ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Validity and validation in archetype analysis: practical assessment framework and guidelines

ENVIRONMENTAL RESEARCH
LETTERS

LETTER

Validity and validation in archetype analysis: practical assessment framework and guidelines

Luigi Piemontese[1,2], Regina Neudert[3], Christoph Oberlack[4,5], Simona Pedde[6], Matteo Roggero[7], Ana Buchadas[8,9], Dominic A Martin[10], Richard Orozco[11], Kara Pellowe[1], Alcade C Segnon[12,13], Lucía Zarbá[14] and Diana Sietz[15]

[1] Stockholm Resilience Centre, Stockholm University, Kräfttriket 2B, 106 91 Stockholm, Sweden
[2] Department of Agriculture, Food, Environment and Forestry (DAGRI), University of Florence, Florence, Italy
[3] Chair of Landscape Economics and Institute of Botany, Greifswald University, Soldmannstrasse 15, 17487 Greifswald, Germany
[4] Centre for Development and Environment (CDE), University of Bern, Bern, Switzerland
[5] Institute of Geography, University of Bern, Bern, Switzerland
[6] Soil Geography and Landscape Group, Wageningen University and Research, Wageningen, The Netherlands
[7] Resource Economics Group, Humboldt Universität zu Berlin, Berlin, Germany
[8] Geography Department, Humboldt-University Berlin, Unter den Linden 6, 10099 Berlin, Germany
[9] Integrative Research Institute on Transformations of Human-Environment Systems (IRI THESys), Unter den Linden 6, 10099 Berlin, Germany
[10] Wyss Academy for Nature, University of Bern, Bern, Switzerland
[11] Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB), Max-Eyth-Alle 100, 14469 Potsdam, Germany
[12] Faculty of Agronomic Sciences, University of Abomey-Calavi, 01 BP 526 Cotonou, Benin
[13] CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), BP 320 Bamako, Mali
[14] Instituto de Ecología Regional, UNT-CONICET, Tucumán, Argentina
[15] Thünen Institute of Biodiversity, Federal Research Institute for Rural Areas, Forestry and Fisheries, Braunschweig, Germany

E-mail: luigi.piemontese@unifi.it

OPEN ACCESS

RECEIVED
1 December 2021

REVISED
25 January 2022

ACCEPTED FOR PUBLICATION
26 January 2022

PUBLISHED
8 February 2022

Keywords: archetype analysis, validation, validity, sustainability science, social-ecological systems

Supplementary material for this article is available online

Abstract

Archetype analysis is a promising approach in sustainability science to identify patterns and explain mechanisms shaping the sustainability of social-ecological systems. Although considerable efforts have been devoted to developing quality standards and methodological advances for archetype analysis, archetype validation remains a major challenge. Drawing on the insights from two international workshops on archetype analysis and on broader literature on validity, we propose a framework that identifies and describes six dimensions of validity: conceptual; construct; internal; external; empirical; and application validity. We first discuss the six dimensions in relation to different methodological approaches and purposes of archetype analysis. We then present an operational use of the framework for researchers to assess the validity of archetype analysis and to support sound archetype identification and policy-relevant applications. Finally, we apply our assessment to 18 published archetype analyses, which we use to describe the challenges and insights in validating the different dimensions and suggest ways to holistically improve the validity of identified archetypes. With this, we contribute to more rigorous archetype analyses, helping to develop the potential of the approach for guiding sustainability solutions.

1. Introduction

From early mentions in a wide range of disciplines, including philosophy, psychology and arts (Eisenack *et al* 2021), the concept of archetype has evolved into a novel approach used in sustainability science to generate knowledge across cases and places (Oberlack *et al* 2019, Eisenack *et al* 2021). Building on previous assessment frameworks providing comprehensive information to support policy making in the context of social-ecological systems (Ash *et al* 2010, Fallon *et al* 2021), archetypes are defined as recurrent patterns in variables and processes that determine the sustainability of social-ecological systems, identified by a diverse portfolio of methods (Sietz *et al* 2019). They can provide guidance on how

to address sustainability problems in complex socio-ecological systems. Archetype analysis aims at striking a balance between two antithetic needs in sustainability sciences: to achieve well-contextualised solutions and to learn across (different) cases (Poteete *et al* 2010, Beach and Pedersen 2016). To that end, archetype analysis provides guidance on how to navigate the middle ground between idiosyncratic approaches, where every case stands on its own, and nomothetic approaches, where universal solutions rely on oversimplified diagnostics (Ragin 2000, Basurto and Ostrom 2009, Oberlack *et al* 2019). Given these advantages, archetype analysis has also been suggested as an approach for future research on biodiversity and ecosystem services in the recent IPBES report, especially for scenario development (IPBES 2016).

Due to the complexity of socio-ecological systems (Folke *et al* 2005, Pahl-Wostl 2007) and the interdisciplinary nature of sustainability sciences (Karlqvist 1999, Lam *et al* 2014), with its diverging and often conflicting scientific traditions (Nagatsu *et al* 2020, Boda 2021), no silver bullet exists that balances the benefits of situated knowledge with those of strong generalisation. For that reason, some researchers have gone to great lengths to test the reliability of their findings about archetypes, devising creative solutions that test the consistency of archetypes identified against real world observations (Sietz *et al* 2012, Vidal Merino *et al* 2019). Yet, validation in archetype analysis has often been conducted in rather ad-hoc ways, making it an analytical frontier for advancing the meaningfulness of archetype analysis. More meaningful archetypes, in turn, will make for better integration in decision-making on sustainable development (Sietz *et al* 2019).

The aim of the present paper is to advance archetype analysis by proposing an overarching approach to validation in archetype analysis, focusing on design, analysis and application. Relying on the outcomes of two expert workshops in 2019 and 2021 (Václavík *et al* 2019, Piemontese *et al* 2021a), we review different types of validity relevant for archetype analysis and integrate them into an overarching framework for archetype validation with six dimensions of validity for the corresponding archetype analysis steps. Subsequently, we apply the framework to 18 published archetype analyses, identify the strengths and weaknesses with respect to the six validity dimensions, and provide guidance for increasing validity.

2. The six dimensions of validity in archetype analysis

The validation of research results is a common challenge in science. Confounding factors, spurious relations, measurement errors all pose potential threats to research inference, calling for ways of ensuring the validity of research results. Adding to that, sustainability science is characterized by a problem-driven nature (Kates *et al* 2001, Clark and Dickson 2003, Pauliuk 2020) as well as by pervasive uncertainty (Messerli *et al* 2019, Clark and Harley 2020). This forces researchers to make compromises between tractability and data availability considerations on one hand, and what is meaningful and actionable for research stakeholders on the other hand. Further challenges are specific to archetype analysis, particularly the challenge of exploring the middle ground between idiosyncratic and nomothetic approaches (Oberlack *et al* 2019). This pushes searches to engage with causality rather than simply detecting correlations (Mahoney 2008, Rohlfing and Zuber 2021); doing so, however, inevitably forces researchers to engage with mixed methods, raising questions of internal, epistemological consistency for a coherent research design (Morse *et al* 2006, Seawright 2016, Gibson 2017, Sietz *et al* 2019).

With these challenges in mind, archetype validation featured prominently on the agenda of the 2019 and 2021 workshops on archetype analysis (Václavík *et al* 2019, Piemontese *et al* 2021a). The workshops and follow-up discussions identified six dimensions which characterize the concept of validity in archetype analysis and correspond to six archetype analysis steps, i.e. actual stages of the scientific process involving considerations of validation. The validity dimensions described in detail below are: conceptual; construct; internal; external; empirical; and application validity (figure 1). Among these, conceptual validity corresponds to the step of problem framing, when research questions and framing are formulated. Construct validity relates to the step of attribute selection, which operationalizes the framing into quantifiable attributes. These two analysis steps constitute the foundation for designing the archetype analysis. Following the analysis phase includes the methods of analysis, which uses the attributes to reveal the archetypes, and the generalizability of outcomes. These two steps correspond to internal and external validity, respectively. The last phase concerns the application of the archetypes identified, which embraces the empirical relevance of the archetypes in the context of the issues relevant to society and considered in the framing steps (corresponding to empirical and application validity). The validity dimensions and steps may overlap and interact with each other as explained in later sections, given the circular nature of archetype analysis.

2.1. Conceptual validity

Conceptual validity refers to the salience of the research framing, including the formulation of the problem and research questions from the perspective of scientists as well as partners from policy and

practice (when relevant). This is a key component of archetype validity since research problems and questions affect all other dimensions of validity at later stages of the knowledge production cycle (Locke 2012).

From a sustainability science point of view, archetype analysis needs to be rooted in research that is relevant to sustainability and guided by appropriate theoretical framing to produce scientifically-sound knowledge (Magliocca *et al* 2018). At the same time, since sustainability science is a problem-driven research field, research questions and framing should build on problems relevant to society, and may also involve non-academic stakeholders in the problem formulation. Conceptual validity in archetype analysis is based on the integration of both scientifically-sound research framing and problem-driven research questions.

2.2. Construct validity

Construct validity refers to the overall quality, fit, and representativeness of the variables chosen to transform the concepts studied into the attributes used to build the archetypes (Yin 2009, Gerring 2016). Attributes can be characteristics, variables, qualities, factors, or other properties chosen at an intermediate level of abstraction to achieve a balance between case-based validity and generalisation. As pointed out by Eisenack *et al* (2019), achieving a selection of attributes at an appropriate level of abstraction is a major challenge, making it difficult to provide strict selection procedures.

To ensure a coherent and plausible selection, attributes need to be well justified and properly linked with the conceptual framing and generalizability objectives. In the case of deductive research approaches, where researchers hypothesise based on existing theories or frameworks, the attributes should be properly related to the theoretical and conceptual framing (Maxwell 2010). For inductive research, where researchers look for patterns in the data to develop a theory, the attributes should be empirically grounded with the problem addressed (Blackstone 2012). For example, building the research hypothesis and selecting the attributes may be done through participatory approaches in collaboration with stakeholders (Yin 2009). Other examples of scientifically-sound methods for attribute selection are meta-analysis of literature or multi-stakeholder approaches (Oberlack *et al* 2016).

2.3. Internal validity

Internal validity in archetype analysis refers to the appropriateness of the applied method (or combination of methods) for analysing the attributes within the research framing. This is particularly important in archetype analysis, as it can rely on diverse methodological tools from quantitative, qualitative, and mixed methods approaches (Sietz *et al* 2019). The use of appropriate methods for the goals of the research—methods which represent the state of the art in validation of inferences and are backed by detailed documentation—improves the trustworthiness and replicability of the research approach and results.

For straightforward, single-method analyses, this dimension of validity can be addressed by using conventional method-specific validation procedures (Salciccioli *et al* 2016). For example, in quantitative archetype analysis methods, such as cluster analysis, there are standard methods to ensure robust cluster results (e.g. based on a consistency test, see Kok *et al* 2016, Sietz *et al* 2017, Sietz *et al* 2019, Segnon *et al* 2021), and to measure the closeness and representativeness of the clusters (Rocha *et al* 2020). In qualitative archetype analysis methods, quality standards are ensured, for example, via triangulation of information and sources, codebook construction, intercoder reliability checks and assessment of rival explanations (Yin 2009). However, in interdisciplinary or transdisciplinary studies, mixed-method approaches are commonly used to capture and analyse different types of qualitative and quantitative data and provide complementary information to gain insights. Although there are limited specific guidelines to assess internal validity in mixed-method analysis, these typically focus on an integrative framework where methods have a singular internal validation process, as well as the validation of meta-inferences generated through the combination of methods (Tashakkori and Teddlie 2008, Ihantola and Kihn 2011, Neudert *et al* 2019). Finally, to strengthen internal validity, method-specific techniques for uncertainty analysis should be used to define limitations and increase the reliability and usability of results in those cases where it is possible to do so (Bleijenbergh *et al* 2011).

2.4. External validity

External validity pertains to the generalizability of the claims emerging from archetype analysis, or the confidence in extending the results outside of the study sample (Yin 2009, Maxwell 2010, Magliocca *et al* 2018). This is a key challenge in archetype analysis, since an important purpose is to find an intermediate level of abstraction able to provide context-specific generalisations across or even beyond the studied cases. This is particularly important for archetype analyses that aim for the transferability of solutions across contexts. Here, it is important to ensure that a study clearly states the domain of generalizability, both spatially and temporally (Eisenack *et al* 2019). However, defining clear boundaries of generalizability is a challenging and debated task (Magliocca *et al* 2018).

To assess external validity, we suggest that two aspects must be present: (a) the generalizability boundaries need to be stated and discussed in the

paper and guide the study design; and (b) at least one method needs to be used to validate whether the claims hold beyond the studied cases, if applicable, and fall within the same archetypes.

2.5. Empirical validity

Inspired by validation approaches used in modelling (Bossel 1994), archetypes prove to be empirically valid if they correspond to reported sustainability outcomes (e.g. food or livelihood security), and if the causal mechanisms leading to these outcomes are consistent and plausible (Sietz *et al* 2012, 2017, 2019, Vidal Merino *et al* 2019, Sterzel *et al* 2020, Segnon *et al* 2021). Empirical validity is important because the term 'archetypes', rather than 'ideal types', follows the detailed configurations of factors embodied in archetypes, as opposed to the extreme, hypothetical constructs embodied in ideal types (Oberlack *et al* 2019). In both quantitative and qualitative approaches, a real-world check of the archetypes identified can be achieved through stakeholder-based assessments and workshops.

2.6. Application validity

Application validity refers to the usefulness of results for application by final knowledge users (Bossel 1994), for example non-governmental organisations or government officials. Application validity is a key dimension in sustainability research with the aim of fostering real-world actions that increase sustainability (Sietz *et al* 2019), for example, the scaling up of sustainable solutions (Coe *et al* 2014). Applications also include investigations of the relevance of archetypal patterns for improved understanding of sustainability outcomes, e.g. distribution of armed conflicts, complementing traditional statistical methods used to explore the occurrence or lack of armed conflicts (Sterzel *et al* 2014).

Scaling of sustainability solutions is a prime research frontier in archetype analysis to systematically support the transfer of knowledge and insights from one place to another (Sietz *et al* 2019, Eisenack *et al* 2021). The scalability of sustainability strategies rests on the assumption that similar social-ecological conditions translate into comparable interventions (Sietz *et al* 2019, Piemontese *et al* 2020). Archetypes prove to have application validity if the transferability of strategies to foster sustainable land use and other studied phenomena can be shown within a given archetypal pattern and within the stated domain of validity (Václavík *et al* 2016). This can increase the acceptance of findings derived from archetypes (e.g. potential strategies and solutions) by final users who may use these findings to address real-world problems, guide transformations, or further scientific knowledge and innovation (Bleijenbergh *et al* 2011).

3. From concept to application: an operational framework for validity assessment in archetype analysis

The six dimensions of validity are milestones of a highly-interconnected circular process, which are connected to corresponding analytical steps in archetype analysis, designed following the criteria presented by Eisenack *et al* (2019) (figure 1). As in other circular frameworks within sustainability science (Holling and Gunderson 2002, Pahl-Wostl 2009, Ostrom 2011), the circularity implies that there is neither fixed hierarchy nor a predefined 'first step' among the six dimensions. For example, addressing conceptual validity dimension in the problem framing is not always the first step of the validation process. The problem framing and research questions of the analysis may strongly depend on the needs and purposes of stakeholders (e.g. policy makers, local communities, and practitioners), which depend on the impact that the analysis is intended to produce. In this case, addressing application validity and not conceptual validity may be the first step of the validation process. Each step in this process is broad, in the sense that the analysis is interdisciplinary and often involves iterations between scientists and stakeholders.

Each analytical step can strengthen or weaken the significance of the other steps, hence the relevance of the validity dimensions. The relative significance can generate both trade-offs and synergies among dimensions and depends on the purpose (real-world application or theoretical contribution), approach (inductive or deductive), and the methodological lens (qualitative, quantitative, or mixed-methods; Davies and Elder 2005, Koro-Ljungberg 2008). For example, a study designed for real-world application (thus identifying practical knowledge needs first) and consisting of a well-planned participatory approach may increase impact by targeting and framing problems as perceived by stakeholders, thus strengthening the application validity dimension. A stronger application validity demands a clear problem framing, which strengthens the conceptual validity dimension in a way that is immediately understandable and useful for users. This could be the case also for studies where specific methods, either qualitative, quantitative, or mixed, are well-known and suitable for case studies with similar characteristics. In such cases, stronger links between attribute selection and analysis can facilitate the construct and internal validity dimensions, strengthening impact and therefore application validity, particularly when stakeholders are involved. In clustering, for example, establishing internal validity includes the selection of a suitable number of clusters to optimise variance measures or reproducibility. Stakeholder

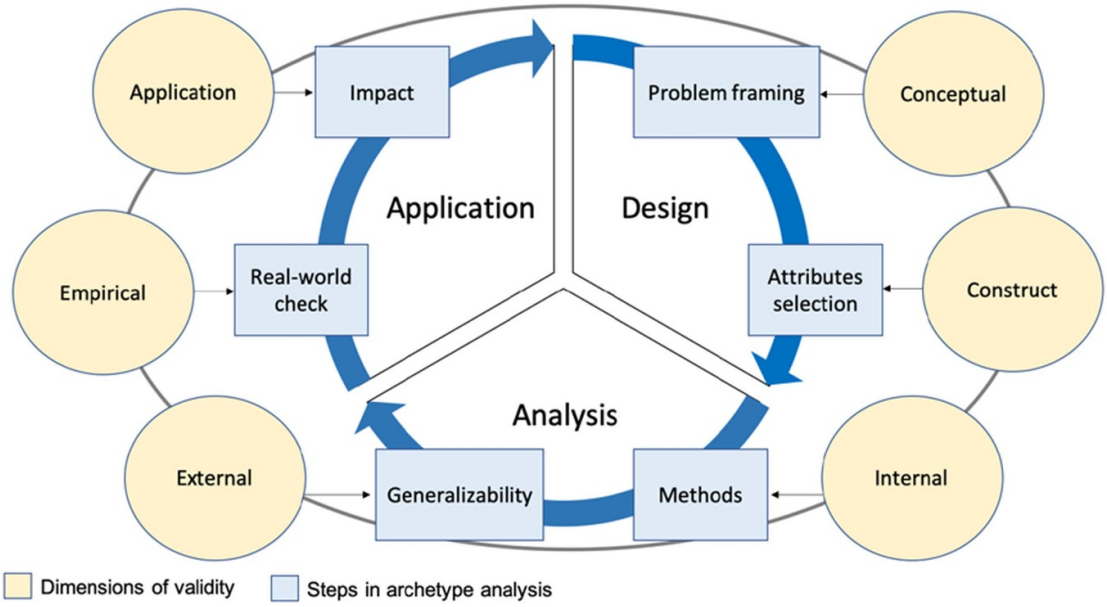

Figure 1. The six dimensions of validity within the archetype analysis process, which we derived from the expert workshops and follow-up discussions.

involvement in interpretation can support decisions on appropriate numbers of clusters, strengthening the link between internal and application validity.

At this point, it is important to keep in mind that different steps may have trade-offs. For instance, strengthening internal validity through greater model specification and tighter scope conditions will unavoidably weaken the external validity of an analysis (i.e. the generalizability beyond the case studies; Jimenez-Buedo and Miller 2010, Druckman *et al* 2011).

We applied and tested the six validity dimensions on a selection of published archetype analyses across different sustainability research fields. In its application, the six dimensions provide an operational framework to guide the validity assessment in archetype analyses. The framework consists of the definitions of validity standards for the six dimensions, with evaluation criteria ranging from weak ('W'), medium ('M'), to strong ('S'). We draw the evaluation criteria of validity dimensions (table 1) based on the definitions of the six dimensions (see section 2), cross-cutting characteristics of archetype analysis and interdisciplinary requirements of sustainability science. For example, conceptual validity in archetype analysis is based on the level of fit and integration between the two salient conceptual components of (a) scientifically-sound research framing and (b) problem-driven research questions. Thus, we define conceptual validity as weak if both components are absent or poorly addressed, medium if one of the two components is absent or poorly addressed, and strong if both components are properly addressed.

4. Best-practice guideline to improve validity in archetype analysis

We applied the validity assessment to 18 published archetype analyses (supplementary material, section S2 available online at stacks.iop.org/ERL/17/025010/mmedia) to demonstrate the usefulness of the framework and provide overarching guidelines and lessons learned to support a wider and informed use of the framework within the archetype analysis research community. Each assessment was performed by one co-author and reviewed by all the co-authors of this paper to minimise biases. Nonetheless, the exemplary assessments are not meant to be completely objective, but rather to show the applicability of the framework and help draw generalised guidelines and insights on its usability. In the following sections we refer to these assessments using a code made of a letter and the name of the first author.

The overall results of the assessments indicate that almost all studies encounter the trade-offs described in section 3; none of the selected archetype analyses scored 'Strong' on all validity dimensions (figure 2). Overall, external and application validity are the weakest dimensions. One reason for this finding may be that deliberate planning of validation is not yet a common good practice in archetype analysis, and many publications do not report on their practice of validation. This points to the need for a comprehensive, multi-dimensional validation plan, designed before the start of the analysis, and for its reporting in publications. The advantage of implementing a validation plan for single studies is that it clarifies, step by step, which phase might be weaker, thus offering an opportunity to increase validity for all

Table 1. Definition of evaluation criteria for validity assessment in archetype analysis.

Validity dimension	Evaluation criteria		
	Weak	Medium	Strong
Conceptual	Both of the two salient conceptual components (scientifically-sound research framing and problem-driven research questions) are absent or poorly addressed.	Either of the two salient conceptual components (scientifically-sound research framing and problem-driven research questions) is absent or poorly addressed.	Both of the two salient conceptual components (scientifically-sound research framing and problem-driven research questions) are properly addressed.
Construct	No clear connection between the selected attributes and the conceptual framing of the study (e.g. no clear theoretical nor empirical justification of attribute selection).	Attributes and the conceptual framing are connected, but attributes are selected without rigorous methods (e.g. meta-analysis).	Clear connection between the attributes and the conceptual framing and attributes are selected using rigorous methods.
Internal	The method of analysis is inappropriate for the objectives and data used, or the method is not state-of-the-art, or it is unclear how results emerged from primary data.	Methods are appropriate and state-of-the-art, but there is inadequate information on data sources and/or analysis. Method-specific validity checks are not applied or are insufficiently applied.	State-of-the art methods are used, data sources and analysis procedures are sufficiently described to ensure replicability; method-specific sensitivity or reliability measures are used; validity of research results is critically discussed.
External	No consideration of generalizability boundaries and no methods are used to cross-check for external validity.	Generalizability boundaries are considered, but there are no methods used to cross-check for external validity.	Generalizability boundaries are considered and at least one method is used to cross-check for external validity.
Empirical	No clear correspondence between archetypes, reported outcome and causal mechanisms.	Archetypes correspond to some extent to reported outcomes and are partially consistent with causal mechanisms leading to these outcomes.	Archetypes clearly correspond to reported outcomes and are consistent with causal mechanisms leading to these outcomes.
Application	The archetype application, e.g. transferability of sustainability solutions within ATs, is not indicated or cannot be proven.	The archetype application, e.g. transferability of sustainability solutions within ATs, is partially indicated, based on independent data, case studies and/or stakeholders' perceptions.	The archetype application, e.g. transferability of sustainability solutions within ATs, is explicitly indicated, based on independent data, case studies and/or stakeholders' perceptions.

dimensions. Using these insights, we developed best-practice guidelines to improve validation in archetype analysis. These guidelines include considerations for developing a validation plan, as well as strengthening the links and reducing trade-offs between dimensions.

4.1. Developing a validation plan

Validation needs to be planned as part of the research design before conducting the actual archetype analysis. The validation plan should outline the desired level of validation with regard to all six dimensions, it identifies in which phase of archetype analysis validation is needed, for which dimensions, and present methods for testing and increasing validity. The validation plan should distinguish between independent and interdependent dimensions. When dimensions need to be validated independently, it could help to use complementary approaches. For example, in quantitative archetype analysis, a viable approach to increase external, empirical, and application validity is to define an independent set of validation data to compare and cross-check the archetypes identified with other sources (e.g. A-Sietz, B-Vidal-Merino, C-Piemontese and M-Segnon). When dimensions are interdependent, the addition of complementary approaches could include details on existing data and/or plan the collection of additional data for validation. For example, the inclusion of stakeholder interviews and workshops can support validation of several dimensions in qualitative archetype analysis (G-Karrash, S-Neudert) and scenario archetypes (E-Pedde, P-Pinnegar). The value of synergies deriving from interdependent dimensions is discussed more in detail in section 4.2.

To design a comprehensive validation plan, we suggest conducting archetype analysis and validation in iterative steps, with sufficient resources and time scheduled to enable several iterations. Decisions in the iterations include: (a) adjusting or adding hypotheses, underlying processes, or data sets used in the archetype analysis to obtain a closer agreement with observed system behaviour; (b) revising archetype results based on insights obtained from validation;

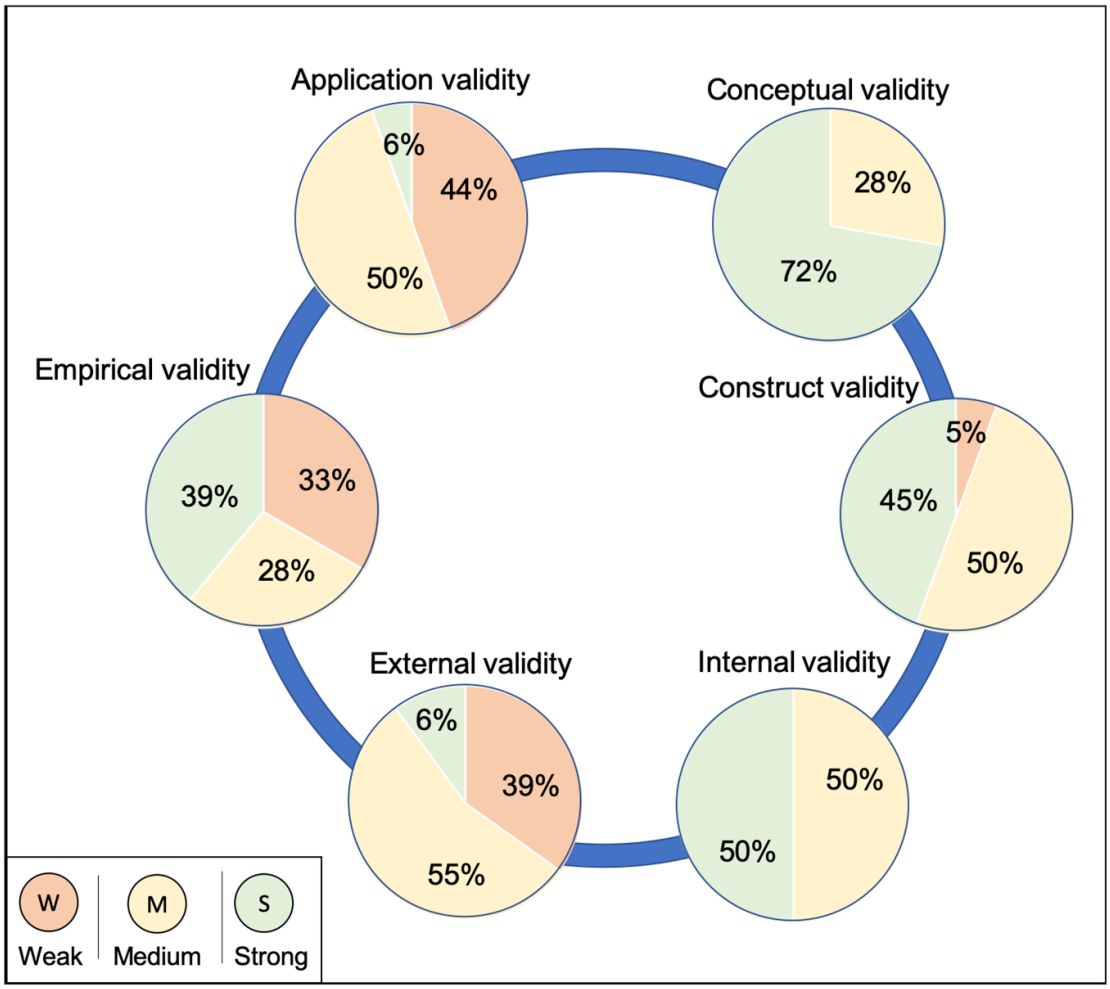

Figure 2. Overall score on the validity of the six dimensions from our assessments of 18 published archetype analysis. See section S2 of supplementary material for an overview of all assessments.

and (c) confirming archetype credibility with regard to each validity dimension.

4.2. Validation requirements

The first step in designing a validation procedure is defining and weighing the degree to which and how archetypes are to be validated. Desired levels of validation precision include the definition of functional and spatial resolution, user perceptions, and temporal dynamics that the identified archetypes need to reflect. This means that levels of allowable or unavoidable imprecision must be specified. For example in conceptual validity, there appears to be a trade-off between case-specific research questions, which can be validated and refined with stakeholders, and broader-scale, scientifically-driven research questions of broad relevance to the scientific community (e.g. the case-specific framing in S-Neudert vs the broad global scale in T-Oberlack). The problem framing (conceptual validity) in global studies can be informed by literature and/or theories, review of case studies, and stakeholder participation. Iteratively, conceptual validity influences application validity in that a research problem framed with stakeholders enhances the relevance of the research results for stakeholders (e.g. F-Thorn). Transferability (i.e. application validity), typically a key motivation of archetype analysis, depends on the extent to which scientists and practitioners from multiple contexts perceive problem settings to be similar (e.g. see empirical validity in I-Levers).

If well-documented local case studies are used to validate broad spatially-explicit archetypes, the question arises as to whether the small local extent and the high functional resolution of case studies represents the large extent and functional aggregation of broader archetypes. Ideally, there would be a sufficiently-large number of independent case studies available for validation that capture the variations within an archetype. However, these are often not available. Alternatively, available case studies can be used to illustrate important features, processes, and sustainability outcomes (e.g. A-Sietz, C-Piemontese). Moreover, follow-up case studies are of critical importance to revise or validate knowledge about archetypes from previous studies.

Moreover, the level of consistency with independently-reported causal mechanisms and

sustainability outcomes (i.e. empirical validity) can be determined by stakeholder evaluation (e.g. I-Levers), inter-coder reliability in model-centred meta-analyses of case studies (Oberlack *et al* 2016) and statistical approaches (Sietz *et al* 2012). Working at a local level is advantageous for quantitative validation (Sietz *et al* 2012) given that limitations in regional or global observational data (e.g. spatial resolution) often constrain such validation on larger scales (e.g. C-Piemontese). Stakeholder evaluation can improve qualitative consistency, e.g. by merging cluster pairs which are qualitatively similar (e.g. I-Levers).

External, empirical, and application validity often require additional research steps, e.g. verifying identified archetypes with additional data or approaches (e.g. external and empirical validity) or consulting stakeholders with research results (i.e. empirical and application validity). If these steps cannot be carried out or reported in an archetype analysis, the limitations and the required additional steps should be mentioned (e.g. empirical and application validity in T-Oberlack).

The validity of archetypes can be improved by investing more time in testing validity and adjusting the analysis. Yet at some point, a further investment results only in a very small improvement in validity (i.e. saturation effect). Therefore, it is important to determine as part of the decision procedure whether the precision level reached is sufficient for the objective of the study.

4.3. Directions for future validation

4.3.1. Improving weaker validity dimensions

Application validity is one of the weakest dimensions in the present assessment, with 44% of the studies scoring weak and only 6% scoring strong (figure 2). One likely reason is that assessing the transferability of knowledge across contexts is very challenging within the scope of a single study (Vaclavik *et al* 2016). Future research should therefore increasingly test whether sustainability solutions can be transferred across cases or regions that share an archetype, i.e. similarities in specific attributes and/or causal effects (Adler *et al* 2018, Eisenack *et al* 2019, Sietz *et al* 2019). To achieve this, researchers would need to test, challenge, or refine earlier identified archetypes, rather than identifying new ones. Such a programmatic approach in archetype analysis will build more cumulative knowledge about archetypes of sustainability over time.

Weak application validity can also arise if the research problem and questions are framed from a purely disciplinary or academic perspective (Beran *et al* 2021). Future archetype analyses may improve their application validity by embracing transdisciplinary co-creation of knowledge through participatory research approaches (Wuelser *et al* 2021, Jacobi *et al* 2022), as discussed in section 4.3.2. Finally, future studies aiming for high application validity should define application requirements, including clear definitions of intended users, spatial scale, and the temporal horizon of application, because these are prerequisites for determining the degree to which archetypes accurately represent knowledge needs and real-world situations.

The 2nd weakest dimension is external validity (39% scoring weak and only 6% strong, figure 2). While external validity generally increases with the number and diversity of cases considered, rich context-specific knowledge, the idiosyncrasy of each case and high data requirements of each case often limit the total number of cases, thus limiting external validity of a single study (e.g. J-Banson) (Wuelser *et al* 2021). Therefore, approaches to increase external validity include a careful and unbiased selection of cases from different contexts or the use of different sets or types of data in the delineation and the validation of archetypes respectively.

Another reason for low external validity arises if a study does not specify the conditions under which a knowledge claim is expected to hold true (Magliocca *et al* 2018). Therefore, a requirement for external validity in archetype analyses is to state the limits of generalizability and establish a clear domain of validity. Furthermore, if future research increasingly adopts a programmatic approach to archetypes, by testing, challenging or refining earlier identified archetypes for new contexts (e.g. different countries), external validity of knowledge claims about archetype will increase cumulatively.

The scale of archetypes needs to be considered in validation best practises. A small set of broad spatially-explicit archetypes may have high internal validity at the specific scale at which they are identified (Sietz 2014, Sietz *et al* 2017). Yet, archetypes observed at one scale may manifest differently at other scales and findings can only be generalised when underlying relationships are consistent across scales. Nested archetypes identified at various scales are one way to operationalize a multi-scale approach providing refined opportunities for transferring sustainability solutions from one place to another (A-Sietz).

4.3.2. Encouraging participatory approaches to strengthen validity

Many of the weaknesses that we found in our validity assessments could be improved by involving stakeholders in the research process through participatory methods. For example, involving stakeholders from the problem framing and hypothesis formulation phase can increase not only the conceptual validity, but can also enhance the application validity and final utility of archetypes (Vidal Merino *et al* 2019). Stakeholder involvement can prove useful also for increasing construct validity, because a participatory selection of attributes could be a strategy to identify the most relevant attribute and operationalize the

conceptual problem (Piemontese *et al* 2021b), especially if used after participatory problem framing. For internal validity, participatory methods are only appropriate for specific, mostly qualitative, research methods. Using participatory methods can also benefit the external validity of archetypes, for example, by guiding and assessing the extent of and relevant conditions for generalizability.

A relevant role for participatory approaches is also in improving empirical and application validity. Insights derived from archetype analysis should be translated and diffused in appropriate communication channels in a way that stakeholders can make use of them (Lang *et al* 2012). Involving final users can increase their acceptance of findings derived from archetypes (e.g. potential strategies and solutions), and facilitate the translation of findings to solve real-world problems or to make scientific progress (Bleijenbergh *et al* 2011, Jacobi *et al* 2022). Finally, the acceptance of archetypes is key to generating actionable knowledge. For that, stakeholders should feel ownership over and commitment to the archetypes identified, and consider as legitimate any insights resulting from archetype analysis. Such results are best rendered through transdisciplinary research efforts (Bleijenbergh *et al* 2011, Lang *et al* 2012).

Nonetheless, using participatory methods is not a panacea to justify choices made in analysis without providing the rationale, motivation, and boundaries for the proper use of participatory methods and careful explanation of how they contribute to archetype validity. We note that participatory methods require sufficient resources, access in social networks and stakeholder interest in research (Tribaldos *et al* 2020). Finally, archetypes are meant to provide transferable and generalizable knowledge claims to support sustainability. Thus, any generalised claims need to carefully account for a diverse and inclusive group of stakeholders so as to reduce biased information and to produce effective and just recommendations.

5. Conclusions

Validating archetypes identified in sustainability science is to date an under-utilised opportunity to demonstrate their credibility and hence usefulness for informing decision making. To advance validation in archetype analysis in a structured way, we introduced six dimensions of validity and illustrated how they may be used. We discussed how to analyse validity and plan for archetype validation considering the interdependence of validity dimensions. We provided recommendations to support a more comprehensive validation process, as well as evaluation criteria to identify weak, medium, and strong levels of validity for the six dimensions.

In contrast to the classical notion of validation as a check of results against an objective reality, we provided a holistic approach that seeks to achieve validity throughout the research process. We recommend designing each stage of the validation process in such a way that synergies between the six validity dimensions are considered, thereby increasing the overall validity and societal relevance of resulting archetypes. Future archetype validation can best be advanced based on (a) systematic planning of validation procedure, (b) better availability of functionally- and spatially-resolved data, and (c) encouragement of multi-stakeholder participation. Our framework provides guidance to those seeking to enhance the validity of archetypes by systematically understanding and using the opportunities associated with particular analytical steps.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

We cordially thank all participants of the international research workshops on 'Archetype analysis in sustainability research' (2019–2021) providing the basis for this paper. We are particularly thankful to Klaus Eisenack, Zuzana Harmackova and Tomas Vaclavik for the useful discussions in the early stage of the work. This paper contributes to the GLP science plan (Global Land Programme, https://glp.earth).

ORCID iDs

Luigi Piemontese https://orcid.org/0000-0002-1600-5450
Ana Buchadas https://orcid.org/0000-0003-4219-108X
Dominic A Martin https://orcid.org/0000-0001-7197-2278
Alcade C Segnon https://orcid.org/0000-0001-9751-120X
Lucía Zarbá https://orcid.org/0000-0002-9118-0566

References

Adler C, Hadorn G H, Breu T, Wiesmann U and Pohl C 2018 Conceptualizing the transfer of knowledge across cases in transdisciplinary research *Sustain. Sci.* **13** 179–90

Ash N *et al* 2010 *Ecosystems and Human Well-Being: A Manual for Assessment Practitioners* (Washington, DC: Island Press)

Basurto X and Ostrom E 2009 Beyond the tragedy of the commons *Econ. delle fonti di energia e dell'ambiente* **52** 35–60

Beach D and Pedersen R B 2016 *Causal Case Study Methods: Foundations and Guidelines for Comparing, Matching, and Tracing* (Ann Arbor, MI: University of Michigan Press)

Beran D *et al* 2021 Rethinking research processes to strengthen co-production in low and middle income countries *BMJ* **372** m4785

Blackstone A 2012 *Principles of Sociological Inquiry—Qualitative and Quantitative Methods* (Saylor Foundation) (available at: https://saylordotorg.github.io/text_principles-of-sociological-inquiry-qualitative-and-quantitative-methods/) Shared under CC-BY-NC-SA 3.0 License (https://creativecommons.org/licenses/by-nc-sa/3.0/) (https://doi.org/10.1177/0003122412451728)

Bleijenbergh I, Korzilius H and Verschuren P 2011 Methodological criteria for the internal validity and utility of practice oriented research *Qual. Quant.* **45** 145–56

Boda C S 2021 Values, science, and competing paradigms in sustainability research: furthering the conversation *Sustain. Sci.* **16** 2157–61

Bossel H 1994 *Modeling and Simulation* (Wiesbaden: Springer)

Clark W C and Dickson N M 2003 Sustainability science: the emerging research program *Proc. Natl Acad. Sci.* **100** 8059–61

Clark W C and Harley A G 2020 Sustainability science: toward a synthesis *Annu. Rev. Environ. Resour.* **45** 331–86

Coe R, Sinclair F and Barrios E 2014 Scaling up agroforestry requires research 'in' rather than 'for' development *Curr. Opin. Environ. Sustain.* **6** 73–77

Davies A and Elder C 2005 Validity and validation in language testing *Handbook of Research in Second Language Teaching and Learning* ed E Hinkel (New York: Routledge) p 20

Druckman J N, Greene D P, Kuklinski J H and Lupia A (eds) 2011 *Cambridge Handbook of Experimental Political Science* (Cambridge: Cambridge University Press)

Eisenack K, Oberlack C and Sietz D 2021 Avenues of archetype analysis: roots, achievements, and next steps in sustainability research *Ecol. Soc.* **26** 31

Eisenack K, Villamayor-Tomas S, Epstein G, Kimmich C, Magliocca N, Manuel-Navarrete D, Oberlack C, Roggero M and Sietz D 2019 Design and quality criteria for archetype analysis *Ecol. Soc.* **24** 6

Fallon A, Lankford B and Weston D 2021 Navigating wicked water governance in the 'solutionscape' of science, policy, practice, and participation *Ecol. Soc.* **26** 37

Folke C, Hahn T, Olsson P and Norberg J 2005 Adaptive governance of socio-ecological systems *Annu. Rev. Environ. Resour.* **30** 441–73

Gerring J 2016 *Case Study Research: Principles and Practices* 2nd edn (Cambridge: Cambridge University Press)

Gibson C B 2017 Elaboration, generalization, triangulation, and interpretation: on enhancing the value of mixed method research *Organ. Res. Methods* **20** 193–223

Holling C S and Gunderson L H 2002 *Panarchy: Understanding Transformations in Human and Natural Systems* (Washington, DC: Island Press)

Ihantola E M and Kihn L A 2011 Threats to validity and reliability in mixed methods accounting research *Qual. Res. Account. Manage.* **8** 39–58

IPBES 2016 Summary for policymakers of the methodological assessment of scenarios and models of biodiversity and ecosystem services by the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services *Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services* ed S Ferrier, K N Ninan, P Leadley, R Alkemade, L A Acosta, H R Akçakaya, L Brotons, W Cheung, V Christensen, K A Harhash, J Kabubo-Mariara, C Lundquist, M Obersteiner, H Pereira, G Peterson, R Pichs-Madruga, N H Ravindranath, C Rondinini and B Wintle (Bonn) p 32

Jacobi J, Llanqe A, Mukhovi S M, Birachi E, von Groote P, Eschen R, Hilber-Schoeb D I, Frossard E and Robledo-Abad C 2022 Transdisciplinary co-creation increases the utilization of knowledge from sustainable development research *Environ. Sci. Policy* **129** 107–15

Jimenez-Buedo M and Miller L M 2010 Why a trade-off? The relationship between the external and internal validity of experiments *THEORIA Int. J. Theory History Found. Sci.* **25** 301–21

Karlqvist A 1999 Going beyond disciplines: the meanings of interdisciplinarity *Policy Sci.* **32** 379–83

Kates R W, Clark W C, Corell R, Hall J M, Jaeger C C, Lowe I and Svedin U 2001 Sustainability science *Science* **292** 641–2

Kok M, Lüdeke M, Lucas P, Sterzel T, Walther C, Janssen P, Sietz D and de Soysa I 2016 A new method for analysing socio-ecological patterns of vulnerability *Reg. Environ. Change* **16** 229–43

Koro-Ljungberg M 2008 Validity and validation in the making in the context of qualitative research *Qual. Health Res.* **18** 983–9

Lam J C K, Walker R M and Hills P 2014 Interdisciplinarity in sustainability studies: a review *Sustain. Dev.* **22** 158–76

Lang D J, Wiek A, Bergmann M, Stauffacher M, Martens P, Moll P, Swilling M and Thomas C J 2012 Transdisciplinary research in sustainability science: practice, principles, and challenges *Sustain. Sci.* **7** 25–43

Locke E A 2012 Construct validity vs concept validity *Hum. Resour. Manage. Rev.* **22** 146–8

Magliocca N R, Ellis E C, Allington G R H, de Bremond A, Dell'Angelo J, Mertz O, Messerli P, Meyfroidt P, Seppelt R and Verburg P H 2018 Closing global knowledge gaps: producing generalized knowledge from case studies of social-ecological systems *Glob. Environ. Change* **50** 1–14

Mahoney J 2008 Toward a unified theory of causality *Comp. Polit. Stud.* **41** 412–36

Maxwell J 2010 Understanding and validity in qualitative research *Harv. Educ. Rev.* **62** 279–301

Messerli P, Kim E M, Lutz W, Moatti J P, Richardson K, Saidam M and Furman E 2019 Expansion of sustainability science needed for the SDGs *Nat. Sustain.* **2** 892–4

Morse J M, Niehaus L, Wolfe R R and Wilkins S 2006 The role of the theoretical drive in maintaining validity in mixed-method research *Qual. Res. Psychol.* **3** 279–91

Nagatsu M, Davis T, DesRoches C T, Koskinen I, MacLeod M, Stojanovic M and Thorén H 2020 Philosophy of science for sustainability science *Sustain. Sci.* **15** 1807–17

Neudert R, Salzer A, Allahverdiyeva N, Etzold J and Beckmann V 2019 Archetypes of common village pasture problems in the South Caucasus: insights from comparative case studies in Georgia and Azerbaijan *Ecol. Soc.* **24** 5

Oberlack C *et al* 2019 Archetype analysis in sustainability research: meanings, motivations, and evidence-based policy making *Ecol. Soc.* **24** 26

Oberlack C, Tejada L, Messerli P, Rist S and Giger M 2016 Sustainable livelihoods in the global land rush? Archetypes of livelihood vulnerability and sustainability potentials *Glob. Environ. Change* **41** 153–71

Ostrom E 2011 Background on the institutional analysis and development framework *Policy Stud. J.* **39** 7–27

Pahl-Wostl C 2007 The implications of complexity for integrated resources management *Environ. Model. Softw.* **22** 561–9

Pahl-Wostl C 2009 A conceptual framework for analysing adaptive capacity and multi-level learning processes in resource governance regimes *Glob. Environ. Change* **19** 354–65

Pauliuk S 2020 Making sustainability science a cumulative effort *Nat. Sustain.* **3** 2–4

Piemontese L, Castelli G, Fetzer I, Barron J, Liniger H, Harari N, Bresci E and Jaramillo F 2020 Estimating the global potential of water harvesting from successful case studies *Glob. Environ. Change* **63** 102121

Piemontese L, Kamugisha R N, Tukahirwa J M B, Tengberg A, Pedde S and Jaramillo F 2021b Barriers to scaling sustainable land and water management in Uganda: a cross-scale archetype approach *Ecol. Soc.* **26** 6

Piemontese L, Pedde S, Oberlack C and Roggero M 2021a *Fourth International Research Workshop 'Archetypes of Sustainable Development', Virtual Workshop Report* (Humboldt-Universität zu Berlin) (available at: www.agrar.hu-berlin.de/de/institut/departments/daoe/ress/forneu/4th_atana_workshop_report.pdf)

Poteete A R, Janssen M A and Ostrom E 2010 *Working Together: Collective Action, the Commons, and Multiple Methods in*

Practice (Princeton, NJ: Princeton University Press)
Ragin C C 2000 *Fuzzy-Set Social Science* (Chicago, IL: University of Chicago Press)
Rocha J, Malmborg K, Gordon L, Brauman K and DeClerck F 2020 Mapping social-ecological systems archetypes *Environ. Res. Lett.* **15** 034017
Rohlfing I and Zuber C I 2021 Check your truth conditions! Clarifying the relationship between theories of causation and social science methods for causal inference *Sociol. Methods Res.* **50** 1623–59
Salciccioli J D, Crutain Y, Komorowski M and Marshall D C 2016 Sensitivity analysis and model validation *Secondary Analysis of Electronic Health Records* ed M I T C Data (Cham: Springer International Publishing) pp 263–71
Seawright J 2016 Better multimethod design: the promise of integrative multimethod research *Secur. Stud.* **25** 42–49
Segnon A C, Totin E, Zougmoré R B, Lokossou J C, Thompson-Hall M, Ofori B O, Achigan-Dako E G and Gordon C 2021 Differential household vulnerability to climatic and non-climatic stressors in semi-arid areas of Mali, West Africa *Clim. Dev.* **13** 697–712
Sietz D 2014 Regionalisation of global insights into dryland vulnerability: better reflecting smallholders' vulnerability in Northeast Brazil *Glob. Environ. Change* **25** 173–85
Sietz D, Frey U, Roggero M, Gong Y, Magliocca N, Tan R, Janssen P and Václavík T 2019 Archetype analysis in sustainability research: methodological portfolio and analytical frontiers *Ecol. Soc.* **24** 34
Sietz D, Mamani Choque S E and Lüdeke M K B 2012 Typical patterns of smallholder vulnerability to weather extremes with regard to food security in the Peruvian Altiplano *Reg. Environ. Change* **12** 489–505
Sietz D, Ordoñez J C, Kok M T J, Janssen P, Hilderink H B M, Tittonell P and van Dijk H 2017 Nested archetypes of vulnerability in African drylands: where lies potential for sustainable agricultural intensification? *Environ. Res. Lett.* **12** 095006
Sterzel T, Lüdeke M K B, Walther C, Kok M T, Sietz D and Lucas P L 2020 Typology of coastal urban vulnerability under rapid urbanization *PLoS One* **15** e0220936
Sterzel T, Lüdeke M, Kok M, Walther C, Sietz D, de Soysa I, Lucas P and Janssen P 2014 Armed conflict distribution in global drylands through the lens of a typology of socio-ecological vulnerability *Reg. Environ. Change* **14** 1419–35
Tashakkori A and Teddlie C 2008 Quality of inferences in mixed methods research: calling for an integrative framework *Advances in Mixed Methods Research* ed M M Bergman (London: SAGE Publications Ltd) pp 101–19
Tribaldos T, Oberlack C and Schneider F 2020 Impact through participatory research approaches: an archetype analysis *Ecol. Soc.* **25** 15
Václavík T, Langerwisch F, Cotter M, Fick J, Häuser I, Hotes S, Kamp J, Settele J, Spangenberg J H and Seppelt R 2016 Investigating potential transferability of place-based research in land system science *Environ. Res. Lett.* **11** 095002
Václavík T, Oberlack C, Eisenack K and Sietz D 2019 *The 3rd Workshop on Archetype Analysis in Sustainability Research* (*30 October–1 November 2019*) (Czech Republic: Palacký University Olomouc) (available at: www.agrar.hu-berlin.de/de/institut/departments/daoe/ress/forneu/3rd_archetype_workshop_report.pdf)
Vidal Merino M, Sietz D, Jost F and Berger U 2019 Archetypes of climate vulnerability: a mixed-method approach applied in the peruvian andes *Clim. Dev.* **11** 418–34
Wuelser G, Adler C, Breu T, Hirsch Hadorn G, Wiesmann U and Pohl C 2021 On which common ground to build? Transferable knowledge across cases in transdisciplinary sustainability research *Sustain. Sci.* **16** 1891–905
Yin R K 2009 *Case Study Research: Design and Methods* 4th edn (Los Angeles, CA: SAGE)