

Christopher Luederitz, Niko Schöpke, Arnim Wiek, Daniel J. Lang, Matthias Bergmann, Joannette J. Bos, Sarah Burch, Anna Davies, James Evans, Ariane König, Megan A. Farrelly, Nigel Forrest, Niki Frantzeskaki, Robert B. Gibson, Braden Kay, Derk Loorbach, Kes McCormick, Oliver Parodi, Felix Rauschmayer, Uwe Schneidewind, Michael Stauffacher, Franziska Stelzer, Gregory Trencher, Johannes Venjakob, Philip J. Vergragt, Henrik von Wehrden, Frances R. Westley

Learning through evaluation – A tentative evaluative scheme for sustainability transition experiments

*Originally published in:
Journal of Cleaner Production,
169 (2017), 61-76
DOI: 10.1016/j.jclepro.2016.09.005*

*Christopher Luederitz a,**
Niko Schaepke a
Arnim Wiek a,b,c
Daniel J. Lang a,c,d
Matthias Bergmann a,e
Joannette J. Bos f
Sarah Burch g
Anna Davies h
James Evans i
Ariane Koenig j
Megan A. Farrelly k
Nigel Forrest b
Niki Frantzeskaki l
Robert B. Gibson m
Braden Kay b
Derk Loorbach l
Kes McCormick n
Oliver Parodi o
Felix Rauschmayer p
Uwe Schneidewind q
Michael Stauffacher r
Franziska Stelzer q
Gregory Trencher s
Johannes Venjakob q
Philip J. Vergragt t
Henrik von Wehrden c,d,u
Frances R. Westley v

Learning through evaluation – A tentative evaluative scheme for sustainability transition experiments

-
- a. Institute of Ethics and Transdisciplinary Sustainability Research, Faculty Sustainability, Leuphana University Lüneburg, Scharnhorststr. 1, 21335 Lüneburg, Germany
 - b. School of Sustainability, Arizona State University, Tempe, AZ 85287-5502, USA
 - c. Center for Global Sustainability and Cultural Transformation, Scharnhorststr. 1, 21335 Lüneburg, Germany
 - d. FuturES Research Center, Leuphana University Lueneburg, Scharnhorststr. 1, 21335 Lueneburg, Germany
 - e. ISOE - Institute for Social-Ecological Research, 60486 Frankfurt, Germany
 - f. Monash Sustainability Institute, Monash University, Clayton, VIC 3800, Australia
 - g. Department of Geography and Environmental Management, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada
 - h. Department of Geography, Trinity College Dublin, Dublin 2, Ireland

- i. The School of Environment, Education and Development, The University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom
 - j. University of Luxembourg, Campus Belval, Maison des Sciences Humaines, 11 Portes des Sciences, 4366 Esch-sur-Alzette, Luxembourg
 - k. School of Social Sciences, Monash University, Clayton, VIC 3800, Australia
 - l. DRIFT e Dutch Research Institute for Transitions, Faculty of Social Sciences, Erasmus University, Rotterdam, Burge-meester Oudlaan 50, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands
 - m. School of Environment, Resources and Sustainability, University of Waterloo, Waterloo, Canada
 - n. International Institute for Industrial Environmental Economics (IIIEE) at Lund University, P.O. Box 196, 22100 Lund, Sweden
 - o. Institute for Technology Assessment and Systems Analysis (ITAS), Karlsruhe Institute of Technology (KIT), P.O. Box 3640, 76021 Karlsruhe, Germany
 - p. UFZ Helmholtz Centre for Environmental Research, Department for Environmental Politics, Leipzig, Germany
 - q. Wuppertal Institute for Climate, Environment and Energy, Döppersberg 19, 42103 Wuppertal, Germany
 - r. Department of Environmental Systems Science, USYS TdLab, ETH Zurich, 8092 Zurich, Switzerland
 - s. Clark University, Department of International Development, Community and Environment, Worcester, USA
 - t. Tellus Institute, 11 Arlington Street, Boston, MA 02116-3411, USA
 - u. Centre of Methods, Leuphana University Lüneburg, Scharnhorststr. 1, 21335 Lüneburg, Germany
 - v. School for Environment Enterprise and Development (SEED), University of Waterloo, Canada
- * Corresponding author:
E-mail address: christopherluederitz@gmail.com
(C. Luederitz)

This is the author's version of a work that was accepted for publication. Changes resulting from the publishing process, such as editing, corrections and structural formatting, may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in the Journal cited above.

Learning through Evaluation – A Tentative Evaluative Scheme for Sustainability Transition Experiments

Abstract

Transitions towards sustainability are urgently needed to address the interconnected challenges of economic development, ecological integrity, and social justice, from local to global scales. Around the world, collaborative science-society initiatives are forming to conduct experiments in support of sustainability transitions. Such experiments, if carefully designed, provide significant learning opportunities for making progress on transition efforts. Yet, there is no broadly applicable evaluative scheme available to capture this critical information across a large number of cases, and to guide the design of transition experiments. To address this gap, the article develops such a scheme, in a tentative form, drawing on evaluative research and sustainability transitions scholarship, alongside insights from empirical cases. We critically discuss the scheme's key features of being generic, comprehensive, operational, and formative. Furthermore, we invite scholars and practitioners to apply, reflect and further develop the proposed tentative scheme – making evaluation and experiments objects of learning.

1 1. Introduction

2 Sustainability problems of economic development, ecological integrity, and social justice
3 jeopardize human and social wellbeing around the world (Parris and Kates, 2003; Steffen et
4 al., 2015). Considering the extent of the problems, viable solutions need to yield
5 *transformational* changes, i.e., large-scale transitions of priorities, practices, and
6 infrastructures (McAlpine et al., 2015; McCormick et al., 2013; Westley et al., 2011).

7 Around the world, collaborative initiatives have emerged that design, implement, and monitor
8 experiments in real-world settings in support of sustainability transitions (Evans and
9 Karvonen, 2011; Trencher et al., 2014a; van den Bosch, 2010). Such experiments differ with
10 regard to their actor constellation, topical focus and governance structure (e.g. Castán Broto
11 and Bulkeley, 2013; Voytenko et al. 2015). While in the past a large number of experiments
12 have been led by citizens and local government organizations, a specific type of transition
13 experiment has emerged during the last decade. The new type of transition experiment is
14 characterized by cross-organizational collaboration between actors from academia and
15 society (government, industry and citizenry) with the aim of collaboratively fostering
16 transformational change and progress towards greater sustainability (Neuens et al., 2013;
17 Voytenko et al., 2015). Although often framed differently, such initiatives can be understood
18 to jointly experiment with a range of sustainability solutions, including but not limited to food
19 production (e.g. Victorian Eco Innovation Lab, Australia), energy consumption (e.g. Campus
20 as a Living Laboratory, Canada), urban living (e.g. Low Carbon Labs, Lund) and mobility
21 (e.g. Delft Design Labs, the Netherlands). Transition experiments are essential to the
22 scientific field of sustainability transitions (Caniglia et al., this issue) and are often carried out
23 by real-world laboratories or labs, in contrast to isolated scientific laboratories, including but
24 not limited to living labs, transition labs, and social innovation labs (e.g. Frantzeskaki et al.,
25 2014; Westley et al., 2014; McCormick and Kiss, 2015, cf. supplementary material A). Thus,
26 a given real-world laboratory can conduct various sustainability transition experiments for
27 testing transformational changes. While different labels are used for describing this process,
28 they all provide “spaces that facilitate explicit experimentation and learning based on
29 participation and user involvement” (Voytenko et al., 2015, p. 4). Accordingly, sustainability
30 transition experiments function also as an umbrella term for transformational interventions as
31 they build on existing efforts, create new actions and add orientation to transitions. They
32 follow a transdisciplinary research approach, integrating various actors into the
33 experimentation process for reconciling diverging preferences and practices, as well as
34 create ownership for sustainability problems and solutions (Lang et al. 2012). Importantly,
35 the sustainability practices experimented on do not concern mere modification or “tinkering”
36 of elements already present. Instead, they are radically different from the status quo, in both
37 process and outcomes (Bernstein et al., 2014; Davies and Doyle, 2015; Evans and
38 Karvonen, 2014).

39 Sustainability transition experiments often focus on defined small-scale settings, specific to a
40 particular location and socio-cultural context (Evans and Karvonen, 2014; Voytenko et al.,
41 2015). Following the notion of experimentation, the intention is to create positive outcomes
42 that are replicable, transferable, and scalable to society at large (Bernstein et al., 2014; Bos
43 et al., 2015;; Ryan, 2013). Experiments focus, for example, on socio-technical innovations
44 (e.g. in the energy or food sector) (e.g. van der Laak et al., 2007), on networks (e.g. political
45 and technical coalitions) (e.g. Bos et al., 2015), or on small spatial or organizational units
46 (e.g. a neighborhood or a building) (e.g. Brown and Vergragt, 2008). In addition to having
47 real-world impacts, such experiments are *research* endeavors to the extent that they
48 produce *evidence* regarding both the persistent unsustainability of dominant regimes and the
49 possible solutions to given sustainability problems within the bounded space of a laboratory
50 (Evans and Karvonen, 2011; Wiek et al., 2015). Thus, this article posits that sustainability
51 experiments (i) define a baseline and a goal for their evaluation, (ii) create a specific set-up
52 to administer interventions, (iii) measure the effects of interventions against the baseline and
53 the goal, (iv) evaluate the effects against sustainability criteria, and (v) offer evidence-

54 supported recommendations on how to mainstream solutions (Karvonen and van Heur,
55 2014; Laakso and Lettenmeier, 2015; Wiek et al., 2015).

56 Transitions scholarship has long recognized the significant potential of transition
57 experiments in generating new knowledge and promoting social learning (e.g. Bos et al.,
58 2013; Farrelly and Brown, 2011; Pahl-Wostl, 2007). Iterative and reflexive monitoring and
59 evaluation needs to be an integral part of sustainability transition experiments to support
60 individual and organizational *learning* promoting ongoing change and up-scaling impact
61 (Forrest and Wiek, 2014; Taanman, 2014; van Mierlo et al. 2010). By addressing the
62 broader systemic transition context within which such initiatives sit, the opportunities for
63 deepening, broadening, and scaling-up of such experiments could be increased (Raven et
64 al., 2010). While the framing of actions, projects, and initiatives as experiments has become
65 popular around the world and they are being positioned as drivers of wider transition their
66 impacts are poorly understood (Caniglia et al., this issue). Therefore, scholars are calling for
67 greater cross-case learning from different sustainability transition experiments (Forrest and
68 Wiek, 2015; McCormick et al., 2013; Raven et al., 2011). Undertaking evaluative research
69 supports conclusions regarding the success of particular interventions, aids generalizing
70 insights, and enables the improved design and operation of experiments, helping them to
71 become more effective and efficient (Wiek et al., 2015).

72 Evaluation of sustainability transition experiments is faced with various challenges.
73 Transitions initiatives are no longer conducting 'projects' but aim to create a new setting for
74 transforming conventional practices and informal power structures (Nevens et al., 2013;
75 Kemp 1998; Geels and Ravens, 2006). Nevertheless, sustainability transition experiments
76 often remain the most tangible approach (Nevens et al., 2013). Their objective is to initiate
77 and facilitate radical long-term transitions (Rotmans and Loorbach, 2009; Loorbach, 2010),
78 but orchestrate this through specific experiments, which aim to challenge the status quo.
79 Scholars argue that aligning experimentation alongside prevalent structures and paradigms
80 is necessary in the short-term, while ultimately aiming towards a long-term transformation
81 (Schot and Geels, 2008; Robinson et al. 2011).

82 Reflexive evaluation of experiment enables learning-by-doing; a critical mechanism
83 supporting sustainability transitions (Taanman, 2014). Thus, evaluation emerges as a core
84 activity in transitions, periodically informing experiments to adapt, extend and revise the
85 envisioned pathway. To achieve this requires: ex-ante evaluation prior to the implementation
86 of experiments to inform their design; formative evaluation to adjust and improve ongoing
87 experiments; and, ex-post evaluation to appraise the contribution of experiments to
88 sustainability after completion. Evaluations scrutinize assumptions, structures, and values as
89 well as related and unrelated changes in society in order to inform future actions (Schot and
90 Geels, 2008; Rotmans and Loorbach, 2010; Robinson 2003). Embedded within these
91 different modes of evaluation are reflexive learning processes which continually assess the
92 transformational potential of experiments and the evaluation itself. As sustainability transition
93 experiments are embedded within structures and power relations, advanced reflexivity within
94 an evaluation is required (Avelino and Rotmans, 2009).

95 A number of studies have explored ways to appraise the outcomes of transition experiments,
96 but coordinating efforts are widely lacking (Bai et al., 2010; Ferguson et al., 2013; Forrest
97 and Wiek, 2014; Hart et al., 2015; König, 2015; Loorbach et al., 2015; Moloney and Horne,
98 2015; Moore et al., 2014; Seyfang and Longhurst 2016; Taanman, 2014; Trencher et al.,
99 2014b). Although these studies provide useful insights into aspects of sustainability
100 transition experiments, none of them comprehensively covers a broad array of aspects
101 critical to (different types of) experiments. This partly arises from the diversity of the different
102 types of initiatives surveyed, which extend from, for example, transition policy programs,
103 transition management projects, technical innovation projects, to community initiatives or
104 social innovation processes. In addition, learning and coordination across various transition
105 experiments is constrained by the use of different, case-specific evaluative schemes, if one
106 exists at all.

107 Other fields, such as international development and resource management, have
108 demonstrated how evaluative schemes, if used jointly, can successfully facilitate and
109 accelerate learning and progress, as they allow learning and coordination across similar
110 case studies (Banerjee et al., 2010; Ostrom, 2009). For instance, the diagnostic social-
111 ecological systems framework for analyzing elements and their interrelation in coupled
112 social-ecological system is a pivotal example of such efforts. The framework – developed
113 and advanced by Elinor Ostrom and others (e.g. Ostrom, 2007; Ostrom and Cox, 2010;
114 McGinnis and Ostrom, 2014; Leslie et al., 2015; Vogt et al., 2015) – departs from conditions
115 in common-pool resource systems that are considered crucial for enabling self-organization.
116 While the framework provides a common terminology for understanding socio-ecological
117 systems, without implying causal relations, it is sensitive to context specifics and supports
118 generalization and theory building (Partelow, 2015). This facilitates interdisciplinary
119 collaborations and invites different theories for explaining observed dynamics (McGinnis and
120 Ostrom, 2014). The framework is widely used in research on water, food, and forestry
121 systems (e.g. Vogt et al., 2015; Partelow and Boda, 2015; Marshall, 2015).

122 In this article, we present a tentative evaluative scheme for sustainability transition
123 experiments, with the notion that when applied, this would facilitate learning across different
124 transition experiments, and help fostering sustainability transitions. We aim to systematically
125 support designing and improving transition experiments as well as tracing their influence on
126 learning and transformational efforts while ensuring reflexivity regarding the limitation of
127 such undertakings. Overall, this paper seeks to identify the essential characteristics of a
128 tentative evaluative scheme which will increase its: broad applicability; readiness to be
129 applied; comprehensiveness; and, its capacity to improve the performance of experiments.

130 The purpose of this article is to provide a conceptual basis for further discussions on the
131 potentials, needs, restrictions, and drawbacks of experiments evaluation efforts. This applies
132 to academic work on evaluation such as the publication of findings from various
133 sustainability transition experiments. It also applies to practical work such as the
134 collaborative application of the scheme involving researchers and practitioners to facilitate
135 mutual learning. We emphasize the tentative nature of the evaluative scheme inviting
136 participants of experiments – both in research and practice – to critically reflect upon its
137 potentials and limitations and take part in learning from and improving transition efforts. This
138 involves continuous changes in the evaluative features and processes of evaluation (see
139 McGinnis and Ostrom, 2014).

140 This article departs from an evaluative scheme developed in a study on urban sustainability
141 experiments (Wiek et al., 2015). Here, we further develop and expand on this study, drawing
142 on the existing literature that deals more generally with transition experiments and initiatives.
143 With support from this literature, the evaluative scheme ought to be:

- 144 (i) *Generic*, i.e., applicable to different types of sustainability transition experiments;
- 145 (ii) *Comprehensive*, i.e., capturing the ultimate outcomes as well as the intermediate
146 and mediating attributes (inputs, processes, outputs) of experiments;
- 147 (iii) *Operational*, i.e., ready to be applied (including guidance on how to specify it for
148 application to particular cases and contexts); and,
- 149 (iv) *Formative*, i.e., support experiments in becoming more effective and efficient.

150 The method of this article is as follows. After developing the conceptual framework for the
151 evaluative scheme, a literature review was conducted. This drew on an array of reported
152 sustainability experiments to illustrate and define the evaluation schemes' various
153 dimensions. This process followed a four-step procedure. First, we identified and pooled
154 suitable publications on experiments from Scopus and Google Scholar (see supplementary
155 material A). The search was limited to peer-reviewed case studies to ensure some degree of
156 scientific rigor and quality control in the analyzed material. Selection criteria were that the

157 articles (i) were empirical studies, that (ii) reported on collaborative science-society
158 initiatives, (iii) explicitly focused on sustainability, and (iv) employed transition approaches
159 with an experimental character. Selected studies range from intervention studies in which
160 the authors present their own experiments (e.g. Bernstein, et al., 2013) to case studies in
161 which the authors report on an experiment (e.g. Evans and Karvonen, 2014). Since our
162 literature review includes only peer-reviewed articles in English and overlooks non-refereed
163 publications, we are cognizant of particular biases created; from excluding certain types of
164 studies (i.e. non-refereed or non-English). Yet we consider it sufficient for the purpose of
165 developing a tentative evaluative scheme as the reviewed literature reports on a broad
166 range of initiatives, including possible contestation and further enrichment of the literature
167 used in following sections. Second, we extracted information from 61 unique case studies for
168 conceptualizing inputs, processes, outputs, and outcomes as basic categories of the
169 evaluation scheme. Third, we identified features and related definitions, exemplified typical
170 indicators, illustrated examples, and presented literature in support of each of the above
171 categories. In the spirit of a *tentative* scheme, the collection of examples and indicators is
172 not exhaustive. The presented examples of the developed features are selected according
173 to their respective suitability intending to support operationalization of the scheme and
174 experimental designs. The indicators, although not fully operationalized, serve as reminders
175 and placeholders to identify and translate features into measurable parameters when
176 operationalizing the scheme. Fourth and finally, in the process of finalizing the evaluation
177 scheme, preliminary versions have been presented, discussed and revised according to in-
178 depth feedback from audiences at numerous international conferences (see
179 Acknowledgements). The input enabled initial appraisal of the scheme's applicability and
180 comprehensiveness as well as supported deliberation regarding its use in cross-case
181 analysis.

182 The article is structured as follows. In Section 2, we present the conceptual framework,
183 followed by the evaluative scheme in Section 3. We then conclude by critically reflecting on
184 the evaluative scheme against the four guidelines presented above.

185

186 **2. Conceptual Framework for the Evaluation Scheme**

187 The evaluative scheme presented below (Figure 1; Section 3) is used to appraise the extent
188 to which a sustainability transition experiment generates desired effects, and how this was
189 accomplished (i.e., through what kind of interventions). The scheme is based on the basic
190 logic model of evaluation (McLaughlin and Jordan, 2010; Rossi et al., 2004), which is
191 organized according to four evaluative dimensions: *inputs* that are invested into the
192 experiment, *processes* that are performed by the experiment, *outputs* that are generated by
193 the experiment, and sustainability *outcomes* that are accomplished by the experiment.
194 However, there are two important modifications. First, we change the sequence of items
195 from the *experiment* rationale (Inputs → Processes → Outputs → Outcomes) to the
196 *evaluation* rationale with the primary interest in outputs and outcomes, and from there
197 tracking back processes and inputs (Forrest and Wiek, 2014). Second, we depict the logical
198 model components as parallel and interdependent, which requires iterative evaluation
199 among the four dimensions. In other words, inputs are not only needed for initiating an
200 experiment nor are outputs only produced after completion of a project. For example,
201 outputs might initiate new processes or generate new investments of additional resources
202 amid the experimentation. Thus, the presented scheme aims at being capable of capturing
203 complex dynamic processes with overlapping and parallel interferences. The evaluation
204 scheme is guided by the following four questions:

- 205 1. *What was generated?* – Identify the produced *outputs* and related features
206 including direct results of the interventions; namely built capacities (results of
207 learning processes), actionable knowledge, accountability, structural changes, up-

208 take of experiments, as well as generalizable insights with regards to specific
209 issues or methods.

210 2. *What was accomplished?* – Identify achieved *outcomes* in terms of sustainability.
211 This explores the extent to which generated changes support progress towards
212 sustainability, namely socio-ecological integrity, livelihood sufficiency and
213 opportunities, intra- and intergenerational equity, resource maintenance and
214 efficiency, socio-ecological stewardship and democratic governance, as well as
215 precaution and adaptation (Gibson, 2006).

216 3. *How was it completed?* – Identify what *processes* led to outputs and outcomes
217 such as sequence of actions, sound methodology, collaboration, reflexivity and
218 learning, and transparency.

219 4. *What was invested?* – Identify *inputs* that enabled actions and processes and
220 related features, i.e. initial awareness, commitment, expertise, trust, and support
221 (incl. financial and human resources).

222 These guiding questions can inform all types of evaluation: *ex-ante* evaluation to inform the
223 design of experiments, *formative* evaluation to adjust and improve experiments, or *ex-post*
224 evaluation to appraise the contribution of experiments to sustainability.

225 [--- Insert here Figure 1 ---]

226 [--- Insert here the legend of Figure 1 ---]

227 Figure 1: Dimensions of the evaluative scheme for appraising sustainability transition experiments.
228

229 3. Evaluative Scheme for Sustainability Transition Experiments

230 This section further describes the four evaluative dimensions (outputs, outcomes,
231 processes, and inputs) and presents for each identified feature definitions, typical indicators,
232 illustrative examples, and evaluative questions. We present instructive definitions of each
233 evaluative feature as well as formative evaluative questions in Box 1.

234 3.1 Output Features

235 Outputs are *direct* results of sustainability transition experiments, including built capacities,
236 actionable knowledge, structural changes, as well as the up-take of experiments (Wiek et al.,
237 2015). These key outputs may have differing importance depending on the experiment and
238 can be interconnected in various ways. For example the capacities built in participants
239 enable them to generate actionable knowledge and increase accountability for the realized
240 structural changes. Additional features include the generalization of evidence for generating
241 outputs to support the up-take of the experiment to broader application, as well as the
242 integration of generalizable knowledge into the scientific discourse

243 3.1.1 Built capacities

244 Sustainability transition experiments build capacities such as skills, abilities, and crafts that
245 foster or embrace sustainability (Bos et al., 2013; Loorbach et al., 2015; Wiek and Kay,
246 2015). Such capacities go beyond skillfully conversing on sustainability issues towards
247 enabling people to *act sustainably* in their everyday decision-making and practices. Built
248 capacities include strategic competence in developing effective interventions (Schreuer et
249 al., 2010), practical skills, such as creating and maintaining a community garden (Bernstein
250 et al., 2014), and interpersonal competence for building coalitions and alliances
251 (Frantzeskaki et al., 2014; Wittmayer et al., 2014). Experiments can also be used as learning
252 settings for educating students (Bernstein et al., 2014; Ryan, 2013; Trencher et al., 2016) as
253 well as for educating practitioners on new solutions and (possibly) new roles and
254 responsibilities for sustainability transitions (Farrelly and Brown, 2011). Typical indicators for
255 built capacities are post-experiment activities and practices carried out by participants that

256 have the potential to address the given sustainability problem such as community gardening
 257 and food distribution systems, consumption of organic food products, launching of new
 258 sustainability-based businesses, expansion of networks, and incorporation of sustainability
 259 into decision-making in the public or private sector.

260 An illustrative example of built capacity as output of a transition experiment is the capacity
 261 built in planners and other participants to develop long-term sustainability plans in Phoenix,
 262 United States, as reported by Wiek and Kay (2015).

263 The evaluative question for this feature is: *Does the transition experiment build capacities in*
 264 *participants to generate sustainability solutions?*

265 **3.1.2 Actionable knowledge**

266 Actionable knowledge is evidence-supported guidance for practical application that has been
 267 tested in successful efforts to solving (or at least mitigating) a sustainability problem within
 268 the defined experimental setting (Forrest and Wiek, 2014; Frantzeskaki and Kabisch, 2016).
 269 Three knowledge types are relevant to sustainability transition experiments. The first two are
 270 analytical-descriptive knowledge about the given sustainability problem (Wittmayer et al.,
 271 2014) and anticipatory, normative knowledge about the sustainability goals (Davies et al.,
 272 2012; Frantzeskaki and Tefrati, 2016). The third knowledge output of experiments is
 273 transformational knowledge on the most effective means of fostering transitions from the
 274 current to a (more) sustainable state (Ceschin, 2014; Wittmayer et al., 2014; Bos and Brown,
 275 2012). This feature includes scientific output as well as knowledge generated by
 276 practitioners. Typical indicators for actionable knowledge may include scientific output as well
 277 as context specific transition pathways that identify strategic actions for implementing
 278 transformational change and building agreement on the problem framing.

279 An illustrative example of actionable knowledge as output of a transition experiment is the
 280 developed transition management approach for coordinating ambitious strategies for the City
 281 of Aberdeen, UK, as reported by Frantzeskaki and Tefrati (2016). Civil servants from the city
 282 department and participants from civil society valorized the knowledge gained in
 283 implementing experimental settings for opening a center for developing skills that are
 284 required for a low-carbon economy.

285 The evaluative question for this feature is: *Does the transition experiment generate*
 286 *actionable knowledge that provides evidence on how to generate sustainability solutions?*

287 **3.1.3 Accountability**

288 Accountability refers to participants' commitment, maybe even formalized through
 289 agreements and agreed-upon sanctions, to implement results generated by the experiment
 290 and dedication to positive change (Wiek and Kay, 2015). Participants develop confidence
 291 about being able to implement the selected actions when actively participating in the
 292 experiments. Participants' commitment to the identified actions is enhanced as the
 293 participants learn about the actions' effectiveness in the process of pursuing sustainability
 294 transitions. Confidence and commitment can be built especially well through transition
 295 experiments that try novel practices and experience positive results (Wittmayer et al., 2014).
 296 Allowing for ownership of the vision and promoting transition experiments as the stepping-
 297 stones for realizing sustainability goals support accountability (Frantzeskaki et al., 2014).
 298 Typical indicators for accountability are the participants' attitudes, but also more formalized
 299 commitments towards the implementation of the results.

300 An illustrative example of accountability as output of a transition experiment is the
 301 community center that was reopened by active citizens in Rotterdam (neighborhood of
 302 Carnisse), the Netherlands as reported by Wittmayer and Schöpke (2014). The center
 303 continued operation based on the positive results of the experiment.

304 The evaluative question for this feature is: *Does the transition experiment build confidence*
 305 *and commitment for generating and realizing sustainability solutions?*

306 **3.1.4 Structural changes**

307 Sustainability transition experiments generate an array of structural changes to foster rapid
308 transformations (Evans and Karvonen, 2011; Trencher et al., 2014b). Such outputs of
309 experiments can be subdivided into physical change (transformation of infrastructure), and
310 societal change (transformation of institutions).

311 *Changes in physical structures*

312 Change of physical structures refers to the creation of new or transformation of existing
313 buildings, infrastructures, technologies and products. These real-world changes are often
314 radically different from the existing structures (Vergragt and Brown, 2007) and can include
315 sustainable buildings (Trencher et al., 2014a; Vergragt and Brown, 2012), green
316 infrastructure (Bernstein et al., 2014), innovative energy systems (Hart et al., 2015), and new
317 vehicles (Brown et al., 2003). However, real-world changes in physical structures may also
318 correspond to changed understandings, priorities, practices, and behavior (see below).
319 Typical indicators for physical transformation would incorporate modified or newly built forms
320 such as new bicycle lanes, rooftops, novel or improved products arising from new scientific
321 knowledge and innovations. Other indicators would be commercialization of patents; shifts in
322 the design, production and manufacturing of goods; and changes in the natural environment,
323 for example, afforested areas or increasing green spaces in urban areas.

324 An illustrative example of physical changes as output of a transition experiment is the
325 bicycle-based transport technology for elderly people that changed mobility behavior in Cape
326 Town, South Africa reported by Ceschin (2014).

327 The evaluative question for this feature is: *Does the transition experiment generate physical*
328 *changes that support solutions for the identified sustainability problem?*

329 *Changes in societal realms*

330 Sustainability experiments are also undertaken to deliver societal change. Societal change
331 refers to the creation of new or transformation of existing networks and organizations, values
332 and norms, rules and policies, decision-making processes, behavior and practices, and
333 discourses, often radically different from existing ones (Bos and Brown, 2012; Davies and
334 Doyle, 2015; Schreuer et al., 2010). Societal changes induced by experiments include
335 changed norms (Davies et al., 2015), policies (Loorbach and Rotmans, 2010), mobility
336 practices (Ceschin 2014), and political discourses (Loorbach and Rotmans, 2010). Typical
337 indicators for societal change are new or altered activities, practices, routines, as well as
338 social relations and partnerships.

339 An illustrative example of societal real-world changes as output of a transition experiment is
340 the organizational innovation in health care in the Netherlands reported by Loorbach and
341 Rotmans (2010). Contrary to conventional practices, the “Buurtzorg” (District Care)
342 establishes small nurse teams that are responsible for a small group of clients, have their
343 own budget and possess freedom to self-organize their professional practices.

344 The evaluative question for this feature is: *Does the transition experiment generate societal*
345 *changes that support solutions for the identified sustainability problem?*

346 **3.1.5 Facilitate up-take**

347 The ultimate objective of conducting transition experiments is to provide generalizable
348 evidence that a solution works beyond overly specific and narrow circumstances (Bos and
349 Brown, 2012; Vandevyvere and Nevens, 2015). A transition experiment is intended to
350 facilitate the up-take of its results. This anticipates that the results of an experiment can be
351 either transferred or scaled for broader use. This allows the participants and affected
352 stakeholders to utilize the results of the experiment for formulating solutions to similar
353 challenges, either in other contextual settings (transferability) or in system wide applications
354 (scalability) (Ceschin, 2014). More specifically, transferability refers to the potential that the

355 experiment can be replicated – whether application of the experiment in a different context
356 would generate similar results. Scalability refers to the potential that the experiment can be
357 expanded - whether nurturing the experiment in the given context would generate desired
358 results throughout the system. This can be achieved through ‘scaling out’ which refers to
359 repeating the experiment in the same context or through ‘scaling up’ which refers to
360 integrating and applying the experiment at a higher system level. Facilitating the take-up
361 requires generalizing insights gained through the experimentation including the anticipation
362 of potential negative side effects. Furthermore, experiments allow for additional insights that
363 can enrich the scientific discourse, including substantiation of methods for or theories of
364 socio-ecological transformations.

365 *Transferability*

366 Transferability refers to generalized lessons learned from an experiment that can be applied
367 in different contexts (Ceschin, 2014). This requires extraction of generic, process-related
368 factors and case specific knowledge that have supported application (Brown and Vergragt,
369 2008; Forrest and Wiek, 2015; Westley et al., 2014). Indications of transferability can best be
370 generated through feasibility and comparative studies. It should be noted that replicating the
371 experiment in similar or different contexts (e.g. Ray, 2013) is *actually transferring* the
372 insights and thus goes beyond the indication of transferability. Exemplary insights for
373 transferability can be gained through related feasibility studies, comparative studies, or
374 contextualization of an experiment through conceptual reasoning. Related typical indicators
375 are reliability of insights in other contexts or validity of cause and effect assumptions in
376 various settings.

377 An illustrative example of transferability as output of a transition experiment is reported by
378 Bos and Brown (2012). Following the implementation of an experiment in a catchment area
379 in Sydney, Australia, a project was initiated to transfer and extend sustainable water
380 management planning into other areas.

381 The evaluative question for this feature is: *Does the transition experiment indicate how the*
382 *sustainability solution can be transferred to different contexts?*

383 *Scalability*

384 Scalability refers to generalizable knowledge that facilitates the up-take of experiment
385 results. This can concern system-wide applications through “scaling out” in the initial system,
386 or applications at a larger system level through “scaling up” (Bos and Brown, 2012; Ceschin,
387 2014; Smith et al., 2014;). In both cases, translating and applying small-scale processes into
388 a larger scale entails collaboration with more actors (Laakso and Lettenmeier, 2015) as well
389 as translational competence (Smith, 2007). Scalability can be demonstrated through the
390 evaluation of scalable properties of solutions. Exemplary insights with regards to scalability
391 can be gained via related feasibility studies including engagement of actors working at
392 targeted scales. Actual efforts to take experimental results and scaling them out or up go
393 beyond mere indication of scalability. A typical indicator is the independence of measures
394 from changing governance systems on different scales.

395 An illustrative example of scalability as an output of an experiment is reported by Trencher et
396 al. (2014a) where results from building and mobility experiments in the 2000-Watt Society
397 Basel Pilot Region are shared with industry and government stakeholders across
398 Switzerland, to foster change in policy and industry practice on the national level.

399 The evaluative question for this feature is: *Does the transition experiment indicate the*
400 *potential for and how outputs can be scaled out to broader applications or up to higher*
401 *hierarchical levels?*

402 *Accounting for unintended consequences associated with up-take*

403 In some contexts, up-take of sustainability solutions may generate both positive and
404 negative unintended consequences (Evans and Karvonen, 2011; Smith et al., 2014). Careful

405 consideration of potential interactive effects is necessary for anticipation and evaluation of
 406 the risks and opportunities related to transferring and scaling experiments. In particular,
 407 when processes of an experiment are applied in contexts with different characteristics or if
 408 up-taking exposes an experiment to changed dynamics. Typical indicators are consideration
 409 of rebound effects, long-term consequences, and the potential for co-optation and offsetting
 410 of sustainability gains.

411 An illustrative example for reducing the risks of unintended consequences as outcome of a
 412 transition experiment is the self-build construction package for harvesting rain-water in north
 413 eastern Brazil reported by Smith et al. (2014). The up-take of the experiment contained self-
 414 build aspects to enhance community interactions and empower people instead of creating
 415 dependencies on local elites.

416 The evaluative question for this feature is: *Does the transition experiments account for*
 417 *unintended consequences that are associated the up-take of sustainability solutions?*

418 3.2 Outcome Features

419 Outcomes refer to sustainability-related accomplishments of the experiment, and provide a
 420 basis for examining the extent to which a transition experiment contributed to sustainability
 421 (Forrest and Wiek, 2014; Wiek et al., 2015). Reporting on sustainability transition
 422 experiments often fails to provide a comprehensive appraisal of the resulting sustainability
 423 effects. Good appraisals are not easy because they face two competing demands. They
 424 need to apply a consistent set of criteria to allow comparison of outcomes among
 425 experiments. But they must also recognize that the outcomes may vary depending on the
 426 focus of the experiment (e.g. on water, food, energy or neighborhood development) and the
 427 specifics of the context. We have therefore chosen to evaluate sustainability outcomes by
 428 adopting an established set of comprehensive criteria as a common framework and then
 429 specify the criteria for the particular cases and contexts (Gibson, 2006; Gibson et al., 2005).
 430 Bearing in mind that not all features apply to every experiment, this approach supports
 431 evaluations that deliver comparable findings about sustainability outcomes.

432 3.2.1 Socio-ecological integrity

433 Socio-ecological integrity is a sustainability requirement that recognizes the interdependence
 434 of human well-being and bio-physical conditions (Gibson et al., 2005, p. 95-98).
 435 Operationalizing this feature for sustainability transition experiments in urban planning
 436 requires for instance harmonizing physical structures and respective human activities
 437 (Section 3.1.4) with biophysical processes and elements (Luederitz et al., 2013). It involves
 438 preventing degradation or compromising of ecosystem services and reducing overall
 439 demands on already stressed life-support systems, enhancing the regenerative capacity of
 440 natural resources, and as a last resort offsetting unavoidable adverse impacts (Lamorgese
 441 and Geneletti, 2013). Typical indicators are new green walls and roofs, ecosystem-based
 442 spatial planning including adapted user behavior, and new, improved or prioritized habitat
 443 (i.e. blue and green infrastructure).

444 An illustrative example for ensuring socio-ecological integrity as outcome of a transition
 445 experiment is the tree and shade program that was implemented to mitigate negative urban
 446 sprawl effects and ensure recreation of life-support functions in Phoenix, United States
 447 reported by Bernstein et al. (2014).

448 The evaluative question for this feature is: *Do the transition experiment's outputs strengthen*
 449 *socio-ecological integrity?*

450 3.2.2 Livelihood sufficiency and opportunity

451 Human well-being depends on sufficient access of individuals and communities to what is
 452 needed for a decent life. This includes ensuring availability of opportunities for exercising
 453 positive human powers and capabilities in the specific context (Gibson et al., 2005, p. 98-

454 101). In water governance cases, for example, operationalizing this feature requires that
 455 built capacities (Section 3.1.1) and structural changes (Section 3.1.4) support human
 456 prosperity. It includes providing long-term access to water with sufficient quality and quantity
 457 to satisfy people's basic livelihood needs, enhance their psycho-physical well-being, and
 458 pursue economic activities while also maintaining ecological functions (Larson et al., 2013).
 459 Typical indicators are access to potable water and availability of water.

460 An illustrative example for livelihood sufficiency and opportunity as an outcome of a
 461 transition experiment is the LED lighting introduction initiative implemented by Columbia
 462 University in the Millennium Villages Project in Malawi. Adkins et al. (2010) report that
 463 following the experiment village inhabitants saved significantly in kerosene expenditures and
 464 reported higher levels of satisfaction regarding lighting quality.

465 The evaluative question for this feature is: *Do the transition experiment's outputs enhance*
 466 *livelihood sufficiency and opportunity?*

467 **3.2.3 Intra- and intergenerational equity**

468 This feature refers to sufficient and effective choices that reduce disparity between the rich
 469 and the poor and enhances future generations' opportunities to pursue sustainable lives
 470 (Gibson et al., 2005, p. 101-105). Again in water governance cases, operationalizing intra-
 471 and intergenerational equity for water management requires that actionable knowledge
 472 (Section 3.1.2), built capacity (Section 3.1.1), and structural changes (Section 3.1.4) improve
 473 equity. It includes enhancing life-support systems to meet everyone's basic needs and
 474 sharing social and economic benefits and costs between upstream and downstream users.
 475 In addition, decision-making is required that improves long-term renewability of freshwater
 476 resources and supports efficient and wise use of water (Shah and Gibson, 2013). As such,
 477 experiments go beyond inclusion and participation of a diverse array of social groups into
 478 creating opportunities in actively empowering them to be part of on-going and future
 479 sustainability transitions. Typical indicators are the creation of opportunities for various social
 480 groups, particularly those least privileged, and ensuring equity between providers and
 481 beneficiaries.

482 An illustrative example for intra- and intergenerational equity as an outcome of a transition
 483 experiment is the Community Watershed Stewardship Program in Portland, United States,
 484 as reported by Miller et al. (2015). In collaboration with the university the program
 485 experimented with application procedures, messaging and outreach to increase the number
 486 of projects that involved underrepresented communities while producing watershed health
 487 benefits.

488 The evaluative question for this feature is: *Do the transition experiment's outputs improve*
 489 *intra- and intergenerational equity?*

490 **3.2.4 Resource maintenance and efficiency**

491 Creation of sustainable livelihoods for all requires the reduction of demands on the
 492 biosphere that jeopardize long-term socio-ecological integrity. That in turn entails cutting
 493 material and energy use per unit of benefit (Gibson et al., 2005, p. 105-107).
 494 Operationalizing this feature for agricultural energy production requires that structural
 495 changes (Section 3.1.4) ensure benign production, support soil fertility, reduce greenhouse
 496 gas emissions and consider rebound effects. Key means include the application of cleaner
 497 production technologies and sustainable agricultural practices. Maximizing the use of
 498 resources through co- and by-production, restoring soil fertility of production land, and
 499 minimizing greenhouse gas emissions along the production chain are also crucial
 500 components. It is critical to consider rebound effects that occur where material or energy
 501 efficiency gains facilitate greater consumption (e.g. when increased vehicle efficiencies
 502 encourage more car travel) (Duarte et al., 2013). Typical indicators are cradle-to-cradle or
 503 "Benign by Design" approaches, reduction in resource consumption, and efficiency gains in
 504 agricultural energy production.

505 An illustrative example for resource maintenance and efficiency as an outcome of a
 506 transition experiment is the replacing of halide lamps with Light Emitting Diode lights at Yale
 507 University, United States reported by Cole and Srivastava (2013).

508 The evaluative question for this feature is: *Do the transition experiment's outputs contribute*
 509 *to overall resource maintenance and efficiency?*

510 **3.2.5 Socio-ecological stewardship and democratic governance**

511 This feature refers to arrangements that support individual and collective engagement in
 512 sustainability decision-making (Gibson et al., 2005, p. 107-111). Operationalization to
 513 municipal planning and policy-making requires participants to address related aspects in
 514 actionable knowledge (Section 3.1.2), built capacities (Section 3.1.1), accountability (3.1.3)
 515 and structural changes (Section 3.1.4). Improving governance for sustainability may involve
 516 creating and maintaining a flexible decision-making framework and fostering ongoing
 517 collaborative decision-making processes with actors at the municipal level. In addition, social
 518 inclusion, involvement and a shared sense of ownership of collective decisions as well as
 519 human-nature relations need to be ensured in all facets of everyday life through government
 520 actors, business, and civil society (Stuart et al., 2014). Experiments also function as safe
 521 operating spaces for socio-ecological innovations (Frantzeskaki and Tefrati, 2016) that can,
 522 amongst others, foster literacy for self-governance and expression of democratic beliefs in
 523 alignment with sustainability values. Typical indicators are participatory settings,
 524 collaboration among different actors, knowledge co-production, strengthened human-nature
 525 relationships, and effective public input into municipal decision-making.

526 An illustrative example for improved socio-ecological stewardship and democratic
 527 governance as an outcome of a transition experiment is the re-opening of a community
 528 center in Rotterdam, Netherlands reported by Wittmayer et al. (2014). Inhabitants of a
 529 deprived neighborhood were empowered to engage in self-maintenance of community
 530 space.

531 The evaluative question for this feature is: *Do the transition experiment's outputs build or*
 532 *support socio-ecological understanding and democratic governance?*

533 **3.2.6 Precaution and adaptation**

534 The feature of precaution and adaptation captures the importance of acknowledging
 535 uncertainty and of anticipating and avoiding unpredictable risks. Precautionary approaches,
 536 creation of learning opportunities and preparation for surprises are essential for
 537 operationalization (Gibson et al., 2005, p. 111-113). The application of this feature in the
 538 evaluation of an aquaculture operation requires actionable knowledge (Section 3.1.2), built
 539 capacities (Section 3.1.1) and structural changes (3.1.4) to reflect on uncertainties and apply
 540 adaptive approaches. Key considerations include capturing the impacts of changes in fishing
 541 practices, enhancing capacities to monitor changes over time, and generating knowledge on
 542 future demands (Vincent and Morrison-Saunders, 2013). Typical indicators are risk-averse
 543 and cautious approaches, comprehensive risk analysis, and measures that explicitly address
 544 environmental degradation.

545 An illustrative example for precaution and adaptation as an outcome of a transition
 546 experiment is reported by Voytenko et al. (2015) in an initiative to integrate use of green and
 547 blue infrastructure to cope with storm water in New Kiruna City, Sweden. Contrary to the
 548 conventional approach to use piped networks, multifunctional green areas are utilized. With
 549 regards to current and future climate change impacts and other urban challenges,
 550 knowledge and tools were also developed for integrated urban storm water management.

551 The evaluative question for this feature is: *Do the transition experiment's outputs ensure*
 552 *precaution and adaptation?*

553 **3.3 Process Features**

554 Processes are a sequence of actions conducted in sustainability transition experiments. The
 555 particular actions and their sequence are critical for creating desired outputs. Process
 556 features are structured sequence of actions, sound methodology, collaboration, reflexivity
 557 and learning, and transparency (Forrest and Wiek, 2014). Since process and outputs often
 558 become intertwined during the experimentation, performed processes are as important as
 559 the generated outputs.

560 3.3.1 Sequence of actions

561 The sequence of actions in experimentation needs to include (Bernstein et al., 2014;
 562 Karvonen and van Heur, 2014; Laakso and Lettenmeier, 2015):

- 563 (i) Defining a baseline and a goal for the interventions
- 564 (ii) Creating a specific set-up to administer interventions
- 565 (iii) Measuring the effects of the interventions against the baseline and the goal
- 566 (iv) Evaluating the effects against sustainability criteria
- 567 (v) Offering evidence-supported recommendations on how to implement the results

568 Actions include scientific activities as well as, for example, managerial tasks when
 569 administering interventions. Action (v) includes processes and mechanisms that stimulate
 570 considering the experiment from a whole system perspective (Westley and Miller, 2003).
 571 Typical indicators are the adequate planning of actions and their interference in the timeline
 572 of the experiment, the completeness of actions as well as engaging the right participants and
 573 the right information.

574 An illustrative example for a sequence of action in a transition experiment is reported by
 575 Laakso and Lettenmeier (2015). Following the quantification of household consumption and
 576 the definition of sustainable material footprints, household specific visions were co-created
 577 and roadmaps developed through backcasting. The results from household experimentation
 578 were evaluated against the co-created visions and sustainable material footprints. Finally, a
 579 “Future Workshop” was conducted with relevant practitioners and decision-makers offering
 580 evidence supported recommendation on how to mainstream solutions.

581 The evaluative question for this feature is: *Is the transition experiment structured into a*
 582 *meaningful sequence of actions?*

583 3.3.2 Sound methodology

584 Sound methodology comprises the methods that are applied in each action of the
 585 experiment (see above). The pool includes, among others, methods for intervention design
 586 (e.g. problem analysis, visioning, strategy development, etc.), assessment, monitoring and
 587 evaluation (Bernstein et al., 2014; Ceschin, 2014; Davies and Doyle, 2015). This gives
 588 emphasis to *rigorous but broad and flexible* methods that promote transformational change
 589 over conventional approaches with a narrower focus on collecting and analyzing data.
 590 Typical indicators are structured procedures for generating outputs and the adequacy of
 591 methods for the respective action.

592 An illustrative example for a sound methodology in a transition experiment can be reviewed
 593 in Davies and Doyle (2015) reporting on an experiment to transform household consumption
 594 across the Republic of Ireland and Northern Ireland. The methodology included sound
 595 methods for baseline and goal definition, intervention design, as well as monitoring and
 596 evaluation.

597 The evaluative question for this feature is: *Does the transition experiment adopt a sound*
 598 *methodology to conduct the experiment?*

599 3.3.3 Collaboration

600 Collaboration in the context of transition experiments refers to: the participants of
 601 experiments (the collaborators), the mechanisms through which collaboration is facilitated
 602 (the participatory-setting) and the modes of interactions (the intensity of collaboration)
 603 (Juujärvi and Pessa, 2013; Tams and Wadhawan, 2012; Trencher et al., 2014a).
 604 Participants of experiments vary according to the focus and phase but typically include,
 605 among others, researchers, practitioners, and the public (Brown et al., 2003; Iwaniec and
 606 Wiek, 2014; Wittmayer et al., 2014). Participants need to be carefully selected to avoid
 607 power imbalance or excluding marginalized groups from the experiment (Wittmayer and
 608 Schöpke 2014). Participatory settings are the engagement procedures including focus
 609 groups, stakeholder workshops and more dynamic processes such as participatory modeling
 610 (Bernstein et al., 2014; Liedtke et al., 2015; Schreuer et al., 2010). In the preparation and the
 611 core phase of the experiment scientific and non-scientific actors collaborate through inter-
 612 and transdisciplinary approaches. Respective modes of interactions include information
 613 sharing, consultation, collaboration, and empowerment (Bernstein et al., 2014; Vandevyvere
 614 and Nevens, 2015). This feature also captures educational settings in which students
 615 participate in the experiments (Ceschin, 2014; Trencher et al., 2014b; Wiek and Kay, 2015).
 616 Typical indicators are affiliations of participants and their roles, information flows, decision-
 617 making procedures, and interactions.

618 An illustrative example for collaboration in a transition experiment is the revitalization of
 619 public space in Phoenix, United States, as reported by Wiek et al. (2015). The experiments
 620 were designed and conducted with various external stakeholders including an elementary
 621 school, the school district, the county department on public health, and the city service
 622 department who provided funds, helped in the co-design, and were active in the
 623 implementation (e.g. painting, planting, negotiating, etc.).

624 The evaluative question for this feature is: *Does the transition experiment facilitate*
 625 *collaboration among relevant stakeholders in the experimentation process?*

626 3.3.4 Reflexivity and learning

627 Reflexivity and learning refer to the iterative analysis of all components of the experiment
 628 (Evans and Karvonen, 2014; van Mierlo and Beers, 2015). This involves the components,
 629 processes and actors involved in the experiment as well as it demands recognizing and
 630 reflecting upon the broader institutional context, issues of power, privileges, legitimacy and
 631 aspects rendering salience (Loorbach et al., 2015). Learning based on reflexivity throughout
 632 the experiment allows for changing and adapting processes to generate desired outputs
 633 (Moore et al., 2005; van Buuren and Loorbach, 2009; Vergragt and Brown, 2007). In this
 634 context, first order learning refers to changing given processes making them more efficient
 635 and effective. Second order learning involves developing new processes as well as
 636 reinterpreting the purpose and function of given activities – often crucial for transformational
 637 change. Second order learning can occur if participants with different worldviews collaborate
 638 in the experiment. Typical indicators are the presence of a shared learning agenda and
 639 dedicated points of reflections such as meetings to explicitly reflect on the experiment,
 640 review processes, as well as changes of the experimentation process.

641 An illustrative example for reflexivity and learning in a transition experiment are the activities
 642 related to the piloting of eco-innovations in Paris, France, as reported by Audet and
 643 Guyonnaud (2013). For example, the innovation experiments conducted by the Fondaterra
 644 Foundation were remodeled and framed as transition initiatives based on collaborative
 645 educational seminars to strategically promote and harness change.

646 The evaluative question for this feature is: *Does the transition experiment foster reflexivity*
 647 *and learning throughout the process?*

648 3.3.5 Transparency

649 Transparency refers to open and truthful reporting on intentions and pursued actions in the
 650 experimentation process. It includes documentation and publishing of the process, data,

651 decision-making and conclusions ensuring the possibility for all actor groups to access
652 related information (Evans and Karvonen, 2014; Iwaniec and Wiek, 2014; Ryan, 2013). It
653 also captures indication of researchers' accountability for the experimentation process.
654 Typical indicators are openly published results, reports that explicate assumptions and
655 intentions, and documentation of the decision-making process.

656 An illustrative example of transparency as part of the process of a transition experiment is to
657 explicitly highlight the underlying assumptions on which interventions in Melbourne,
658 Australia, were based, as reported by Ryan (2013). Such transparency enhancing processes
659 prevented antagonism regarding the outputs of the urban experiment amid polarized political
660 debates.

661 The evaluative question for this feature is: *Does the transition experiment ensure*
662 *transparency throughout the process?*

663 3.4 Input Features

664 Inputs are contributions to and investments in the sustainability transition experiment
665 including awareness, commitment, expertise, trust, as well as financial, and other types of
666 support (Wiek et al., 2015; Forrest and Wiek, 2014). Although inputs are often thought of as
667 prerequisites that need to be in place prior to experimentation, inputs remain of vital
668 importance throughout experimentation.

669 3.4.1 Awareness

670 Awareness refers to the ability and consciousness of participants to acknowledge the need
671 for radical real-world changes prior to and during their engagement in the experiment (Bos
672 and Brown, 2012; Nevens and Roorda, 2014). It involves the motives and intentions of
673 participants to participate and helps protect experiments from loss of momentum during later
674 phases (Moore et al., 2005; Wiek et al., 2014). Typical indicators are sustainability-related
675 track records of participants, and participants' general awareness of the sustainability issues
676 tackled by the experiment.

677 An illustrative example of awareness as input into a transition experiment is declaration of
678 the city council to become a carbon neutral city four years before related experiments were
679 initiated in the City of Ghent, Belgium, as reported by Nevens and Roorda (2014).

680 The evaluative question for this feature is: *Does the transition experiment involve*
681 *participants that are aware of the need for transformational change pursued through the*
682 *experiment?*

683 3.4.2 Commitment

684 Commitment refers to willingness, promises, positive attitudes and interests of involved
685 participants to explore "intentionally radical" instead of "incremental or entropic" changes
686 (Karvonen and van Heur, 2014, p. 387). This includes researchers and non-academic
687 participants' motivation to exceed monetary or reputational benefits and pursue
688 collaboratively taken decisions driven by intrinsic motivations to contribute to a common goal
689 (Ceschin, 2014; Moore, et al., 2005). Accountability as a transition experiment output is often
690 dependent on a critical level of initial commitment (as input feature). Typical indicators are
691 that participants' agreement to deliver tasks on time, participants' engagement in decision-
692 taking, and continuous participation in the experimentation.

693 An illustrative example of commitment as input into a transition experiment is the intrinsic
694 interests of participants in the integrated urban water management in Sydney, Australia,
695 reported by Bos and Brown (2012). Participants' commitment facilitated a meaningful
696 dialogue between different interests, which resulted in political commitment towards the
697 initiative.

698 The evaluative question for this feature is: *Does the transition experiment involve*
699 *participants committed to carrying out the experiment?*

700 **3.4.4 Expertise**

701 Expertise, including professional skills and experiences, is a critical input for sustainability
702 transition experiments (Wiek et al., 2015). It includes recognized professional skills and
703 experiential techniques to research, craft, guide, decide and judge experimentation.
704 Furthermore, it refers to reflexive capacities and abilities to learning from the experiment as
705 well as expertise in issues of ethics, transparency, and power relations (Wittmayer and
706 Schöpke 2014). Typical indicators include related work experience and academic and
707 professional degrees and training of the participants.

708 An illustrative example of expertise as input into a transition experiment is a participatory
709 technology assessment in Graz, Austria, reported by Schreuer et al. (2010). Expertise was
710 provided by professionals from the municipal department for energy, fuel cell development,
711 research institutes and an energy network – critical for designing an experiment on fuel cells.

712 The evaluative question for this feature is: *Does the transition experiment involve*
713 *participants who possess the necessary skills and knowledge to carry out the experiment?*

714 **3.4.5 Trust**

715 Trust refers to the mutual willingness to collaborate on equal footing, reconcile divergent
716 worldviews, as well as acknowledge different interests (Bernstein et al., 2014; Vandevyvere
717 and Nevens, 2015). Since experiments are particularly susceptible to failure (Nevens et al.,
718 2013), engendering trust amongst participants is important for building participants'
719 confidence in the processes and the potential outcomes of the experiment, making a
720 collaborative experiment and joint addressing of potential difficulties possible. In addition, the
721 process of co-creating knowledge and shared evaluation of the experiments demands trust
722 as a source of open, truthful and collaborative exchange, particularly as interests and
723 reputation are potentially at stake (Trencher et al., 2015). Typical indicators are participants'
724 attitudes toward other participants, ability to speak one's mind, and willingness to rely on
725 others' judgments and capacities.

726 An illustrative example of trust as input into a transition experiment is the engagement of
727 university researchers in interventions in Melbourne, Australia, as reported by Ryan (2013).
728 The implementation of future exhibitions and tours was welcomed by local councils because
729 they were incorporated into long-term visions and short-term actions proposed by an
730 institution that was seen as independent from commercial developers and the government.

731 The evaluative question for this feature is: *Does the transition experiment involve*
732 *participants who trust each other?*

733 **3.4.6 Support**

734 Support refers to structural, financial and nonfinancial resources as well as assistance from
735 public and private authorities in preparing and executing sustainability transition experiments
736 (Bos and Brown, 2012; Vandevyvere and Nevens, 2015). It also includes voluntary and in-
737 kind contributions and donation of work beyond normal obligations (Moore et al., 2005; Wiek
738 et al., 2015). Typical indicators are available funds, positions, hours of voluntary
739 contributions and endorsements from actors and institutions.

740 An illustrative example of support as input into a transition experiment is reported by
741 Frantzeskaki et al. (2014). A "Floating Pavilion" was constructed as pilot project for testing
742 social, technological and economic aspects of floating apartments that are planned for the
743 regeneration of Rotterdam's harbor (the Netherlands). Besides in-kind funding and support
744 by private companies, public authorities and research institutes, the financial investments
745 amounted to 5.5 €million.

746 The evaluative question for this feature is: *Does the transition experiment secure sufficient*
747 *support for the experimentation?*

748 3.5 Summary

749 Overall, the above scheme provides a structured appraisal to assist with sustainability
750 transition experiments becoming more effective and efficient. In addition, we intend to
751 facilitate and accelerate learning across different experiments. Since the description of the
752 evaluative scheme is generic, application to empirical experiments requires contextualizing,
753 concretizing and adapting each feature. We summarize the presented features in box 1 and
754 through instructive definitions provide tentative principles for designing sustainability
755 transition experiments.

Box 1: The tentative evaluation scheme for appraising sustainability transition experiments

Criteria Set: Outputs (I)	Criteria Set: Processes (III)
<p>Built capacities Empower participants to act sustainably in everyday decision-making and practices through educating them in cognitive, practical and interpersonal competencies and enable to internalize required skills and activate new behavioral patterns. Evaluative question: <i>Does the transition experiment build capacities in participants to generate sustainability solutions?</i></p> <p>Actionable knowledge Generate evidence-supported instructions that have been tested on effectively solving a sustainability problem within the defined experimental setting including guidelines on how to most effectively transition from the current to the desired state. Evaluative question: <i>Does the transition experiment generate actionable knowledge that provides evidence on how to generate sustainability solutions?</i></p> <p>Accountability Ensure confidence and commitment of participants to implement results generated by the experiment and their dedication to positive change. Evaluative question: <i>Does the transition experiment build confidence and commitment for generating and realizing sustainability solutions?</i></p> <p>Changes in physical structures Create new or transform existing buildings, infrastructures, technologies and products that are radically different from existing ones. Evaluative question: <i>Does the transition experiment generate physical changes that support solutions for the identified sustainability problem?</i></p> <p>Changes in social structures Create new or transform existing networks and organizations, values and norms, rules and policies, behavior and practices, and discourses that are radically different from existing ones. Evaluative question: <i>Does the transition experiment generate societal changes that support solutions for the identified sustainability problem?</i></p> <p>Transferability Create generalizable lessons learned regarding processes through to outcome of the experimentation that are applicable to different contexts. Evaluative question: <i>Does the transition experiment indicate how the sustainability solution can be transferred to different contexts?</i></p> <p>Scalability Create generalizable knowledge that facilitates the up-take of experiment results in system-wide applications Evaluative question: <i>Does the transition experiment indicate the potential for and how outputs can be scaled out to broader applications or up to higher hierarchical levels?</i></p> <p>Accounting for unintended consequences associated with up-take Reflect on and identify circumstances that have the potential to generate unintended consequences through the up-take of sustainability solutions. Evaluative question: <i>Does the transition experiments account for unintended consequences that are associated with the up-take of sustainability solutions?</i></p>	<p>Socio-ecological stewardship and democratic governance Provide arrangements that support individual and collective sustainability decision-making fostering ongoing collaborative actions, social inclusion and ownership. Evaluative question: <i>Do the transition experiment's outputs build or support socio-ecological understanding and democratic governance?</i></p> <p>Precaution and adaptation Acknowledge uncertainty and avoid uncomprehended risks, creating learning opportunities and preparing for surprises and change. Evaluative question: <i>Does the transition experiment's outputs ensure precaution and adaptation?</i></p>
<p>Criteria Set: Outcomes (II)</p> <p>Socio-ecological integrity Harmonize human well-being with the biophysical processes and elements, preventing degradation of ecosystems and reducing overall impacts and threads to the life-support system. Evaluative question: <i>Do the transition experiment's outputs strengthen socio-ecological integrity?</i></p> <p>Livelihood sufficiency and opportunity Ensure sufficient access of individuals and communities to what is needed for a decent life and create opportunities for positively exercising power and capabilities. Evaluative question: <i>Do the transition experiment's outputs enhance livelihood sufficiency and opportunity?</i></p> <p>Intra- and intergenerational equity Ensure sufficient and effective choices that reduce gaps between the rich and the poor and enhance opportunities of future generation to pursue sustainable lives. Evaluative question: <i>Do the transition experiment's outputs improve intra- and intergenerational equity?</i></p> <p>Resource maintenance and efficiency Create sustainable livelihoods for all while reducing threats that jeopardize the long-term socio-ecological integrity and cutting material and energy use per unit of benefit. Evaluative question: <i>Do the transition experiment's outputs contribute to overall resource maintenance and efficiency?</i></p>	<p>Criteria Set: Inputs (IV)</p> <p>Awareness Enable participants' consciousness of and ability to acknowledge the need for radical real-world changes prior to their engagement in the experiment. Evaluative question: <i>Does the transition experiment involve participants that are aware of the need for transformational change pursued through the experiment?</i></p> <p>Commitment Cater for willingness, promises, positive attitudes and interests of involved participants to explore intentionally radical instead of incremental changes Evaluative question: <i>Does the transition experiment involve participants committed to carrying out the experiment?</i></p> <p>Expertise Ensure expertise of participants in sustainability transition experiments including widely recognized professional skills and experiential techniques to research, craft, guide, decide and judge experimentation. Evaluative question: <i>Does the transition experiment involve participants who possess the necessary skills and knowledge to carry out the experiment?</i></p> <p>Trust Cater for mutual willingness of and between researchers and non-academic participants to rely on actions of other members of the sustainability transition experiment. Evaluative question: <i>Does the transition experiment involve participants who trust each other?</i></p> <p>Support Ensure structural, financial and nonfinancial resources as well as assistance from public and private authorities in preparing and executing sustainability transition experiments. Evaluative question: <i>Does the transition experiment secure sufficient support for the experimentation?</i></p>

755 4. Discussion

756 Although differences in transition approaches have been highlighted on the theoretical level
757 (Markard et al., 2012; van den Bergh et al., 2011), little attention has been paid to the
758 diversity of practical sustainability and transition experiments around the world (Trencher et
759 al., 2014b). Currently undertaken transition experiments come in various shapes and forms.
760 The presented evaluative scheme is designed to be applicable to a broad range of
761 sustainability transition experiment types. The presented features are not based on a single
762 theoretical interpretation of transition experiments. Rather, the scheme includes a broad
763 array of features that are of importance across different framings of sustainability transition
764 experiments. Thus, the evaluative scheme allows for comparative evaluations of various
765 experiments to identify critical success factors (cf. Forrest and Wiek, 2014, 2015). It offers a
766 coherent set of principles for designing experiments (see the instructive definitions of each
767 feature in box 1) and evaluative questions that can enhance the reflexive nature of initiatives
768 and their contribution to sustainability transitions. The following discussion is framed by the
769 four criteria that informed the development of the scheme, i.e. being generic,
770 comprehensive, operational, and formative.

771 4.1 *Is the evaluative scheme generic?*

772 Cross-case learning between and among different sustainability transition experiments
773 requires generically defined features (Macmillan et al., 2001; Rogers, 2008). The presented
774 scheme was developed with regards to transition experiments framed through various
775 approaches. The features cover a broad range of requirements intended to be applicable to
776 sustainability transition experiments independent from their specific conceptual framing.

777 Application of the scheme requires contextualization of the outlined features. While generic
778 attributes guide the evaluation independent of the context, application to a particular
779 experiment does require the integration of certain needs and context specifics (Gibson,
780 2006). The illustrative examples are intended to facilitate this process. In addition, local
781 concerns and characteristics need to be drawn from studies in similar contexts, relevant
782 public documents and integration of local knowledge. Contextualization, however, should not
783 jeopardize the common ground required for cross-case comparison. For this purpose it
784 suffices that evaluations only capture the essential characteristics of the experiment.

785 The scheme is an invitation to researchers and practitioners to engage in reflexive
786 evaluations and advance the presented features. Since the scheme is intended as a
787 “working list” of general requirements, features could be merged, subdivided, or revised. The
788 scheme is a “living” construct open to critical application, learning, and improvement. In this
789 spirit, the evaluative scheme serves as a starting point for a platform of exchange on the
790 experiences of researchers and practitioners with the evaluation of sustainability transition
791 experiments.

792 4.2 *Is the evaluative scheme comprehensive?*

793 A comprehensive evaluative scheme needs to cover the different dimensions including *all*
794 features critical to the nature of sustainability transition experiments (Forrest and Wiek,
795 2014; McLaughlin and Jordan, 2010). We adopted the established logical model of
796 evaluation to ensure basic comprehensiveness (Figure 1). The scheme is comprehensive as
797 it describes the different dimensions of the experiment: the use of resources (*inputs*) in
798 *processes* that generate *outputs* and evaluate them with regards to sustainability
799 (*outcomes*), including a tentatively comprehensive collection of critical features from a broad
800 range of experiment types.

801 The scheme will only be useful if the evaluation is rigorous. This implies applying the
802 scheme to the full extent in order to capture *all* features critical to a transition experiment and
803 to allow for cross-case comparison between different experiments. The evaluative questions

804 need to be answered with scrutiny to support honest evaluation. The objective of being
805 comprehensive also implies that sufficient reasons are being provided if features are added
806 or dismissed. All features are justified with relevant literature to reduce arbitrariness – and
807 this should be a rule for proposed changes, too. Following the presented scheme would also
808 reduce getting caught in the politics of evaluation (see e.g. Bulkeley and Betsill, 2013).
809 However, the presented scheme is only practical when there is commitment to rigorous
810 evaluation and capacity to use the results.

811 There are three limitations to the comprehensiveness of the scheme. First, it focuses on
812 experiments, even if they aim at a larger goal (sustainability transition), which requires
813 cumulative evaluations. Sustainability outcomes will be at least complementary or even
814 mutually reinforcing. Encouraging and reproducing positive effects is the intent of
815 sustainability transition experiments. However, accomplishing only a small selection of
816 outcome features will not be sufficient for leveraging sustainability. Transition experiments are
817 often conducted through transition labs. If the overall contribution of a sustainability transition
818 lab is evaluated, all outcome features need to be integrated in the immediate and long-term
819 for seeking reinforcing benefits and multiplying gains (Gibson, 2006). Thus, carefully
820 choosing the right timing for evaluation is important. However, not every type of evaluation is
821 capable of capturing time delays. Since not all downstream activities may fall within the
822 range of evaluation, the successful on-going up-take of experiments may exceed the scope
823 of evaluation timeframes. Finally, ex-post evaluation should be planned for from the start of
824 an experiment to ensure that required actions are carried out (e.g. baseline assessment).

825 Second, actors may evaluate a given experiment in different ways, depending on their
826 normative orientation and respective judgment (Smith and Raven, 2012; Leach et al., 2010).
827 The appraisals might vary depending on the framing of the experiment, too (Smith et al.,
828 2014; Fressoli et al., 2014). This applies to the outcomes – whether an experiment is
829 successful or not – as well as to the processes – whether they are appropriate and just,
830 leading to different judgments on features critical for the experiment. Processes and content
831 are intertwined in transition experiments, which means that the generated outcomes are as
832 important as the process through which they are produced (Rotmans and Loorbach, 2009;
833 Robinson 2003). Independent of the actor groups involved, vested interests, power relations,
834 and political realities will influence evaluation efforts. The presented scheme is intended to
835 facilitate a structured debate regarding the proposed features and process, functioning as a
836 guiding tool for learning. In addition, the comprehensive character of the scheme supports
837 the uncovering of issues not adequately addressed through the evaluation or the
838 experiment.

839 Third, although the presented scheme can inform the design of experiments, it does not
840 account for causal relations among different features. However, based on our experience
841 and the reviewed literature, features of one dimension may follow a logic order (see Section
842 3.1), but features of different dimensions may as well be connected through causal relations.
843 For example, a functional technology as an output of an experiment (Section 3.1.2) is
844 achieved by adopting a sound methodology (Section 3.3.1) and through collaboration
845 (Section 3.3.3), but ultimately depends on participants' awareness (Section 3.4.1) and
846 commitment (Section 3.4.2). Application to multiple experiments will allow identifying the
847 influencing factors, relations, and weights. Studies applying the scheme may also identify
848 causal mechanisms through process tracing from inputs to outcomes via intermediate
849 processes and outputs (Forrest and Wiek, 2014; George and Bennett, 2005). Such causal
850 mechanisms, plus cumulative data from multiple studies provide the basis for theory building
851 and designing further evaluative studies targeting specific hypotheses about what makes an
852 experiment succeed or fail (Yin, 2009). The focus on experiments as the smallest unit or
853 stepping stone of sustainability transitions provides possibilities to inform long-term transition
854 processes (Rotmans, 2005).

855 4.3 Is the evaluative scheme operational?

856 Operationalization is required to enable practical application of the scheme (Bornmann,
857 2013). We intend to facilitate this through typical indicators and evaluative questions.
858 Following the numbering in Figure 1, evaluators are equipped with the essential questions
859 for appraising experiments and provided with specific sources for operationalization.
860 Additional research is needed to further operationalize the scheme and provide samples of
861 exemplary operationalization.

862 The operationalization of generic features poses reflexive questions, including: “Who
863 evaluates whom and for what purpose?” We argue for the application of the scheme by core
864 members of the experiment or at least that they support external evaluation. When being
865 applied by practitioners in a utilization-focused evaluation, the scheme enhances the
866 strategic orientation, coherence and impact of the experiment (Patton, 2012). In addition,
867 participating in the process of evaluation through facilitation of data collection creates
868 dedicated points of reflection. This provides an informal opportunity for learning that
869 otherwise would not be present. For researchers, the scheme could aid evaluation of the
870 transformational potential of experiments, also enabling cross-case comparison of
871 experiments. While evaluation contributes to learning of researchers and practitioners, it
872 may also serve the increasing demands by funders for accountability. However, this creates
873 tensions between short-term accountability and long-term sustainability transitions (Regeer
874 et al. 2016). This reflects conflicts between experiments and their respective contexts (ibid).
875 Accordingly, evaluation is not a neutral, objective task, but influenced by power and interests
876 (Evans and Karvonen, 2014; Smith et al., 2014; Wamsler et al., 2014). Therefore, evaluators
877 need to avoid, for example, framing least privileged groups as beneficiaries without giving
878 them a proper say in the decision-making (Evans and Karvonen, 2014). This raises question
879 of legitimacy (in the social sphere) and accuracy and relevancy (in the scientific sphere) –
880 which call for transparency about goal and process of each evaluation.

881 Making the scheme fully operational and applicable requires to embed it into an evaluation
882 methodology, which requires coping with various challenges as indicated in a study by Wiek
883 et al (2014). Such a methodology needs to specify methods for gathering data on different
884 features as well as for analyzing and visualizing results. It needs to account for challenges
885 related to the politics of evaluation as well as ambiguity related to the purpose and outcome
886 of the evaluation. Such methodology would support coherent, yet reflexive, application of the
887 scheme to a large number of transition experiments. In addition, it would support multi-step
888 evaluation processes and coherent ways of summarizing and aggregating results.
889 Developing an evaluation methodology is a desirable next step, which needs to be informed
890 by application of the scheme.

891 4.4 Is the evaluative scheme formative?

892 An evaluative scheme needs to support sustainability transition experiments to become
893 more effective and efficient. The application of the presented scheme as a formative tool
894 therefore intends to improve designing experiments and improving ongoing experimentation.
895 When the scheme is being used as guideline for designing experiments (*ex-ante evaluation*),
896 evaluators can derive design principles from Box 1. The scheme functions as a checklist that
897 channels the attention to essential items that need to be evaluated regarding their relevance
898 for the experiment in question (e.g. which inputs need to be secured and what processes
899 have to be carried out to generate outputs). Ex-ante evaluation allows the appraisal of
900 prospective outputs with regards to their sustainability outcomes (following the big arrows in
901 Figure 1).

902 The scheme can also be applied to completed experiments (*ex-post evaluation*). Evaluators
903 can utilize the evaluative questions provided in Box 1. The scheme provides orientation for
904 the evaluation by starting from the outputs evaluating them with regards to sustainability
905 (outcomes), and working ‘backwards’ by tracking processes and inputs. Carefully choosing

906 the right timing for evaluation is as important as the evaluation itself since an untimely
907 appraisal might not do justice to an experiment and “out-score” its accomplishments. *Ex-post*
908 *evaluation* should be planned for from the start of an experiment to support experiment
909 design and implementation (e.g. to ensure attention to the need to conduct a baseline
910 assessment).

911 In case of *formative evaluation* for improving on-going sustainability transition experiments,
912 the design guidelines and evaluative questions presented in Box 1 are equally important. It
913 offers the possibility to regularly appraise progress and shortcomings of experiments. To
914 improve design and performance, evaluators can start at any evaluative dimension (Figure
915 1). While they reflect on the tentative design principles as well as on the evaluative
916 questions, they also have to simultaneously work backwards to the inputs, and track
917 forwards towards the targeted outcomes.

918 In addition, extending formative evaluation beyond solely improving experiments efficiency
919 and effectiveness requires re-conceptualizing their contribution to overall societal change
920 processes. This demands participants to engage in open and reflexive processes
921 considering the goals and procedures of an experiment and facilitate cross-case comparison
922 between different experiments. Finally, the presented scheme is only formative if there is
923 commitment to evaluation and capacity to use the outcomes. Evaluation requires financial
924 and human resources and, ideally, is already planned for when designing the experiment
925 proposal.

926 **5. Conclusion**

927 This article presents a tentative evaluative scheme for appraising individual sustainability
928 transition experiments and facilitating their cross-case comparison. We propose a set of
929 characteristics the scheme requires to be broadly applicable, practical, comprehensive and
930 used to improve the performance of contemporary and future experiments. Following the
931 basic logic model of evaluation, we reviewed sustainability transition experiments to identify
932 features in the evaluative dimensions of inputs, processes, outputs and outcomes. Each
933 feature was described (definitions), exemplified (indicators), illustrated (examples) and
934 justified. The resulting evaluative scheme in general and with the discussed limitations is (i)
935 *generic*, i.e., applicable to different types of sustainability transition experiment; (ii)
936 *comprehensive*, i.e., captures all critical features of experiments; (iii) *operational*, i.e., ready
937 for application; and (iv) *formative*, i.e., supports experiments in becoming more effective and
938 efficient. While the presented scheme is neither finished nor a recipe for success, it serves
939 as a basis for structured reflection and strategizing in support of experiments that help
940 society to transition towards sustainability. We emphasize the need for applying the scheme
941 to facilitate learning and accelerate progress across different experiments as well as for
942 advancing evaluation of sustainability transitions. We encourage future research projects
943 that apply, question and improve this framework to expand the evidence base for designing
944 and conducting the next generation of sustainability transition experiments.

945

946 **Acknowledgements**

947 We are grateful for the comments from David Jacobs, Paula Kivimaa, Adrian Smith and
948 three reviewers on previous versions of this article. The final version also benefitted from
949 suggestions and inputs at four international conferences including the INOGOV Workshop
950 2015 in Helsinki, Finland, the ESEE 2015 Conference in Leeds, UK, the IST 2015
951 Conference in Brighton, UK, and the Transformation 2015 Conference in Stockholm,
952 Sweden.

6. References

- Adkins, E., Eapen, S., Kaluwile, F., Nair, G., & Modi, V. (2010). Off-grid energy services for the poor: Introducing LED lighting in the Millennium Villages Project in Malawi. *Energy Policy*, 38(2), 1087–1097. doi:10.1016/j.enpol.2009.10.061
- Audet, R., & Guyonnaud, M.-F. (2013). Transition in practice and action in research. A French case study in piloting eco-innovations. *Innovation: The European Journal of Social Science Research*, 26(4), 398–415. doi:10.1080/13511610.2013.850019
- Avelino, F., Rotmans, J. (2009). Power in Transition: An Interdisciplinary Framework to Study Power in Relation to Structural Change. *Eur. J. Soc. Theory* 12, 543–569. doi:10.1177/1368431009349830
- Bai, X., Roberts, B., & Chen, J. (2010). Urban sustainability experiments in Asia: patterns and pathways. *Environmental Science & Policy*, 13(4), 312–325. doi:10.1016/j.envsci.2010.03.011
- Banerjee, A. V., Duflo, E., Glennerster, R., & Kothari, D. (2010). Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *BMJ: British Medical Journal*, 340, c2220. doi:10.1136/bmj.c2220
- Bernstein, M. J., Wiek, A., Brundiers, K., Pearson, K., Minowitz, A., Kay, B., & Golub, A. (2014). Mitigating urban sprawl effects: a collaborative tree and shade intervention in Phoenix, Arizona, USA. *Local Environment*, (January 2015), 1–18. doi:10.1080/13549839.2014.965672
- 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), 217–233. doi:10.1002/asi.22803
- Bos, J. J., & Brown, R. R. (2012). Governance experimentation and factors of success in socio-technical transitions in the urban water sector. *Technological Forecasting and Social Change*, 79(7), 1340–1353. doi:10.1016/j.techfore.2012.04.006
- Bos, J. J., Brown, R. R., & Farrelly, A. M. (2013). A design framework for creating social learning situations. *Global Environmental Change*, 23(2), 398–412. doi:10.1016/j.gloenvcha.2012.12.003
- Bos, J. J., Brown, R. R., & Farrelly, M. A. (2015). Building networks and coalitions to promote transformational change: Insights from an Australian urban water planning case study. *Environmental Innovation and Societal Transitions*, 15, 11–25. doi:10.1016/j.eist.2014.10.002
- Brown, H. S., Vergragt, P., Green, K., & Berchicci, L. (2003). Learning for Sustainability Transition through Bounded Socio-technical Experiments in Personal Mobility. *Technology Analysis & Strategic Management*, 15(3), 291–315. doi:10.1080/09537320310001601496

- Brown, H. S., & Vergragt, P. J. (2008). Bounded socio-technical experiments as agents of systemic change: The case of a zero-energy residential building. *Technological Forecasting and Social Change*, 75(1), 107–130. doi:10.1016/j.techfore.2006.05.014
- Bulkeley, H., Betsill, M.M. (2013). Revisiting the urban politics of climate change. *Env. Polit.* 22, 136–154. doi:10.1080/09644016.2013.755797
- Castán Broto, V., Bulkeley, H. (2013). A survey of urban climate change experiments in 100 cities. *Glob. Environ. Change* 23, 92–102. doi:10.1016/j.gloenvcha.2012.07.005
- Ceschin, F. (2014). How the Design of Socio-technical Experiments Can Enable Radical Changes for Sustainability. *International Journal of Design*, 8(3), 1–21.
- Cole, C., & Srivastava, C. (2013). Energy Blitz Leads to Measured Reductions on Campus: Students Embrace Campus as a Living Lab at Yale. *Sustainability: The Journal of Record*, 6(1), 37–41. doi:10.1089/SUS.2013.9893
- Davies, A. R., & Doyle, R. (2015). Transforming Household Consumption: From Backcasting to HomeLabs Experiments. *Annals of the Association of American Geographers*, 105(2), 425–436. doi:10.1080/00045608.2014.1000948
- Davies, A. R., Doyle, R., & Pape, J. (2012). Future visioning for sustainable household practices: spaces for sustainability learning? *Area*, 44(1), 54–60. doi:10.1111/j.1475-4762.2011.01054.x
- Duarte, C. G., Gaudreau, K., Gibson, R. B., & Malheiros, T. F. (2013). Sustainability assessment of sugarcane-ethanol production in Brazil: A case study of a sugarcane mill in São Paulo state. *Ecological Indicators*, 30, 119–129. doi:10.1016/j.ecolind.2013.02.011
- Evans, J., & Karvonen, A. (2011). Living Laboratories for Sustainability: Exploring the Politics and Epistemology of Urban Transition. In H. Bulkeley, V. Castán Broto, M. Hodson, & S. Marvin (Eds.), *Cities and Low Carbon Transition* (pp. 126–141.). London: Routledge.
- Evans, J., & Karvonen, A. (2014). “Give Me a Laboratory and I Will Lower Your Carbon Footprint!” - Urban Laboratories and the Governance of Low-Carbon Futures. *International Journal of Urban and Regional Research*, 38(2), 413–430. doi:10.1111/1468-2427.12077
- Farrelly, M. A., & Brown, R. R. (2011). Rethinking urban water management: Experimentation as a way forward? *Global Environmental Change*, 21(2), 721–732. doi:10.1016/j.gloenvcha.2011.01.007
- Ferguson, B. C., Brown, R. R., & Deletic, A. (2013). A Diagnostic Procedure for Transformative Change Based on Transitions, Resilience, and Institutional Thinking. *Ecology and Society*, 18(4), art57. doi:10.5751/ES-05901-180457
- Forrest, N., & Wiek, A. (2014). Learning from success—Toward evidence-informed sustainability transitions in communities. *Environmental Innovation and Societal Transitions*. doi:10.1016/j.eist.2014.01.003

- Forrest, N., & Wiek, A. (2015). Success factors and strategies for sustainability transitions of small-scale communities – Evidence from a cross-case analysis. *Environmental Innovation and Societal Transitions*, 1–19. doi:10.1016/j.eist.2015.05.005
- Frantzeskaki, N., & Kabisch, N. (2016). Designing a knowledge co-production operating space for urban environmental governance lessons from Rotterdam, Netherlands and Berlin, Germany. *Environmental Science & Policy*, (2015), 1–9. doi:10.1016/j.envsci.2016.01.010
- Frantzeskaki, N., & Tefrati, N. (2016). A transformative vision unlocks the innovative potential of Aberdeen City. In *Theory and practice of governance of urban sustainability transitions*. UK: Springer.
- Frantzeskaki, N., Wittmayer, J., & Loorbach, D. (2014). The role of partnerships in “realising” urban sustainability in Rotterdam’s City Ports Area, The Netherlands. *Journal of Cleaner Production*, 65, 406–417. doi:10.1016/j.jclepro.2013.09.023
- Geels, F., Raven, R. (2006). Non-linearity and Expectations in Niche-Development Trajectories: Ups and Downs in Dutch Biogas Development (1973–2003). *Technol. Anal. Strateg. Manag.* 18, 375–392. doi:10.1080/09537320600777143
- George, A. L., & Bennett, A. (2005). *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.
- Gibson, R. B. (2006). Sustainability assessment: basic components of a practical approach. *Impact Assessment and Project Appraisal*, 24(3), 170–182. doi:10.3152/147154606781765147
- Gibson, R. B., Hassan, S., Holtz, S., Tansey, J., & Whitelaw, G. (2005). *Sustainability assessment: Criteria and Processes*. London: Earthscan.
- Hart, D. D., Bell, K. P., Lindenfeld, L. A., Jain, S., Johnson, T. R., Ranco, D., & McGill, B. (2015). Strengthening the role of universities in addressing sustainability challenges: the Mitchell Center for Sustainability Solutions as an institutional experiment. *Ecology and Society*, 20(2), art4. doi:10.5751/ES-07283-200204
- Iwaniec, D., & Wiek, A. (2014). Advancing Sustainability Visioning Practice in Planning—The General Plan Update in Phoenix, Arizona. *Planning Practice & Research*, (November), 1–26. doi:10.1080/02697459.2014.977004
- Juujärvi, S., & Pessa, K. (2013). Actor Roles in an Urban Living Lab: What Can We Learn from Suurpelto, Finland? *Technology*, (November), 22–27.
- Karvonen, A., & van Heur, B. (2014). Urban Laboratories: Experiments in Reworking Cities. *International Journal of Urban and Regional Research*, 38(2), 379–392. doi:10.1111/1468-2427.12075
- König, A. (2015). Towards systemic change: on the co-creation and evaluation of a study programme in transformative sustainability science with stakeholders in Luxembourg. *Current Opinion in Environmental Sustainability*, 16, 89–98. doi:10.1016/j.cosust.2015.08.006

- Kemp, R., Schot, J., Hoogma, R. (1998). Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management. *Technol. Anal. Strateg. Manag.* 10, 175–198. doi:10.1080/09537329808524310
- Laakso, S., & Lettenmeier, M. (2015). Household-level transition methodology towards sustainable material footprints. *Journal of Cleaner Production*, 1–8. doi:10.1016/j.jclepro.2015.03.009
- Lamorgese, L., & Geneletti, D. (2013). Sustainability principles in strategic environmental assessment: A framework for analysis and examples from Italian urban planning. *Environmental Impact Assessment Review*, 42, 116–126. doi:10.1016/j.eiar.2012.12.004
- Lang, D.J., Wiek, A., Bergmann, M., Stauffacher, M., Martens, P., Moll, P., Swilling, M., Thomas, C.J. (2012). Transdisciplinary research in sustainability science: practice, principles, and challenges. *Sustain. Sci.* 7, 25–43. doi:10.1007/s11625-011-0149-x
- Larson, K. L., Wiek, A., & Withycombe Keeler, L. (2013). A comprehensive sustainability appraisal of water governance in Phoenix, AZ. *Journal of Environmental Management*, 116, 58–71. doi:10.1016/j.jenvman.2012.11.016
- Leach, M., Scoones, I., Stirling, A. (2010). *Dynamic Sustainabilities: Technology, Environment, Social Justice*. Earthscan, London. doi:10.4324/9781849775069
- Leslie, H.M., Basurto, X., Nenadovic, M., Sievanen, L., Cavanaugh, K.C., Cota-Nieto, J.J., Erisman, B.E., Finkbeiner, E., Hinojosa-Arango, G., Moreno-Báez, M., Nagavarapu, S., Reddy, S.M.W., Sánchez-Rodríguez, A., Siegel, K., Ulibarria-Valenzuela, J.J., Weaver, A.H., Aburto-Oropeza, O. (2015). Operationalizing the social-ecological systems framework to assess sustainability. *Proc. Natl. Acad. Sci. U. S. A.* 112, 5979–84. doi:10.1073/pnas.1414640112
- Liedtke, C., Baedeker, C., Hasselkuß, M., Rohn, H., & Grinewitschus, V. (2015). User-integrated innovation in Sustainable LivingLabs: an experimental infrastructure for researching and developing sustainable product service systems. *Journal of Cleaner Production*, 97, 106–116. doi:10.1016/j.jclepro.2014.04.070
- Loorbach, D., Frantzeskaki, N., & Lijnis Huffenreuter, R. (2015). Transition Management: Taking Stock from Governance Experimentation. *Journal of Corporate Citizenship*, 2015(58), 48–66. doi:10.9774/GLEAF.4700.2015.ju.00008
- Loorbach, D., & Rotmans, J. (2010). The practice of transition management: Examples and lessons from four distinct cases. *Futures*, 42(3), 237–246. doi:10.1016/j.futures.2009.11.009
- Loorbach, D. (2010). Transition Management for Sustainable Development: A Prescriptive, Complexity-Based Governance Framework. *Governance* 23, 161–183. doi:10.1111/j.1468-0491.2009.01471.x
- Luederitz, C., Lang, D. J., & Von Wehrden, H. (2013). A systematic review of guiding principles for sustainable urban neighborhood development. *Landscape and Urban Planning*, 118, 40–52. doi:10.1016/j.landurbplan.2013.06.002

- Macmillan, S., Road, C., Steele, J., Austin, S., Engineering, B., Kirby, P., & Spence, R. (2001). Development and verification of a generic framework for conceptual design. *Design Studies*, 22(2), 169–191.
- Markard, J., Raven, R., & Truffer, B. (2012). Sustainability transitions: An emerging field of research and its prospects. *Research Policy*, 41(6), 955–967.
doi:10.1016/j.respol.2012.02.013
- Marshall, G.R. (2015). A social-ecological systems framework for food systems research: accommodating transformation systems and their products. *Int. J. Commons* 9, 881.
doi:10.18352/ijc.587
- McAlpine, C. A., Seabrook, L. M., Ryan, J. G., Feeney, B. J., Ripple, W. J., Ehrlich, A. H., & Ehrlich, P. R. (2015). Transformational change: creating a safe operating space for humanity. *Ecology and Society*, 20(1), art56. doi:10.5751/ES-07181-200156
- McCormick, K., Anderberg, S., Coenen, L., & Neij, L. (2013). Advancing sustainable urban transformation. *Journal of Cleaner Production*, 50, 1–11.
doi:10.1016/j.jclepro.2013.01.003
- McGinnis, M.D., Ostrom, E. (2014). Social-ecological system framework: initial changes and continuing challenges. *Ecol. Soc.* 19. doi:10.5751/ES-06387-190230
- McLaughlin, J. A., & Jordan, G. B. (2010). Using logic models. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (3rd ed., pp. 55–88). San Francisco: Jossey-Bass.
- Miller, T., Goodling, E., Herrington, C., & Devlin, J. (2015). The Community Watershed Stewardship Program: experiments in engagement and equity in Portland, OR. *Current Opinion in Environmental Sustainability*, 17, 30–35.
doi:10.1016/j.cosust.2015.08.008
- Moloney, S., & Horne, R. (2015). Low Carbon Urban Transitioning: From Local Experimentation to Urban Transformation? *Sustainability*, 7(3), 2437–2453.
doi:10.3390/su7032437
- Moore, J., Pagani, F., Quayle, M., Robinson, J., Sawada, B., Spiegelman, G., & Van Wynsberghe, R. (2005). Recreating the university from within: Collaborative reflections on the University of British Columbia's engagement with sustainability. *International Journal of Sustainability in Higher Education*, 6(1), 65–80.
doi:10.1108/14676370510573140
- Moore, M.-L., Tjornbo, O., Enfors, E., Knapp, C., Hodbod, J., Baggio, J. A., & Norström, A. (2014). Studying the complexity of change: toward an analytical framework for understanding deliberate social-ecological transformations. *Ecology and Society*, 19(4).
- Nevens, F., Frantzeskaki, N., Gorissen, L., & Loorbach, D. (2013). Urban Transition Labs: co-creating transformative action for sustainable cities. *Journal of Cleaner Production*, 50, 111–122. doi:10.1016/j.jclepro.2012.12.001

- Nevens, F., & Roorda, C. (2014). A climate of change: A transition approach for climate neutrality in the city of Ghent (Belgium). *Sustainable Cities and Society*, 10, 112–121. doi:10.1016/j.scs.2013.06.001
- Ostrom, E. (2009). A general framework for analyzing sustainability of social-ecological systems. *Science*, 325(5939), 419–22. doi:10.1126/science.1172133
- Ostrom, E. (2007). A diagnostic approach for going beyond panaceas. *Proc. Natl. Acad. Sci. U. S. A.* 104, 15181–7. doi:10.1073/pnas.0702288104
- Ostrom, E., Cox, M. (2010). Moving beyond panaceas: a multi-tiered diagnostic approach for social-ecological analysis. *Environ. Conserv.* 37, 451–463. doi:10.1017/S0376892910000834
- Pahl-Wostl, C. (2007). Transitions towards adaptive management of water facing climate and global change. *Water Resources Management*, 21(1), 49–62. doi:10.1007/s11269-006-9040-4
- Patton, M.Q. (2012). *Essentials of utilisation-focused evaluation*. Thousand Oaks, CA: Sage
- Parris, T. M., & Kates, R. W. (2003). Characterizing and Measuring Sustainable Development. *Annual Review of Environment and Resources*, 28(1), 559–586. doi:10.1146/annurev.energy.28.050302.105551
- Partelow, S. (2015). Coevolving Ostrom's social-ecological systems (SES) framework and sustainability science: four key co-benefits. *Sustain. Sci.* doi:10.1007/s11625-015-0351-3
- Partelow, S., Boda, C. (2015). A modified diagnostic social-ecological system framework for lobster fisheries: Case implementation and sustainability assessment in Southern California. *Ocean Coast. Manag.* 114, 204–217. doi:10.1016/j.ocecoaman.2015.06.022
- Raven, R., Bosch, S. Van den, & Weterings, R. (2010). Transitions and strategic niche management: towards a competence kit for practitioners. *International Journal of Technology Management*, 51(1), 57. doi:10.1504/IJTM.2010.033128
- Raven, R. P. J. M., Verbong, G. P. J., Schilpzand, W. F., & Witkamp, M. J. (2011). Translation mechanisms in socio-technical niches: a case study of Dutch river management. *Technology Analysis & Strategic Management*, 23(10), 1063–1078. doi:10.1080/09537325.2011.621305
- Regeer, B.J., de Wildt-Liesveld, R., van Mierlo, B., Bunders, J.F.G. (2016). Exploring ways to reconcile accountability and learning in the evaluation of niche experiments. *Evaluation* 22, 6–28. doi:10.1177/1356389015623659
- Robinson, J. (2003). Future subjunctive: backcasting as social learning. *Futures* 35, 839–856. doi:10.1016/S0016-3287(03)00039-9
- Robinson, J., Burch, S., Talwar, S., O'Shea, M., Walsh, M. (2011). Envisioning sustainability: Recent progress in the use of participatory backcasting approaches for sustainability research. *Technol. Forecast. Soc. Change* 78, 756–768. doi:10.1016/j.techfore.2010.12.006

- Rogers, P. J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*, 14(1), 29–48. doi:10.1177/1356389007084674
- Rossi, P. H., Lipsey, W. M., & Freeman, H. E. (2004). *Evaluation: A systematic approach*. Thousand Oaks: SAGE Publications, Inc.
- Rotmans, J. (2005). *Societal Innovation: between dream and reality lies complexity*. Rotterdam.
- Rotmans, J., Loorbach, D. (2009). Complexity and Transition Management. *J. Ind. Ecol.* 13, 184–196. doi:10.1111/j.1530-9290.2009.00116.x
- Ryan, C. (2013). Eco-Acupuncture: designing and facilitating pathways for urban transformation, for a resilient low-carbon future. *Journal of Cleaner Production*, 50, 189–199. doi:10.1016/j.jclepro.2012.11.029
- Schreuer, A., Ornetzeder, M., & Rohracher, H. (2010). Negotiating the local embedding of socio-technical experiments: a case study in fuel cell technology. *Technology Analysis & Strategic Management*, 22(6), 729–743. doi:10.1080/09537325.2010.496286
- Seyfang, G., Longhurst, N. (2016). What influences the diffusion of grassroots innovations for sustainability? Investigating community currency niches. *Technol. Anal. Strateg. Manag.* 28, 1–23. doi:10.1080/09537325.2015.1063603
- Shah, S. H., & Gibson, R. B. (2013). Large dam development in India: sustainability criteria for the assessment of critical river basin infrastructure. *International Journal of River Basin Management*, 11(1), 33–53. doi:10.1080/15715124.2012.754445
- Smith, A. (2007). Translating Sustainabilities between Green Niches and Socio-Technical Regimes. *Technology Analysis & Strategic Management*, 19(4), 427–450. doi:10.1080/09537320701403334
- Smith, A., Raven, R. (2012). What is protective space? Reconsidering niches in transitions to sustainability. *Res. Policy* 41, 1025–1036. doi:10.1016/j.respol.2011.12.012
- Smith, A., Fressoli, M., & Thomas, H. (2014). Grassroots innovation movements: challenges and contributions. *Journal of Cleaner Production*, 63, 114–124. doi:10.1016/j.jclepro.2012.12.025
- Schot, J., Geels, F.W. (2008). Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy. *Technol. Anal. Strateg. Manag.* 20, 537–554. doi:10.1080/09537320802292651
- Steffen, W., Richardson, K., Rockstrom, J., Cornell, S. E., Fetzer, I., Bennett, E. M., ... Sorlin, S. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science*, 1259855. doi:10.1126/science.1259855
- Stuart, J., Collins, P., Alger, M., & Whitelaw, G. (2014). Embracing sustainability: the incorporation of sustainability principles in municipal planning and policy in four mid-sized municipalities in Ontario, Canada. *Local Environment*, (October), 1–22. doi:10.1080/13549839.2014.936844

- Taanman, M. (2014). Looking for transitions. Monitoring approach for sustainable transition programmes. Rotterdam: cpibooks.
- Tams, S., & Wadhawan, M. (2012). Innovation Labs: Tackling sustainability through systemic collaboration. *The Bath Perspective Magazine*, 17, 10–15.
- Trencher, G., Bai, X., Evans, J., McCormick, K., & Yarime, M. (2014). University partnerships for co-designing and co-producing urban sustainability. *Global Environmental Change*, 28, 153–165. doi:10.1016/j.gloenvcha.2014.06.009
- Trencher, G., Rosenberg Daneri, D., McCormick, K., Terada, T., Petersen, P., Yarime, M., & Kiss, B. (2016). The role of students in the co-creation of transformational knowledge and sustainability experiments: Experiences from Sweden, Japan and the USA. In W., Filho & L., Brandli (Eds.), *Engaging Stakeholders in Education for Sustainable Development at the University Level*. Berlin: Springer. doi: 10.1007/978-3-319-26734-0_13
- Trencher, G., Terada, T., Yarime, M. (2015). Student participation in the co-creation of knowledge and social experiments for advancing sustainability: experiences from the University of Tokyo. *Curr. Opin. Environ. Sustain.* 16, 56–63. doi:10.1016/j.cosust.2015.08.001
- Trencher, G., Yarime, M., McCormick, K. B., Doll, C. N. H., & Kraines, S. B. (2014). Beyond the third mission: Exploring the emerging university function of co-creation for sustainability. *Science and Public Policy*, 41(2), 151–179. doi:10.1093/scipol/sct044
- Van Buuren, A., & Loorbach, D. (2009). Policy innovation in isolation? *Public Management Review*, 11(3), 375–392. doi:10.1080/14719030902798289
- Van den Bergh, J. C. J. M., Truffer, B., & Kallis, G. (2011). Environmental innovation and societal transitions: Introduction and overview. *Environmental Innovation and Societal Transitions*, 1(1), 1–23. doi:10.1016/j.eist.2011.04.010
- Van den Bosch, S. (2010). *Transition Experiments. Exploring societal changes towards sustainability*. Erasmus Universiteit Rotterdam.
- Van der Laak, W. W. M., Raven, R. P. J. M., & Verbong, G. P. J. (2007). Strategic niche management for biofuels: Analysing past experiments for developing new biofuel policies. *Energy Policy*, 35(6), 3213–3225. doi:10.1016/j.enpol.2006.11.009
- Van Mierlo, B., & Beers, P. J. (2015). Reflexivity and learning in the context of system innovation; Prying loose entangled concepts. In *International Sustainability Transition Conference 2015*.
- van Mierlo, B., B. Regeer et al. (2010). *Reflexieve monitoring in actie. Handvatten voor de monitoring van systeeminnovatieprojecten. [Reflexive Monitoring in Action. Guidelines for the monitoring of system innovation projects]*. Oisterwijk: Boxpress.
- Vandevyvere, H., & Nevens, F. (2015). Lost in Transition or Geared for the S-Curve? An Analysis of Flemish Transition Trajectories with a Focus on Energy Use and Buildings. *Sustainability*, 7(3), 2415–2436. doi:10.3390/su7032415

- Vergragt, P. J., & Brown, H. S. (2007). Sustainable mobility: from technological innovation to societal learning. *Journal of Cleaner Production*, 15(11-12), 1104–1115. doi:10.1016/j.jclepro.2006.05.020
- Vergragt, P. J., & Brown, H. S. (2012). The challenge of energy retrofitting the residential housing stock: grassroots innovations and socio-technical system change in Worcester, MA. *Technology Analysis & Strategic Management*, 24(4), 407–420. doi:10.1080/09537325.2012.663964
- Vincent, I. V., & Morrison-Saunders, a. (2013). Applying sustainability assessment thinking to a community-governed development: a sea cucumber farm in Madagascar. *Impact Assessment and Project Appraisal*, 31(3), 208–213. doi:10.1080/14615517.2013.773720
- Vogt, J.M., Epstein, G.B., Mincey, S.K., Fischer, B.C., McCord, P. (2015). Putting the “E” in SES: unpacking the ecology in the Ostrom social-ecological system framework. *Ecol. Soc.* 20, art55. doi:10.5751/ES-07239-200155
- Voytenko, Y., McCormick, K., Evans, J., & Schliwa, G. (2015). Urban living labs for sustainability and low carbon cities in Europe: towards a research agenda. *Journal of Cleaner Production*, 1–10. doi:10.1016/j.jclepro.2015.08.053
- Wamsler, C., Luederitz, C., & Brink, E. (2014). Local levers for change: Mainstreaming ecosystem-based adaptation into municipal planning to foster sustainability transitions. *Global Environmental Change*, 29, 189–201. doi:10.1016/j.gloenvcha.2014.09.008
- Westley, F., Antadze, N., Riddell, D. J., Robinson, K., & Geobey, S. (2014). Five Configurations for Scaling Up Social Innovation: Case Examples of Nonprofit Organizations From Canada. *The Journal of Applied Behavioral Science*, 50(3), 234–260. doi:10.1177/0021886314532945
- Westley, F., Olsson, P., Folke, C., Homer-Dixon, T., Vredenburg, H., Loorbach, D., ... Leeuw, S. (2011). Tipping Toward Sustainability: Emerging Pathways of Transformation. *Ambio*, 40(7), 762–780. doi:10.1007/s13280-011-0186-9
- Westley, F. R., & Miller, P. (2003). *Experiments in Consilience: Integrating Social and Scientific Responses to Biodiversity Conservation Challenges*. Island Press.
- Wiek, A., & Kay, B. (2015). Learning while transforming: solution-oriented learning for urban sustainability in Phoenix, Arizona. *Current Opinion in Environmental Sustainability*, 16, 29–36. doi:10.1016/j.cosust.2015.07.001
- Wiek, A., Kay, B., & Forrest, N. (2015). Worth the Trouble?! An Evaluative Scheme for Urban Sustainability Transition Labs (USTL) and an Application to the USTL in Phoenix, Arizona. In N. Frantzeskaki, L. Coenen, V. Castan Broto, & D. Loorbach (Eds.), *Urban Sustainability Transitions*. Routledge.
- Wiek, A., Talwar, S., O’Shea, M., & Robinson, J. (2014). Toward a methodological scheme for capturing societal effects of participatory sustainability research. *Research Evaluation*, 23(2), 117–132. doi:10.1093/reseval/rvt031

Wittmayer, J. M., & Schöpke, N. (2014). Action, research and participation: roles of researchers in sustainability transitions. *Sustainability Science*, 483–496. doi:10.1007/s11625-014-0258-4

Wittmayer, J. M., Schöpke, N., van Steenbergen, F., & Omann, I. (2014). Making sense of sustainability transitions locally: how action research contributes to addressing societal challenges. *Critical Policy Studies*, 8(4), 465–485. doi:10.1080/19460171.2014.957336

Yin, R. K. (2009). *Case Study Research: Design and Methods* (4th ed.). Thousand Oaks: Sage Publications.

Supplementary Material A

Table 1 presents an exemplary subset of the literature that was pooled and selected as part of the literature review. In total, 61 unique case studies were used for developing the tentative evaluative scheme for sustainability transition experiments. The reviewed literature can be categorized according to seven sustainability laboratories, including (urban) transition labs, socio-technical experiments, (urban) living labs, homelabs, campus as laboratory, social innovation labs, and urban sustainability transition labs. The whole body of literature we draw on for developing each evaluative feature is referenced in Section 3 of the present article.

Table 1: Overview of different Sustainability Transition Laboratories

Sustainability Transition Laboratory	Focus	Underlying Concepts	Exemplary Literature
(Urban) Transition Labs	Various	Transition Management; organizational learning; action research; transdisciplinary	Farrelly and Brown, 2011; Loorbach and Rotmans, 2010; Nevens et al., 2013; Wittmayer et al., 2014
Socio-Technical Experiments	Innovation mainstreaming	Strategic niche management; innovation studies; organizational learning; product service systems; transition management	Brown et al., 2003; Ceschin, 2014; Quist et al., 2011; Schreuer et al., 2010; Vergragt and Brown, 2012
(Urban) Living Labs	Industry and research institutes	Product service systems; transdisciplinary research; action research; innovation studies	Audet and Guyonnaud, 2013; Evans and Karvonen, 2014; McCormick and Kiss, 2015; Ryan, 2013; Voytenko et al., 2015
HomeLabs	Everyday practices	Practice-oriented participatory; second-order learning; organizational learning	Davies and Doyle, 2015; Davies et al., 2012; Laakso and Lettenmeier, 2015; Liedtke et al., 2015
Campus as Laboratory	Universities	Community-based action research; transdisciplinary research; organizational learning	Abbott, 2014; Hart et al., 2015; Lang and Wiek, 2013; Robinson et al., 2013; Rojas et al., 2007
Social Innovation Labs	Grassroots movements, communities	Networks of transformational agency; changes in everyday life	Avelino et al., 2014; Seyfang and Longhurst, 2015; Smith et al., 2015; Westely et al., 2014
Urban Sustainability Transition Labs	Urban environments	Transformational sustainability research; transdisciplinary research; intervention studies	Bernstein et al., 2014; Forrest and Wiek, 2015; Wiek and Kay, 2015; Wiek et al., 2015, 2012

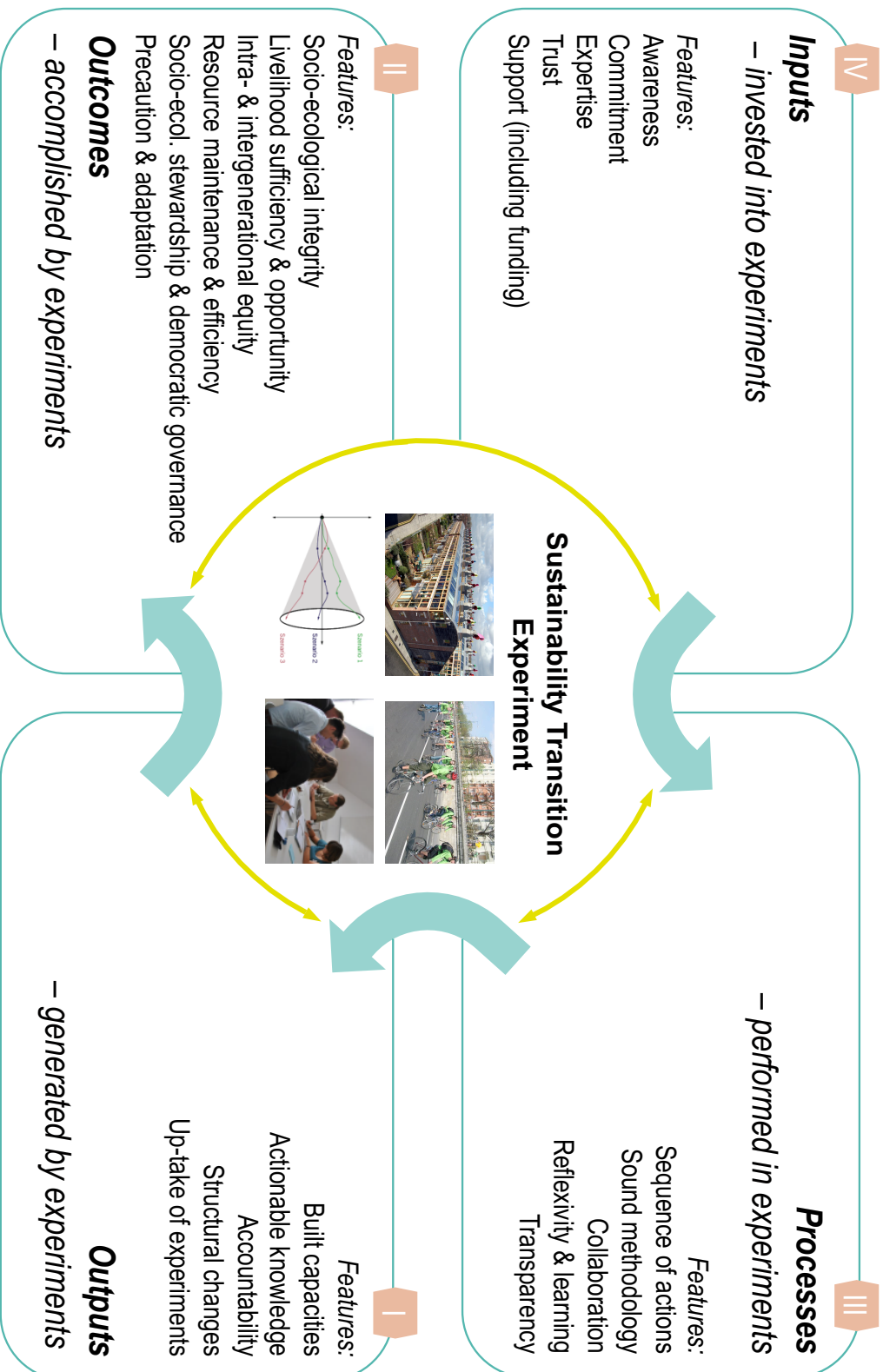
References

- Abbott, A., 2014. The university experiment: Campus as laboratory. *Nature* 514, 288–91. doi:10.1038/514288a
- Audet, R., Guyonnaud, M.-F., 2013. Transition in practice and action in research. A French case study in piloting eco-innovations. *Innov. Eur. J. Soc. Sci. Res.* 26, 398–415. doi:10.1080/13511610.2013.850019
- Avelino, F., Wittmayer, J., Haxeltine, A., Kemp, R., O’Riordan, T., Weaver, P., Loorbach, D., Rotmans, J., 2014. Game Changers and Transformative Social Innovation. The Case of the Economic Crisis and the New Economy, TRANSIT working paper, TRANSIT: EU SSH.2013.3.2-1 Grant agreement no: 613169.
- Bernstein, M.J., Wiek, A., Brundiars, K., Pearson, K., Minowitz, A., Kay, B., Golub, A., 2014. Mitigating urban sprawl effects: a collaborative tree and shade intervention in Phoenix, Arizona, USA. *Local Environ.* 1–18. doi:10.1080/13549839.2014.965672
- Brown, H.S., Vergragt, P., Green, K., Berchicci, L., 2003. Learning for Sustainability Transition through Bounded Socio-technical Experiments in Personal Mobility. *Technol. Anal. Strateg. Manag.* 15, 291–315. doi:10.1080/09537320310001601496
- Ceschin, F., 2014. How the Design of Socio-technical Experiments Can Enable Radical Changes for Sustainability. *Int. J. Des.* 8, 1–21.
- Davies, A.R., Doyle, R., 2015. Transforming Household Consumption: From Backcasting to HomeLabs Experiments. *Ann. Assoc. Am. Geogr.* 105, 425–436. doi:10.1080/00045608.2014.1000948
- Davies, A.R., Doyle, R., Pape, J., 2012. Future visioning for sustainable household practices: spaces for sustainability learning? *Area* 44, 54–60. doi:10.1111/j.1475-4762.2011.01054.x
- Evans, J., Karvonen, A., 2014. “Give Me a Laboratory and I Will Lower Your Carbon Footprint!” - Urban Laboratories and the Governance of Low-Carbon Futures. *Int. J. Urban Reg. Res.* 38, 413–430. doi:10.1111/1468-2427.12077
- Farrelly, M., Brown, R., 2011. Rethinking urban water management: Experimentation as a way forward? *Glob. Environ. Chang.* 21, 721–732. doi:10.1016/j.gloenvcha.2011.01.007
- Forrest, N., Wiek, A., 2015. Success factors and strategies for sustainability transitions of small-scale communities – Evidence from a cross-case analysis. *Environ. Innov. Soc. Transitions* 1–19. doi:10.1016/j.eist.2015.05.005
- Hart, D.D., Bell, K.P., Lindenfeld, L.A., Jain, S., Johnson, T.R., Ranco, D., McGill, B., 2015. Strengthening the role of universities in addressing sustainability challenges: the Mitchell Center for Sustainability Solutions as an institutional experiment. *Ecol. Soc.* 20, art4. doi:10.5751/ES-07283-200204
- Laakso, S., Lettenmeier, M., 2015. Household-level transition methodology towards sustainable material footprints. *J. Clean. Prod.* 1–8. doi:10.1016/j.jclepro.2015.03.009
- Lang, D.J., Wiek, A., 2013. The role of universities in fostering urban and regional sustainability, in: Mieg, H.A., Töpfer, K. (Eds.), *Institutional and Social Innovation for Sustainable Development*. Routledge, London, pp. 394–411.
- Liedtke, C., Baedeker, C., Hasselkuß, M., Rohn, H., Grinewitschus, V., 2015. User-integrated innovation in Sustainable LivingLabs: an experimental infrastructure for researching and developing sustainable product service systems. *J. Clean. Prod.* 97, 106–116. doi:10.1016/j.jclepro.2014.04.070

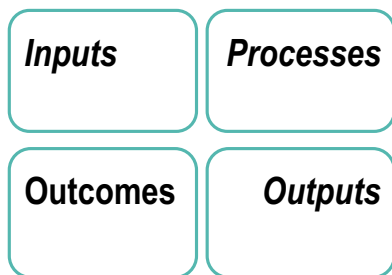
- Loorbach, D., Rotmans, J., 2010. The practice of transition management: Examples and lessons from four distinct cases. *Futures* 42, 237–246. doi:10.1016/j.futures.2009.11.009
- McCormick, K., Kiss, B., 2015. Learning through renovations for urban sustainability: the case of the Malmö Innovation Platform. *Curr. Opin. Environ. Sustain.* 16, 44–50. doi:10.1016/j.cosust.2015.06.011
- Nevens, F., Frantzeskaki, N., Gorissen, L., Loorbach, D., 2013. Urban Transition Labs: co-creating transformative action for sustainable cities. *J. Clean. Prod.* 50, 111–122. doi:10.1016/j.jclepro.2012.12.001
- Quist, J., Thissen, W., Vergragt, P.J., 2011. The impact and spin-off of participatory backcasting: From vision to niche. *Technol. Forecast. Soc. Change* 78, 883–897. doi:10.1016/j.techfore.2011.01.011
- Robinson, J., Berkhout, T., Cayuela, A., Campbell, A., 2013. Next Generation Sustainability at The University of British Columbia: The University as Societal Test-Bed for Sustainability, in: Koenig, A. (Ed.), *Regenerative Sustainable Development of Universities and Cities: The Role of Living Laboratories*. Edward Elgar, Cheltenham.
- Rojas, A., Richer, L., Wagner, J., 2007. University of British Columbia Food System Project: Towards Sustainable and Secure Campus Food Systems. *Ecohealth* 4, 86–94. doi:10.1007/s10393-006-0081-1
- Ryan, C., 2013. Eco-Acupuncture: designing and facilitating pathways for urban transformation, for a resilient low-carbon future. *J. Clean. Prod.* 50, 189–199. doi:10.1016/j.jclepro.2012.11.029
- Schreuer, A., Ornetzeder, M., Rohracher, H., 2010. Negotiating the local embedding of socio-technical experiments: a case study in fuel cell technology. *Technol. Anal. Strateg. Manag.* 22, 729–743. doi:10.1080/09537325.2010.496286
- Seyfang, G., Longhurst, N., 2015. What influences the diffusion of grassroots innovations for sustainability? Investigating community currency niches. *Technol. Anal. Strateg. Manag.* 7325, 1–23. doi:10.1080/09537325.2015.1063603
- Smith, A., Hargreaves, T., Hielscher, S., Martiskainen, M., Seyfang, G., 2015. Making the most of community energies: Three perspectives on grassroots innovation. *Environ. Plan. A*. doi:10.1177/0308518X15597908
- Vergragt, P.J., Brown, H.S., 2012. The challenge of energy retrofitting the residential housing stock: grassroots innovations and socio-technical system change in Worcester, MA. *Technol. Anal. Strateg. Manag.* 24, 407–420. doi:10.1080/09537325.2012.663964
- Voytenko, Y., McCormick, K., Evans, J., Schliwa, G., 2015. Urban living labs for sustainability and low carbon cities in Europe: towards a research agenda. *J. Clean. Prod.* 1–10. doi:10.1016/j.jclepro.2015.08.053
- Westley, F., Antadze, N., Riddell, D.J., Robinson, K., Geobey, S., 2014. Five Configurations for Scaling Up Social Innovation: Case Examples of Nonprofit Organizations From Canada. *J. Appl. Behav. Sci.* 50, 234–260. doi:10.1177/0021886314532945
- Wiek, A., Kay, B., 2015. Learning while transforming: solution-oriented learning for urban sustainability in Phoenix, Arizona. *Curr. Opin. Environ. Sustain.* 16, 29–36. doi:10.1016/j.cosust.2015.07.001
- Wiek, A., Kay, B., Forrest, N., 2015. Worth the Trouble?! An Evaluative Scheme for Urban Sustainability Transition Labs (USTL) and an Application to the USTL in Phoenix, Arizona, in: Frantzeskaki, N., Coenen, L., Castan Broto, V., Loorbach, D. (Eds.), *Urban Sustainability Transitions*. Routledge.

- Wiek, A., Ness, B., Schweizer-Ries, P., Brand, F.S., Farioli, F., 2012. From complex systems analysis to transformational change: a comparative appraisal of sustainability science projects. *Sustain. Sci.* 7, 5–24. doi:10.1007/s11625-011-0148-y
- Wittmayer, J.M., Schöpke, N., van Steenbergen, F., Omann, I., 2014. Making sense of sustainability transitions locally: how action research contributes to addressing societal challenges. *Crit. Policy Stud.* 8, 465–485. doi:10.1080/19460171.2014.957336

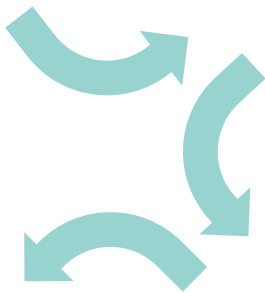
ACCEPTED MANUSCRIPT



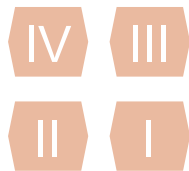
Legend



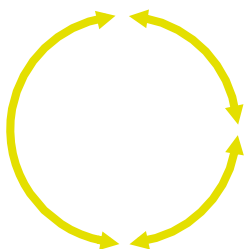
The four dimension of the basic logic model of evaluation adopted for sustainability transition experiments. Depict as parallel and interdependent, evaluation requires iterations among the four dimensions.



Following a sequential order from inputs to outcomes, the arrows display the ideal-typical sequence of activities for designing sustainability transition experiments.



The numbering indicates the sequence of applying the evaluative dimensions following the evaluation rationale



The yellow arrows indicate the interconnectedness of the steps and potential iterations in an experiment.

Highlights:

- A tentative scheme is presented to evaluate sustainability transition experiments
- The scheme is applicable to different types of sustainability transition experiments
- The scheme comprehensively captures the outcomes, inputs and mediating attributes
- It is ready to be applied including guidance for specifying it to particular cases
- It supports experiments in becoming more effective and efficient via reflection and learning