here. Rather, the context in which NewHoRRizon as an interventionist project unfolds is taken to be constantly in-flux, as the actors of which it is comprised exercise agency in using resources to achieve their ends and to act in interaction with others. The context in which they operate is equally dynamic. While structures, orders and institutions cannot act or mechanistically determine the conduct of agents, as Arts and Van Tatenhove put it (2004, p.351), they co-evolve with the actions of actors that draw on and interpret these structures. Those actions become innovative when these actors cease to take these structures, and their own relationship with incumbent 'regimes' of rules and conventions (and associated material artifacts) for granted. When that happens, there is a de-routinization (Giddens, 1984), which may trigger creativity (Hoffman & Loeber, 2016) in designing new types of actions and associated ideas on what is appropriate and wise to do. An interventionist project interferes in these dynamics by strengthening or stirring dynamics causing a further de-routinization and re-routinization, while relating to ongoing processes. Such a practice-oriented, relational perspective in evaluation directs the evaluators' gaze to the constitutive and transformative potential of interventionist action as it *co-evolves* with the very structures it seeks to transform.

With such a relational, dynamic understanding in mind, the respective components of NewHoRRizon can be conceptualized as follows:

Input in the dynamic processes set in motion in NewHoRRizon includes *i)* the background knowledge on R&I system dynamics, in particular of H2020, of the stakeholders selected per Social Lab (solicited first in a range of interviews and subsequently in a series of three Workshops and in interactions between these events); and *ii)* information gathered via document analysis in the initial diagnostic stage of Social Lab preparation.

Process in NewHoRRizon is understood to comprise the activities undertaken in the Social Labs, including *i)* the reflections on, and deliberations of the initial diagnosis per H2020 Programme Line, *ii)* the actual designing, development, refining and (possibly) implementation and/or testing of the Pilot Actions that are informed by the diagnosis; and *iii)* the co-construction of narratives on both Social Lab developments and Pilot Action design and testing/implementation.

Output of NewHoRRizon is understood to include *i)* the actual Pilot Actions (which may take different shapes and forms: a design, event, artefact, etc.) as well as *ii)* a collection of narratives and storylines on both Social Lab methods and experiences, and on a selection of Pilot Actions, together with *iii)* a range of efforts that can be argued to increase the plausibility of procuring future system transformative change, that is, dynamic anchoring (Loeber, 2003; Elzen et al., 2011), linking networks, and ripple effects (Trickett & Beehler, 2015).

See below for a further specification of each of these components.

4.2 Data collection

Data on the various components of NewHoRRizon is gathered in close cooperation with Consortium members of all Work packages, with the aid of specifically designed data collection formats. The methods, with which the data is gathered, vary according to their characteristics and the purpose of data collection. An overview of the means of data collection per aspect of the framework is presented in table 1:

	Focus	Data on	Means of collection
Input ⁹	Awareness of and institutionalisation of RRI per Programme Line	Level to which RRI is institutionalised in each H2020 Programme Line	Diagnosis; Analysis of policy documents, Work programmes, calls, proposal template and evaluation guidelines
	Background knowledge of RRI and application of RRI in Programme Line practices	Level of RRI in research and/or innovations practices in a Programme Line	Diagnosis; Interviews with Programme Line stakeholders
Process ¹⁰	Social Lab design to fit the Programme Line context (diagnosis); reflection and learning on RRI – and the construction of narratives thereon	Participant selection and Workshops design; participants' perceptions on the extent to which the Social Lab is informative and relevant to their professional context	Reflection and Reporting Templates Moments I – VI Parts I (WP8) & II (WP7); Post-Workshop Questionnaires
	Design and development of Pilot Actions – and the construction of narratives thereon	Critical moments in the Pilot Action design process; Pilot Action designs and their underlying assumptions and motivations	Reflection and reporting Templates Moments IV, V and VI - Parts I (WP8) & II (WP7); Post-Workshop Questionnaires; Tool for eliciting Theories of Change per Pilot Action; Interviews with Social Lab managers pre-WS3; Learning Histories exercise
Output	Results of the Pilot Action process; final descriptions in WP7 Pilot Action database	Pilot Actions (design, implemented event, artefact etc.)	Reflection and reporting Template Moments V + VI
	Narratives and storylines per Social Lab / selected Pilot Actions in the <i>Guide to good practices for RRI</i>	Insights in Pilot Action design and development as a means to promote the uptake of RRI, and in the organisation of Social Labs as a setting to do so	Reflection and Reporting Template Moments I- VI; Q-methodology
	Anchoring efforts and ripple effects	Activities related to the Social Labs other than Pilot Actions	Reflection and Reporting Template Moment IV & VI; Moments I – IV Part 3 (WP9); Case studies and associated interviews; Post-Workshops Questionnaire

Table 1. WP8 – foci in data collection

Input

Data on the **level to which RRI is institutionalised in each H2020 Programme Line**, and on the level that RRI is integrated **in research and/or innovation practices** in the Programme Lines is gathered in the diagnoses that served as input in the Social Lab / Pilot Action design process. This data is gathered and reported on outside of WP8 by the Social Lab teams and made available via

⁹ Data gathered and reported on in WP2-5

¹⁰ Data on Social Lab design and Pilot Action design is provided by Social Lab teams, and gathered in a cooperative effort between WP7, 8 and 9.

Deliverables D2.1, D3.1, D4.1 and D5.1 (Diagnoses reports). These reports build on document analysis and interviewing. Documents included Programme line policy documents, Work Programmes, Call texts, Proposal templates and Evaluation forms and guidelines, plus relevant additional documents ranging from Commission's policy documents to individual research project documentation. Semi-structured interviews, using an interview protocol, of about 45 to 90 minutes were conducted with various stakeholders per Programme Line by Social Lab teams.

Process

Data on the **Social Labs and Pilot Action design and development** is provided by Social Lab teams, and collected in cooperation with WP7 and 9, particularly via a jointly designed 'Template for Reflection and Reporting' (see Annexes A, B and C), and a Pilot Action database developed in WP7. The Template is designed to collect information on Social Labs in a way that suits the data needs of WPs 7-9, while acknowledging the limited amount of time available to Social Lab teams. Teams are requested to fill in the template on six moments during the course of a Social Lab: prior to the first Workshop (Moment I, Part I only), shortly after the first Workshop (Moment II, Parts I-III), prior to the second Workshop (moment III, Part I only), shortly after the second Workshop (moment IV, Parts I-III), prior to Workshop 3 (moment V, Part I only) and after workshop 3 (moment VI, Parts I-III). Data on the experiences and perceptions of Social Lab participants is collected and reported on by the respective Social Lab teams (via a Tool for the reconstruction of Theories of Change, developed in WP8, see Annex D) and directly via a post-Workshop Questionnaire developed in WP8, one post-Workshop 2 and one post-Workshop 3. In light of the responsiveness of the evaluation, Social Lab teams are asked to provide input in reply to questions in the Reflection and reporting templates in narrative form that highlight so-called 'critical moments' (see below for an elaboration) and their way of dealing with these:

- Where there any critical moments encountered in the process preceding a particular moment of reflection (pre-/post-WS);
- Which choices did the teams make in relation to the critical developments described, and which assumptions informed these choices and actions? Thus, Social Lab teams are asked to reflect on their choices and underlying assumptions
- What were the results/consequences of the aforementioned choices and action for the discussions in the Social Lab and for the experimentation with RRI? This information sheds a light on possible enablers and barriers for making RRI operational in the context in which the Social Lab is situated.

Output

Results of the **Pilot Action** development process, that is, **descriptions** of their definite form (design, implementation, artefact etc.) are collected in the WP7 database for Pilot Actions. A selection of these, together with their development process and associated storylines, will be described in narrative form in the **Guide to good practices for RRI**. Reflections and reports via Template Moments V and VI, the post-Workshop Questionnaires as well as information collected in WP9 and additional case study research including interviews with selected Pilot Action protagonists and Social Lab team members will serve as a basis for describing **anchoring efforts and ripple effects** in a selection of case studies.

4.3 Analysis and comparative assessment

The combination of NewHoRRizon's project size and the diversity of H2020, and the argued relevance of context-specificity in evaluating the project make comparative assessment of the Social Labs a challenging task. Comparison in the analysis of the Social Labs and of the pilot Actions is achieved in WP8 as follows.

In order to be able to make a comparative assessment of the **Social Labs**, which are per definition experimental project designs that take shape in view of the characteristics of the H2020 Pillars and Programme Lines on which they reflect, they need to be conceptualised as diverse instances of empirical class of phenomena. In WP8, therefore, the Social Labs are understood as specific forms of 'reflexive arrangements' (Loeber and Vermeulen, 2016) that facilitate and support the design and implementation of actions dedicated to furthering system transformative change. A reflexive arrangement seeks to *i*) incite reflection on (aspects of) prevailing regimes relevant to the professional practices of the project (here: Social Lab) participants, to *ii*) develop options for action that potentially challenge standing practices and associated regimes (here: Pilot Actions); and to *iii*) identify leverage points for change on the level of networks and institutions, and to (coordinate and) act on these, in order to 'anchor' the dynamics set in motion by the Social Lab and/or respective Pilot Actions to enhance the potentiality of change, along the lines set, in the longer run (cf. Loeber, 2010).

How the Social Labs in NewHoRRizon give shape to these functions varies according to the characteristics of the Programme Line they focus on, as well as the knowledge and experiences that Social Lab teams have with project design and facilitation, in relation to the guidelines set in the Social Lab manual (Griessler et al., 2019). A comparison between (a selection of) the Social Labs will be made in terms of the relation between selected methods and other design choices on the one hand, and the results in terms of learning effects, deducible from the consecutive reformulations of Pilot Actions in the WP7 database, the changes and results reported via the Reflection and Reporting Template Moments IV-VI, and the self-reported learning effects as shared by Social Lab participants via the Questionnaire. Furthermore, an assessment of a selection of Social Labs will be made in the form of case descriptions that will explore the relation between Social Lab methods and design on the one hand, and learning effects on the other, also in view of the criteria for RRI as developed by Wickson & Carew (2014), where apt.¹¹ In this way, WP8 will comparatively analyse whether and how the Social Labs stimulate reflection on the dominant regime so as to incite experiences of de-routinization, and enable a mobilization of knowledge, skills and resources to design Pilot Actions that may trigger system transformative dynamics in a learning-by-doing manner. This analysis will provide an answer to WP8's first central research question:

- How do Social Labs spur reflection on responsibility in research and innovation within H2020 to inform and enable a context-specific concretization of the RRI concept in dedicated Pilot Actions?

Because of this focus, as outlined in chapters 2 and 3, the analysis of the data provided by Social Lab teams on their work-in-progress will be initially analysed inductively in a qualitative content analysis

¹¹ Wickson & Carew (2014) present quality criteria and an evaluative rubric of performance indicators for assessing RRI in the research. Because they are developed in view of RRI implementation efforts in research, they combine a process orientation with a substantive focus on a research environment which is not applicable in all of NewHoRRizon's Social Labs, and which ignores the interventionist (learning-oriented and impact-oriented) focus that informs NewHoRRizon's Pilot Actions.

to categorise and thematise what is observed as quintessential in developing Pilot Actions that can contribute to promoting the uptake of RRI. To focus the analysis to that end, Social lab teams are to focus, in their reflection and reporting on recent events in their Social Lab efforts in terms of ‘critical moments’. Such moments (potentially) involve an awareness of a mismatch between what was expected and what did in fact occur, and thus enable one to become aware of considerations and understandings of the situation-at-hand that usually remain tacit, which may, in turn, shed a light on the ‘logic of practicality’ and appropriateness that reins in that particular situation. By bringing out how they assessed and handled that situation, practical, ‘tailor-made’ ways of dealing with barriers and surprises in RRI-oriented actions are brought to the fore.

As described in chapter 3, the data collected and analysed via the Reflection and Reporting Template, together with information gathered via the Questionnaire, is used for narrative construction (see §3.2.2). The process of narrative construction will provide an answer to WP8’s third research question:

- How can these experiences and resulting insights in promoting the uptake of RRI be communicated in narrative form?

The analytic process of narrative construction will be practically organised as follows: Initial narratives revolving around inductively selected themes will be constructed in a processes of ‘narrative smoothing’ and plot development, and then be fed back to Social Lab managers and/or other relevant team members before Workshop 3, to be interviewed to ‘test’ the developing narratives via a first ‘respondent check’, and to elicit additional information where necessary. A second respondent check, for validation and for a further elaboration and fine-tuning of the pre-final narratives includes Social Lab participants and will take place in Workshop 3. Guided by the Social Lab Manager, participants will be asked to provide commentary, insights, and lessons gathered in view of the tentative narrative, in a way that reflects the Learning Histories (Kleiner & Roth, 1997) approach.¹² The objective is to make sure that all involved actors can “pose questions about assumptions and implications [ingrained in the narrative], and raise "undiscussable" issues that hover just below the surface” (Kleiner & Roth, 1997). This will bring out participants’ insights and lessons learned in the Social Lab and Pilot Action design process. Instruction on the concrete tenets of this particular part of the evaluation and assessment process will be supplied by WP8. The results help fine-tune the narratives related to the implementation of RRI within particular contexts, as well as construct ‘meta’-narratives on critical moments, common themes and issues that more generally relate to the implementation of RRI in the European research and innovation system.

Finally, the narratives will be analysed by WP8 to condense them into storylines. Storylines present narratives in ‘shorthand’ and help transform an idiosyncratic narrative into a shared story, by producing a ‘communicative miracle’ (Hajer 1995, p. 46) that is, by enabling actors to make sense together and to understand one another even in practices which are fundamentally diverse and which are characterised by discursive diversity. By looking at storylines as condensed versions of narratives that enable actors, despite the great variation in modes of speech, to understand one another, they can play a role in disclosing the narratives as constructed in consultation with Social Lab teams and participants to a wider audience. To assess which storylines communicate adequately what NewHoRRizon’s main findings in narrative form are, an initial set developed by the WP8 team

¹² The Learning History-approach developed by Kleiner & Roth (1997) gives participants in a shared experience the possibility to reflect on a description of a course of events, to add missing information, insights gained and lessons learnt, so as to refine the developing narrative.

will be reduced to a short-list that selected NewHoRRizon stakeholders considered adequate and relevant, with the help of Q-methodology. Q-method is useful in testing the resonance of theses and claims with specific target audiences as it allows for the systematic study of “subjectivity, a person’s viewpoints, opinion, beliefs, attitude, and the like” (Van Exel & De Graaf, 2005, p.1; Brown 1993).

An analysis of the **Pilot Actions** will provide an answer to WP8’s second central research question:

- in which ways do Social Labs help create the conditions under which the RRI-oriented dynamics set in motion by NewHoRRizon and the Pilot Actions are plausibly deemed perpetuable?

Since these are the actual interventions produced by NewHoRRizon, the question of the project’s impact centres on the Pilot Actions. Given the *ex-durante* nature of WP8’s evaluation, it cannot provide insights in the Pilot Actions’ ex-post impact in terms of actual RRI-uptake in H2020 and beyond. WP8 can however shed a light on the plausibility of producing the desired change. Such an assessment builds on theory-based evaluation (see e.g. Rogers, 2008). In the evaluation of policy, the relevant theory in such analytic work is the “policy theory” that informs concrete policy action, that is, the “total of causal, final and normative assumptions underlying a course of action” (Hoogerwerf, 1990; Grin and Van de Graaf, 1996). In case of dedicated interventionist action by actors other than policy makers, as is the case in NewHoRRizon, the phrase ‘action theory’ or ‘theory of change’ is better applicable. In WP8, we opt for the latter: in order to be able to understand the potential of the pilot Actions in promoting the uptake of RRI, the evaluation will focus on the Theory of Change that undergirds each of the Pilot Actions. That implies that data will be gathered on the (often tacit) assumptions, values and arguments by which actors make sense of the problematic situation in their respective Programme Line, their ideas on a solution for the defined problem issue (the actual Pilot Action design) and the strategic and normative considerations that inform the design. Given the idiosyncratic designs, which are ‘tailor-made’ to fit a specific (aspect of) a H2020 Programme Line, these Theories of Change are bound to be widely diverse. In order to grant the various Pilot Actions a minimum of comparability, WP8 will organize this amalgamation of assumptions and arguments under the heading ‘**Theory of Change per Pilot Action**’ in the format developed by Fischer (1995) in view of policy evaluation, for organising the arguments informing policy action.¹³

WP8 to that end developed a **tool that Social Lab teams are advised to use to help Social Lab participants explicate and organise their thoughts and ideas informing the Pilot Action** they work on (see Annex D), for instance during a Social lab workshop. The format explained in the tool also serves as a heuristic in organising Pilot Action descriptions in the WP7 database, elaborating four types of characteristics that define a Pilot Action. Building on Fischer’s framework for the systematic evaluation of policy action, the format for systematizing Pilot Action descriptions comprises of a set of 4 questions, that each ask after a different aspect of Pilot Action design and justification (see table 2; see Box 1 for a practical elaboration of a Pilot Action description, by way of illustration).

¹³ Fischer distinguishes between four categories of such arguments, with each a different type of evaluative questions, on the basis of their epistemological status

1. What does the Pilot Action look like? What kind of interventions / actions does it entail?
2. Why is that a useful thing to do? Which PROBLEM does the Pilot Action address? Which OBJECTIVES does it seek to achieve?
3. Why is that an effective design: Which STRATEGIC considerations inspire the design? (e.g. which actor groups / networks does the Pilot Action target and/or connect? Why? Which actor groups are involved in developing the Pilot Action, and why?)
4. Why is that a good thing to do: Which NORMATIVE considerations inspire the design? (e.g. motives regarding ‘responsibility’ in R & I; which of the 6 keys of RRI / 3O’s in the longer run)

Table 2: Heuristic for mapping the Theory of Change per Pilot Action

An asset of organising information on the ‘what and why’ of the Pilot Actions in this way is that it allows for a systematic assessment and comparison of what is an otherwise widely heterogenetic body of work. Pilot Actions can be related to one another in terms of each type of argumentation (e.g. categorizing them per RRI Key, or per actor group that seek to address, as is the case in NewHoRRizon 2018 Technical Report), and the plausibility of contributing to the uptake of RRI in H2020 and beyond can be assessed in terms of appropriate criteria. In table 3, a list of appropriate assessment criteria is presented.

	Level of argumentation	Pilot Action design and justification	Perspectives by which to assess the design
1.	Evaluation of solutions	Assessment of the Pilot Action	<u>Social Lab participants’ own</u> suggestion / formulation of indicators : which standards, criteria (and, if possible indicators) do the Pilot Action protagonists themselves consider relevant for assessing – in due time – the impact of their design?
2.	Definition of the problem(s)	Practice-based assessment of RRI barriers and enablers	Kuzma & Roberts (2018) categorization of barriers to RRI: barriers with respect to fundamental philosophical differences that actors and stakeholder hold (micro-level); within organizational structures of innovation systems (meso-level); and relating to the larger political, economic, cultural, social contexts (macro-level)

3.	Strategic background theories	Selection of targeted aggregation level, networks and actor (roles)	<p>Using MORRI distinction between levels of aggregation: EU (global) level; national level; regional level; institutional level; individual level¹⁴</p> <p>Using Res-Agora ‘responsibility navigator’s <u>categorization</u>¹⁵ of relevant actor groups: a) Actors and institutions directly concerned with the governance and development of research and innovation agendas: Funding institutions; Universities; Industry and companies conducting research; International organisations; Policy-makers</p> <p>b) Actors not involved in decision making on (R)RI: individual researcher (applicants), individual researchers (grantees), CSOs members / representatives, international organisations representatives; industry associations; citizens/users/consumers</p>
4.	Normative orientation	RRI interpretation	<p><u>EC’s 6 keys / 3 O’s ‘plus’</u>: Gender equality; Public engagement; Science education and science literacy; Open access and open science; Ethics; Governance + Open Innovation (“co-creation” / exchanging knowledge and innovation capacity); Open Science (“sharing knowledge as early as possible”); Open to the World (“Fostering international cooperation in R & I”)</p>

Table 3 – Overview of assessment criteria per Pilot Action characteristic

The approach to analysis of Pilot Actions in relation to their potential towards promoting the uptake of RRI in view of their various underlying assumptions, values and arguments allows the researcher/evaluator to take on board the considerations of those involved themselves in the evaluation. This approach tallies with the observations by the Expert Group on Policy Indicators for Responsible Research and Innovation (Spaapen/Strand, 2015) that indicators are context-dependent, and are therefore in need of a situational explanation and narrative elaboration (KNAW et al., 2014). By enabling those involved, that is, in the case of NewHoRRizon, the Pilot Action protagonists to themselves formulate suitable criteria for assessing the potential added value of their pilot Action, they are invited to make explicit what their line of argumentation is, *and* they can ensure that their actions are properly addressed and assessed (either in a self-assessment, or in an ex-post analysis by a third party). By supporting their Pilot Action plan with such data (e.g. relating to the Pilot Action’s strategic value for promoting RRI, or the resulting artefact or event’s ‘relevance for society’), they

¹⁴ See <http://morri-project.eu>

¹⁵ See <http://responsibility-navigator.eu/navigator/>

turn the narrative of their Pilot Action process and design into a ‘strong story’, that may be convincing to third parties interested in ‘learning vicariously’ yet in need of proof of the relevance of what they read.

Box 1 “RRI Manifesto” developed in SL2.3 on promoting the uptake of RRI in MSCA.

A practical initial elaboration of this Pilot Action in terms of Table 2 would read:

[[design](#)] This Pilot Action is about designing a cartoon-based information clip about why RRI can be informative in thinking about one's research proposal. [[problem / objectives](#)] The cartoon will help make early-career researchers familiar with RRI and its relevance in an engaging and concise manner. Currently, early-career researchers are often not aware of the existence and meaning of the RRI-concept, and if they do, they often take it to be a mere tick-box exercise, that is of relevance only when providing required information on the social validations of one' proposal. [[strategic considerations](#)] The plan is to make the cartoon-based film clip available via the EC Participant Portal to all those who consider submitting a proposal for a Marie Curie grant. [[normative considerations](#)] By doing so, the Pilot Action endorses RRI notions to become integrated in the very practice of research. The underlying ambition is to help further a responsible attitude among researchers in considering societal concerns and relevance in developing their academic work.

An additional asset for the WP8 evaluation is that the explication of each pilot Action's Theory of Change provides a basis for observing learning effects sorted: changes that occur in the reasoning by Pilot Action protagonists about the design and normative and strategic justifications of their Pilot Action speak of critically reflection on the issue and suggest learning, in the sense as defined in chapter 2, as reflection-on-action took place. As a result, the approach developed in WP8 for a practice-oriented theory-based evaluation spurs a process of collective reflection within Social Lab events such as Workshops, which enables learning within the project, as well as helps produce the narratives that form an important output of NewHoRRizon. As opposed to a straightforward *best practices*-approach, this reflexive narrative approach leads to a transfer of knowledge that reaches beyond a copying by others of "lessons learned". The rich stories of narratively described Pilot Actions, ideally with numerical indicators that help them potentially develop into ‘strong stories’, can help disseminate the insights gained in NewHoRRizon to a wide audience in a ranges of fields.

4.4 Inventory of data production requirement and time line

WP8's narrative evaluation approach and the above specification of the process of data gathering and analysis produce a set of request for data production and other forms of cooperation to Social lab teams and other WP-leaders. Below, these are put together for easy reference:

- WP8 collects narrative reflections during the lifetime of the Social Labs (with the help of the Reflection and reporting template on Moments I, II, III, IV, V and VI) on the process of each Social Lab and on its results (i.e. Pilot Actions and the associated Theories of Change);
- WP8 provides a post-WS2 and post-WS3 questionnaire to elicit further feedback from participants;
- On the basis of the narrative reflections, WP8, with the help of Social Lab managers through online interviews/conversations, will reconstruct Social Lab-specific and Pilot Action-specific narratives;
- Preliminary results of this may be discussed in the second Cross-fertilization Workshop;

- Pre-final narratives will be fed back for Social Lab participants to reflect and comment on during WS3 (so-called 'Learning Histories');
- The Learning Histories following from this will be analysed and condensed into storylines by WP8;
- Q-methodology will be applied to select storylines by which to categorize and disclose narratives, which will be collected in a *Guide to good practices for RRI*.

Social Lab teams are asked to provide the following types of data:

- Before the 1st Workshop Social Lab managers and teams report on the process leading up to their first Workshop via the Reflection and Reporting template (version April 2018; Moment I: Part I of the template only). The template unites efforts of WP7/8 and 9 and the pace of reporting is made contingent on the timing of the Workshops.
- After the 1st Workshop Social Lab managers and teams report on the process leading up to their first Workshop via the Reflection and Reporting template (version April 2018: Moment II: Part I, II and III).
- Before the 2nd Workshop Social Lab managers and teams will report on the process in between Workshops via the Reflection and reporting template ([Version 18 December 2018; 1.6]; Moment III).
- During the 2nd Workshop, Social Lab managers and teams will also (ideally) include a session on reflection and evaluation for participants using the Tool for reconstructing the Theory of Change per Pilot Action provided by WP8 (Version 13 November 2018). (Instructions are provided in a Powerpoint on Narrative evaluation [version 18 January 2019]).
- After the 2nd Workshop Social Lab managers and teams will report on the Workshop via the Reflection and reporting template ([version 8 January 2019; 1.8]; Moment IV).
- After the 2nd Workshop Social Lab managers and teams send out an online questionnaire provided by WP8 to Workshop 2 participants to solicit their views and ideas [Version still under discussion].
- Before the 3rd Workshop Social Lab managers and teams will report on the piloting process (and pilots) via the Reflection and reporting template ([Version will be provided in due time]; Moment V).
- Before the 3rd Workshop Social Lab managers and teams will be contacted by WP8 to discuss the developing narratives.
- During the 3rd Workshop Social Lab managers and teams will conduct a session on reflection and evaluation with participants using a Learning History-approach on the basis of materials provided by WP8 [version will be provided in due time]. Instructions are provided in a Powerpoint on Narrative evaluation and Learning Histories [same].
- After the 3rd Workshop Social Lab managers and teams will report on the Workshop via the Reflection and reporting template ([Version will be provided in due time]; Moment VI).
- After the 3rd Workshop Social Lab managers and teams send out an online questionnaire provided by WP8 to participants to solicit their views and ideas.

5. References

- Abma, T.S.A. (Ed.) (1999). *Telling tales: On narrative and evaluation (Advances in program evaluation, Vol. 6: 1-27*. Greenwich, CT: JAI. Introduction. pp.1-27
- Ajzen I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50: 179–211.
- Anzoise, Valentina and Stefania Sardo (2015). Dynamic systems and the role of evaluation: The case of the Green Communities project. *Evaluation and Program Planning* 54: 162–172
- Argyris, Ch. & D. Schön (1994) *Theory in Practice. Increasing professional effectiveness*. San Francisco: Jossey-Bass.
- Arkesteijn, Marlèn Barbara van Mierlo, Cees Leeuwis (2015). The need for reflexive evaluation approaches in development cooperation. *Evaluation*, 21(1): 99-115
- Arts, B., & Van Tatenhove, J. (2004). Policy and power: A conceptual framework between the old and new policy idioms. *Policy Sciences*, 37: 339–356.
- Barthes, R. (1975 [1966]). An Introduction to the Structural Analysis of Narrative, in: R. Barthes and L. Dussit (eds.), *On Narrative and Narratives. New Literary History*, 6(2), 237-272.
- Bedford, L. (2001). Storytelling: The real work of museums. *Curator*, 44(1): 27–34.
- Benford, R.D. and Snow, D.A. (2000). Framing Processes and Social Movements: An Overview and Assessment. *Annual Review of Sociology*, 26, pp. 611-639
- Bernstein, M. et al. (2018) *Diagnosis: RRI in Excellent Science*. NewHorizon-project: EC 741402. Deliverable 2.1
- Blackstock, K. L., G. J. Kelly, and B. L. Horsey (2007). Developing and Applying a Framework to Evaluate Participatory Research for Sustainability. *Ecological Economics*, 60 (4): 726–742.
- Blok, V., Hoffmans, L., & Wubben, E. (2015). Stakeholder engagement for responsible innovation in the private sector: critical issues and management practices. *Journal of Chain and Network Science*, 15(2): 147-164.
- Bornmann, L. (2013). 'What is societal impact of research and how can it be assessed? a literature survey'. *Journal of the American Society for Information Science and Technology* 64 (2), pp. 217–233.
- Bourdieu, P. (1977). *Towards a Theory of Practice*. Oxford: Oxford University Press.
- Brockmeier, J. (2009). Reaching for Meaning: Human Agency and the Narrative Imagination. *Theory Psychology* 19: 213.
- Brown, S. R. (1993). A primer on Q methodology. *Operant subjectivity*, 16(3/4), 91-138.
- Bruner, J. (1987). Life as Narrative. *Social Research*, 54(1): 11-32.
- Clark, K. B. and W.J. Abernathy (1985). Innovation: mapping the winds of creative destruction. *Research Policy* 14: 3-22
- Cook, S. D. N., & Brown, J. S. (1999). Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing. *Organization Science*, 10: 381-400.
- Connelly, F.M. & Clandinin, D.J. (1990). Stories of experience and narrative inquiry. *Educational Researcher*, 19(4), 2-14.
- Constant, N. and Roberts, L. (2017). Narratives as a mode of research evaluation in citizen science : understanding broader science communication impacts. *Journal of science communication*, 16 (4). A03.
- Curtis, C. (2012). Transitioning to transiti-oriented development: The case of Perth, Western Australia. *Urban Policy and research*, 30(3): 275-292

- Davies, J.E. (2002). *Stories of Change. Narrative and Social Movements*. State University of New York Press, Albany.
- De Fina, A. and Georgakopoulou, A. (2008a) Analysing narratives as practices. *Qualitative Research*, 8(3), 379-387.
- De Fina, A. and Georgakopoulou, A. (2008b) Introduction: Narrative analysis in the shift from texts to practices. *Text & Talk*, 28(3), 275-281.
- Delgado, R. (1989). Storytelling for oppositionists: a plea for narrative. *Michigan Law Review*, 87(8): 2411-2441.
- European Commission (2010) Europe 2020: A strategy for smart, sustainable and inclusive growth. Brussels, COM(2010) 2020
- European Commission (2016) Open innovation, open science, open to the world: A vision for Europe. Available at: <http://bookshop.europa.eu/en/open-innovation-open-science-open-to-the-world-pbKI0416263/> Workprogram Swafs
- European Commission. (2017). Grant Agreement Number - 741402 – NewHoRRizon.
- Elias, N. (1984). *What is Sociology?* Columbia University Press.
- Elzen, B., B. van Mierlo, C. Leeuwis (2012). Anchoring of Innovations: Assessing Dutch Efforts to harvest energy from glasshouses. *Environmental Innovation and Societal Transitions*. 5:1– 18
- Fischer, Frank (1995). *Evaluating public policy*. Chicago: Nelson-Hal
- Fisher, W.R. (1987). *Human Communication as Narration*. South Carolina: University of South Carolina Press, Columbia
- Fishbein M. and Ajzen I. (1975). *Belief, Attitude, Intention, and Behaviour: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.
- Fox, C.J. & H.T. Miller (1996). *Postmodern Public Administration*. Towards discourse. Thousand Oaks etc.: Sage
- Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. *Research Policy*, 33(6-7): 897–920.
- Geels, F. W. (2011). The multi-level perspective on sustainability transitions: Responses to seven criticisms. *Environmental Innovation and Societal Transitions*, 1: 24–40.
- Gherardi, S. and D. Nicoloni (2000). To Transfer is to Transform: The Circulation of Safety knowledge. *Organization*. 7(2):329-348.
- Giddens, A. (1984). *The Constitution of Society*. Cambridge: Polity Press
- Gold, J., Devins, D., & Johnson, A. (2003). What is the value of mentoring in a small business? Using narrative evaluation to find out. *British Journal of Guidance and Counselling*, 31(1), 51-62.
- Griessler, E., Hönigsmayer, H, Braun, R, and Frankus, E (2019) . NewHoRRizon Social Lab Manual – Revised Version. NewHoRRizon Deliverable 7.2, Grant Agreement EC 741402
- Grin, J. and H. Van de Graaf (1996). Technology Assessment as learning. *Science, Technology and Human Values*, 21 (1): 73-99.
- Grin, J., Rotmans, J., & Schot, J. (2010). *Transitions to sustainable development: New directions in the study of long term transformative change*. London: Routledge
- Grin, J. and Loeber, A. (2007). Theories of learning. Agency, structure and change, In: Frank Fischer, Gerald J. Miller, Mara S. Sidney (eds.), *Handbook of Public Policy Analysis. Theory, Politics, and Methods*. CRC Press – Taylor & Francis Group, pp. 201-222.
- Guijt, I. J. Brouwers, C. Kusters, E. Prins, B. Zeynalova (2011) Evaluation Revisited. Improving the Quality of Evaluative Practice by Embracing Complexity. Wageningen UR etc.: Conference Report.
- Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Sage.

- Hajer, M. A. (1995). *The politics of environmental discourse: ecological modernization and the policy process*. Oxford: Clarendon Press.
- Hajer, M. (2003). Policy without polity? Policy analysis and the institutional void. *Policy Sciences*, 36: 175–195.
- Halsema A. (2011). The time of the self. a feminist reflection on Ricoeur’s notion of narrative identity. In: Schües C, Orlowski DE, Fielding HA (eds.) *Time in Feminist Phenomenology*. Bloomington: Indiana University Press. pp. 111–131.
- Hargreaves, T. (2011). Practice-ing behaviour change: Applying social practice theory to pro-environmental behaviour change. *Journal of Consumer Culture*, 11: 79-99
- Hellstern, G.-M. (1991). Generating knowledge and refining experience: The Task of Evaluation. In: F.-X. Kaufmann (ed.) *The Public Sector Challenge for Coordination and Learning*. Berlin, New York: Walter de Gruyter.
- Herman, D. (2011). *Basic elements of narrative*. John Wiley & Sons.
- Hoffman, J. (2013). Theorizing power in transition studies: The role of creativity and novel practices in structural change. *Policy Sciences*, 46(3), 257–275.
- Hoffman, J.G. & Loeber, A.M.C., (2016). Exploring the Micro-politics in Transitions from a Practice Perspective: The Case of Greenhouse Innovation in the Netherlands. *Journal of Environmental Policy & Planning*, 18(5): 692-711.
- Hoogerwerf, A. (1990). Reconstructing Policy Theory. *Evaluation and Program Planning*, 13, pp: 285-291.
- Joas, H. (1994). The creativity of action: Pragmatism and the critique of the rational action model. In I. Carlgren, G. Handal, & S. Vaage (Eds.) *Teachers’ minds and actions: Research on teachers’ thinking and practice* (pp. 62–73). London, DC: The Falmer Press
- Ivaldi, Silvia, Giuseppe Scaratti, Gianni Nuti (2015) The practice of evaluation as an evaluation of practices. *Evaluation*, 21(4) 497–512
- Jones, M., E. Shanahan, M. McBeth (eds.) (2014). *The Science of Stories. Applications of the Narrative Policy Framework in Public Policy Analysis*. New York: Palgrave MacMillan
- Joss, S. & S. Bellucci (eds.) (2002) *Participatory Technology Assessment. European perspectives*. London: Centre for the Study of Democracy
- Kleiner, A., & Roth, G. (1997). *Learning histories: A new tool for turning organizational experience into action* (No. 002). MIT Center for Coordination Science.
- Klüver, L. et al. (2000). EUROPTA, European Participatory Technology Assessment; Participatory Methods in Technology Assessment and Technology Decision-Making. Report published on the internet at www.tekno.dk/europta (version October 18, 2000), Copenhagen: The Danish Board of Technology.
- Knorr Cetina, K. (2001). Objectural Practice. In: Schatzki, T. R., Knorr-Cetina, K., & von Savigny, E. (2001). *The practice turn in contemporary theory*. London: Psychology Press. pp.184-197.
- Kuzma, Jennifer & Pat Roberts (2018). Cataloguing the barriers facing RRI in innovation pathways: a response to the dilemma of societal alignment, *Journal of Responsible Innovation*, DOI: 10.1080/23299460.2018.1511329
- Labov, W. and Waletzky, J. (1967) ‘Narrative Analysis: Oral Versions of Personal Experience’, in J. Helm (ed.) *Essays on the Verbal and Visual Arts*. Seattle, WA: University of Washington Press.
- Lave, J. (1988). *Cognition in practice. Mind, mathematics and culture in everyday life*. Cambridge, UK: Cambridge University Press
- Laws, D. and Loeber, A.M.C. (2011). ‘Sustainable Development and Professional Practice: Dilemmas of action and strategies for coping. *Engineering Sustainability*, 164 (ES1): 25-33

Loeber, A. (2003). Inbreken in het gangbare. Transitie management in de praktijk: de NIDO-benadering. Leeuwarden: Nationaal Initiatief Duurzame Ontwikkeling.

Loeber, A. (2004). Practical wisdom in Risk Society. Methods and practice of interpretive analysis on questions of sustainable development. Unpublished PhD thesis, University of Amsterdam

Loeber, A. (2007). Designing for phronèsis: Experiences with transformative learning on sustainable development. *Critical Policy Analysis* [Critical Policy Studies], 1, 389–414.

Loeber, A. (2010). Evaluation as Policy Work: Puzzling and Powering in a Dutch Program for Sustainable Development. In: H. Colebatch, R. Hoppe, M. Noordegraaf (Eds.) *Working for Policy*. Amsterdam: Amsterdam University Press. pp. 131-152.

Loeber, A. (2015, July 1–4). ‘Emni’ - the (re-)construction of normalcy: A methodological exploration. Paper presented at the International Conference on Public Policy, Milan, Italy.

Loorbach, D. (2014). To Transition! Governance Penarch in the New Transformation. Inaugural Speech, Erasmus Universiteit Rotterdam.

McGowan, K. A., F. Westley, E. D. G. Fraser, P. A. Loring, K. C. Weathers, F. Avelino, J. Sendzimir, R. Roy Chowdhury, and M.-L. Moore (2014). The research journey: travels across the idiomatic and axiomatic toward a better understanding of complexity. *Ecology and Society*, 19(3): 37. <http://dx.doi.org/10.5751/ES-06518-190337>

Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy* 39(6): 751-760.

Polkinghorne, D. E. (1995). Narrative configuration in qualitative analysis. *International journal of qualitative studies in education*, 8(1), 5-23.

Polkinghorne, D.E. (1997). *Narrative Knowing and the Human Sciences*. New York: University of New York Press.

Pouliot, V. (2008) The logic of practicality: A theory of practice of security communities. *International Organization*, 62(2): 257-288.

Reckwitz, Andreas (2002) Toward a Theory of Social Practices. A development in culturalist theorizing. *European Journal of Social Theory*, 5(2): 245- 265.

Reynolds, A. (2015). (Breaking) The Iron Triangle of Evaluation. *IDS Bulletin* Vol. 46 No 1 January. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/1759-5436.12122>

Ricoeur, (1992) *Oneself as Another*. Chicago: University of Chicago Press.

Risseuw, C.I. (1988). The fish don't talk about the water. Gender transformation, power and resistance among women in Sri Lanka. Nijmegen: Unpublished PhD Thesis

Rogers, P (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*, 14(1):29-48

Schatzki, Theodore (2010). Materiality and Social Life. *Nature + Culture*, 5(2):123-149.

Schatzki, T. R., Knorr-Cetina, K., & von Savigny, E. (2001). *The practice turn in contemporary theory*. London: Psychology Press

Schon, D. A. (1971). *Beyond the Stable State*. New York: Random House

Schön, D.A. (1983). *The Reflective Practitioner: How professionals think in action*. New York: Basic Books.

KNAW, VSNU & NOW (2014). *Standard Evaluation Protocol 2015 – 2021*. [Standard Evaluation Protocol 2015 – 2021] Amsterdam: KNAW.

Shove, E. & Walker, G. P. (2007). CAUTION! Transitions ahead: politics, practice, and transition management. *Environment and Planning A*, 39(4): 763–770.

Shove, E. (2010). Beyond the ABC: climate change policy and theories of social change. *Environment and Planning A*, 42: 1273-1285

Souto-Manning, M. (2014). Critical narrative analysis: The interplay of critical discourse and narrative analyses. *International Journal of Qualitative Studies in Education*, 27(2): 159-180.

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9): 1568-1580. DOI: 10.1016/j.respol.2013.05.008

Strand, Roger / Jack Spaapen (2015). Indicators for promoting and monitoring Responsible Research and Innovation. Report from the Expert Group on Policy Indicators for Responsible Research and Innovation. With Martin W Bauer, Ela Hogan, Gema Revuelta, Sigrid Stagl. With contributions of Lino Paula, Ângela Guimarães Pereira. Directorate-General for Research and Innovation 2015 Science with and for Society EUR 26866.

Spence, D.P. (1986). Narrative smoothing and clinical wisdom. In T.R. Sarbin (Ed.), *Narrative psychology: the storied nature of human conduct* (pp. 211-232). New York: Praeger

Stake, R.E. (1974). Evaluating Educational Programmes. The Need and the Response. Paris: Organisation for Economic Cooperation and Development.

Switzer, A. (2019). Transitioning the Transport and Land-use system. Unpublished PhD Thesis. University of Amsterdam

Taanman, M. (2012). Working in the Science-Policy Interface: Transition Monitoring in the Dutch Energy Transition Program. In: Geert Verbong, Derk Loorbach (eds.) *Governing the Energy Transition. Reality, Illusion or Necessity?* New York: Routledge

Turnheim, Bruno, Frans Berkhout, Frank Geels, Andries Hof, Andy McMeekin, Björn Nykvist, Detlef van Vuuren (2015). Evaluating sustainability transitions pathways: Bridging analytical approaches to address governance challenges. *Global Environmental Change* 35: 239–253.

Urban, J. B., M. Hargreaves, W.M. Trochim (2014). Evolutionary Evaluation: Implications for evaluators, researchers, practitioners, funders and the evidence-based program mandate. *Evaluation and Program Planning*, 45: 127-139.

Van Exel, J., & De Graaf, G. (2005). *Q methodology: A sneak preview*. [available from www.jobvanexel.nl]

Mierlo, B.C. van, Regeer, B., Amstel, M. van, Arkesteijn, M.C.M., Beekman, V., Bunders, J.F.G., Cock Buning, T. de, Elzen, B., Hoes, A.C., Leeuwis, C. (2010). *Reflexive Monitoring in Action. A guide for monitoring system innovation projects*: Wageningen/Amsterdam: Communication and Innovation Studies, WUR; Athena Institute, VU - ISBN 9789085855996 – 104.

Von Schomberg, R. (2013). A vision of responsible innovation. In R. Owen, J. Bessant & M. Heintz (Eds.), *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*. West Sussex: John Wiley. pp. 51-74.

Wagenaar, H., & Cook, S. D. N. (2003). Understanding policy practices: Action, dialectic and deliberation in policy analysis. In M. Hajer & H. Wagenaar (Eds.) *Deliberative policy analysis: Understanding governance in the network society*. Cambridge, UK: Cambridge University Press. pp.139-171.

Walton, M. (2014). Applying complexity theory: A review to inform evaluation design. *Evaluation and Program Planning*. 45:119-126

Wickson, F., & Carew, A. L. (2014). Quality criteria and indicators for responsible research and innovation: learning from transdisciplinarity. *Journal of Responsible Innovation*, 1(3): 254-273.

Wilk, R. (2009). The Edge of Agency. Time, consumption and everyday life: *Practice, materiality and culture*, 14

Wittmayer, J. M., Backhaus, J., Avelino, F., Pel. B., Strasser, T. and Kunze, I. (2015) Narratives of change: how social innovation initiatives engage with their transformative ambitions (TRANSIT working paper 4). TRANSIT: EU SSH.2013.3.2-1 Grant agreement no: 613169.

Zhai, J. and Dillon, J. (2014). 'Communicating science to students: Investigating professional botanic garden educators' talk during guided school visits'. *Journal of Research in Science Teaching*, 51 (4), pp. 407–429. DOI: 10.1002/tea.21143.

Zwart, H., Landeweerd, L., & van Rooij, A. (2014). Adapt or perish? Assessing the recent shift in the European research funding arena from 'ELSA' to 'RRI'. *Life Sciences, Society and Policy*, 10, 11.

6. Annexes

Annex A – Template for Reflection and Reporting Moment I+II

<https://nextcloud.ihs.ac.at/index.php/s/dCEgqXGzF7tRkY7>

Annex B – Template for Reflection and Reporting Moment III

<https://nextcloud.ihs.ac.at/index.php/s/eMfNijBC8QJFxHp>

Annex C – Template for Reflection and Reporting Moment IV

<https://nextcloud.ihs.ac.at/index.php/s/8BM7ersqdanfeXa>

Annex D – Tool for the reconstruction of Theories of Change

<https://nextcloud.ihs.ac.at/index.php/s/osETk8iT5boeS9X>