Research Notes

Sampling in design research: Eight key considerations

Philip Cash, DTU Technical University of Denmark, Denmark

Ola Isaksson, Chalmers University of Technology, Gothenburg, Sweden

Anja Maier, University of Strathclyde, UK

Joshua Summers, University of Texas at Dallas, USA

How a research team defines their study sample can be decisive in shaping impact on both practice and theory. However, sampling in design research faces several major challenges, including diverse terminology, limited prior literature, and lack of common framework for discussing sampling decisions. We address these challenges by bringing together guidance from across related research fields as well as cross-referring to examples from published design research. We offer a structured process for sample development and present eight key sampling considerations. The paper contributes to research method selection, development, and use, as well as extending discussions surrounding knowledge construction, standards of reporting, and design research impact.

Keywords: research methods, design science, design research, case studies, experimental design

How a research sample is defined forms a key element in establishing the scope of a study and in shaping its potential impact on both theory and practice. Sampling affects both scientific rigour and, in many cases, the perceived value and practical impact of research (Douglas, Noble, & Newman, 1999; Wacker, 2008); and thus forms a key link between considerations related to knowledge construction, research methods, and research impact (Cash, Daalhuizen, & Hay, 2022). However, these key considerations are often left implicit or under-acknowledged in both setting and reporting the work.

Corresponding author:
Philip Cash
pcas@dtu.dk

Design research (the study of design) brings together a diverse set of philosophies and approaches, including constructivism, participatory research through design, and qualitative and quantitative theory and methodology

www.elsevier.com/locate/destud
0142-694X *Design Studies* **78** (2022) 101077
https://doi.org/10.1016/j.destud.2021.101077

development. While this diversity gives rise to much of the richness within the field it also means that any methodological discussion must be carefully contextualised. Further, overarching methodologies such as the Design Research Methodology (DRM) by Blessing and Chakrabarti (2009) or Design Science (Hevner, 2007), require detailed support for specific aspects of the research process in order to provide a robust body of methodological knowledge. Thus, we focus on qualitative and quantitative studies engaged with the scientific theory development cycle, i.e. contributing to theory building and/or theory testing (Cash, 2020; Wallace, 1971, p. 18). This provides a foundation for interpretation and adaptation of methodological discussions found in a similar context in related fields, as well as implications for other research approaches sharing similar concerns.

The methods relevant to this context range from qualitative interview and case studies (where sampling can relate to, for example, the selection of specific cases, as in the work of Crilly and Moroşanu Firth (2019)), to smaller scale quantitative or mixed method studies (where sampling can relate to the specific definition of study participants, as in the study by Nelius, Doellken, Zimmerer, and Matthiesen (2020)), or to large scale quantitative studies (where sampling can be used to reflect whole segments of a population, as in the work of Graff, Meslec, and Clark (2020)). While each method emphasises different aspects of sampling, there are three key challenges to their adaptation and application in design, which we aim to address in this paper.

First, there is a high degree of terminological diversity in the reporting of studies and across the literature. This can contribute to difficulties in study interpretation as well as a perception of siloing between studies (Le Dain, Blanco, & Summers, 2013), such as in experimental versus real world settings (Ball & Christensen, 2018; Crilly, 2019a). Thus, there is a need to clarify relevant sampling terminology.

Second, there has been little specific discussion of sampling terminology or considerations in the design research literature. Hence, the reporting of sampling decisions is often implicit and can appear to be something of a methodological 'black box'. This hampers efforts to develop methodological rigour, as well as the examination of potential synergies across design research approaches.

Finally, most relevant sampling guidance is found in the social science and psychology literatures. This typically emphasises scientific concerns and neglects interactions with practitioners and the design framing. This can lead to a perception of conflicting directives, hindering application to design research. Thus, there is a need to set out a framework to facilitate more structured and detailed discussion and development of sampling considerations in this context.

1 Approach and definitions

Given the above need, we take a first step towards unpacking key sampling considerations. Importantly, while some of these considerations are relevant to a range of research approaches, we do not provide a universal guide to sampling in design research. Further, due to the need to synthesise and abstract sampling knowledge it is necessary to adopt a theory-driven approach to identifying sampling considerations. Specifically, we develop an initial set of considerations adapted from related fields. This establishes a foundation for discussion and a potential point of departure for future review, meta-analysis, and refinement, as well as examination of intersections between design and other fields (McComb & Jablokow, 2022). This follows similar work in related fields where synthesis of methodological knowledge provides an essential basis for subsequent review and analysis (Onwuegbuzie & Leech, 2007), as well as the approach taken by Cash (2018, 2020) in his synthesis and meta-analytical review of theory-development in design research. Thus, our approach has two components: i) to suggest an initial terminology, literature, and framework for discussing sampling in this context, and ii) to contribute to a broader discussion of methodological and research synergies across design research.

To this end, we survey, distil, and adapt sampling guidelines from over twenty research fields, all with a common foundation in theory development. These distilled guidelines form the basis for each of our considerations. To concretise these considerations and bring them into the design research context, we provide illustrative examples drawn from the *Design Studies* journal. Importantly, due to the formative nature of this area of design research methodology we do not carry out a systematic review, rather we contextualise and exemplify our considerations. However, before we start, it is first necessary to establish some basic definitions relevant to sampling discussions in this context (Robson & McCartan, 2011, p. 276):

- ***population***: the total set of relevant cases (e.g. all designers working with digital technologies); and
- ***sample***: a subset of the *population* (e.g. designers working with digital technologies in one company).

Sampling in this context, thus reflects an interplay between population and sample, forming a bridge between theoretical and methodological discussions rooted in the theory development cycle.

2 Sampling: eight key considerations

Design research studies are often motivated by challenges in theory or practice and conclude with contributions that close knowledge gaps, and subsequently enable the next generation of research and insight. This corresponds to a typical progression from qualitative theory building to more quantitative theory testing in an iterative cycle (Cash, 2020; Wallace, 1971, p. 18). Our purpose and the nature of this contribution provides a common foundation for sampling decisions (Wacker, 2008). Thus, we illustrate a *double-loop sampling process*, Figure 1.

The inner loops in Figure 1 comprise: a *definition loop* used to describe and scope a sample, linked to a *refinement loop* pointing to potential combinatory strategies. However, before entering these loops a researcher must consider both the *scientific* and *practice* concerns that define the design framing of the research. Based on the balance between scientific and practice concerns, a researcher might enter the definition loop at any point. For example, a researcher focused on testing general hypotheses, might start at 'theoretical framing' thereby needing the most statistically representative sample possible, prioritising scientific concerns; while a researcher focused on theory building, might also start at 'theoretical framing', but with the aim to identify a particularly interesting sample, balancing scientific and practice concerns. Finally, a researcher focused on issues of application in collaboration with a specific company, might start at 'sample size', prioritising practice concerns, with a secondary reflection on scientific concerns. Whatever the starting point, researchers should examine all considerations to fully delineate their sample.

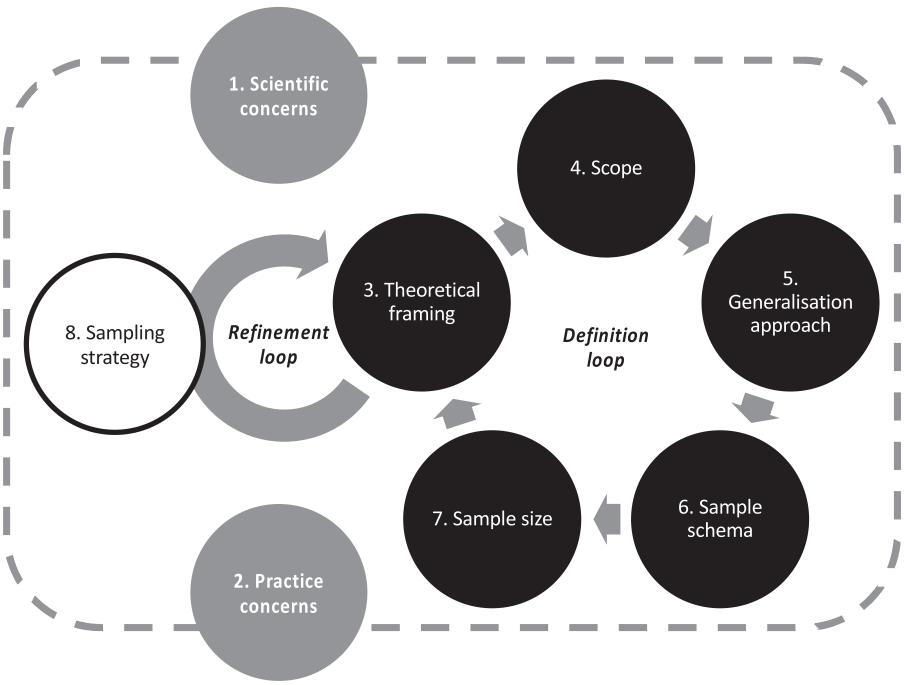

Figure 1 The double-loop sampling process, with consideration numbers

2.1 Framing design research: scientific and practice concerns

Essential to design research is its potential for both scientific and societal impact. Therefore, before entering the double-loop process it is vital to establish the design research framing in terms of both good scientific conduct and link to practice. These form foundational considerations that inform how all other considerations are interpreted and applied.

2.1.1 Good scientific conduct and ethical appropriateness

Key to any research endeavour is a firm grounding in good scientific conduct and research integrity. This positions scientific and ethical appropriateness as the first foundational consideration; which concerns both the integrity of the research and the individuals involved, thereby ensuring sample members are fairly treated and protected (Kitchenham et al., 2002; Onwuegbuzie & Collins, 2007). Focusing on research integrity, this includes ensuring the sample represents the intended population, management of possible biases, and transparency in research decision making, reporting, and communication of limitations (Onwuegbuzie & Collins, 2007). Any mismatch between the proposed sample and population can introduce bias (e.g. overrepresentation of familiar and research-aware participants or excluding portions of the sample or population). A useful and relevant overview can be found in the American Educational Research Association (2011) code of ethics. Particularly critical for design researchers, is the involvement of, and engagement with, practitioners, where there might be ongoing collaborations or limited access to possible sample members. These concerns span the whole research process and can appear to be in tension with practice concerns, thus necessitating careful review and reporting. For example, Ventrella, Zhang, and MacCarty (2020) provide some discussion of ethical concerns impacting their sample. However, they also highlight the need for greater discussion of ethical considerations in general. Given the limited scope here, we recommend the above noted code of ethics, as well as the works of Rosenthal (1994), Robinson (2014), and Creswell (2012).

Consideration 1: Scientific good conduct: what ethical concerns are relevant?

2.1.2 Ensuring impactful research: linking to practice

Key to any design research endeavour is balancing scientific and practice concerns. This positions linking to practice as the second foundational consideration; which puts an emphasis on connecting sampling considerations to a study's framing in design, as well as understanding how sampling can support communication, involvement, and impact in practice (Sjøberg et al., 2002). For example, if a researcher aims to impact practice by helping improve organisational processes in a collaborating company, then their research must foster long–term relationships, company buy-in and (hopefully enthusiastic) engagement, as well as other aspects of organisational change management

(Chakrabarti & Lindemann, 2016) not necessarily required by the contribution to knowledge (Section 2.2.1). Further, design researchers examine a wide range of – often messy – design situations using a variety of approaches and perspectives, which influence where to enter the sample definition loop (Figure 1), as well as how the specific considerations should be approached. Here, practitioners can have diverse roles in, and expectations for, research impact, which must be carefully managed in the context of theory development studies. However, impact can be as simple as providing structured feedback and research will often receive greater attention if the impact on practice is explicit, as discussed by Zielhuis, Sleeswijk Visser, Andriessen, and Stappers (2022). As such, design framing and engagement with practitioners is essential (Cooper, 2019; Easterbrook, Singer, Storey, & Damian, 2008).

Consideration 2: Design framing: what type of impact on practice do you hope to achieve?

2.2 Definition loop

There are five main considerations in this loop: theoretical framing, scope, generalisation approach, sample schema, and sample size.

2.2.1 Theoretical framing

In the theory development context, appropriateness of a population/sample – both qualitative and quantitative – is typically defined with respect to its intended contribution to knowledge (Wacker, 2008). This influences choices for population/sample, as well as the degree of granularity expected in their definition (Lynch, 1999; Wacker, 2008).

In general, the number of potential factors in any situation far exceeds reporting or reasonable conceptualisation. Therefore, definition of population/sample is typically limited to only those factors that affect the concepts under study (Lynch, 1999; Wacker, 2008) and is a key part of the coupling between the research question and method (Goldschmidt & Matthews, 2022). This gives rise to a '*previous literature convention*' (Wacker, 2008), i.e. if a factor has not been related to the theory of interest in a study then it is not required for defining the population/sample. For example, there are multiple factors linked to designer expertise, including diverse skills and domain knowledge. As such, while Kavakli and Gero (2001) are able to robustly differentiate novice (second year students) and expert (more than 25 years of experience) designers' sketching processes, the lack of granularity in the sample definition means that more detailed interpretation of what specific factors determine these differences is impossible. Thus, the more mature the body of theory, the more specifically relevant factors can be identified and defined (Cash, 2018; Melnyk & Handfield, 1998).

Table 1 The theory-building/theory-testing research cycle and its implications for sample definition

	Step	*Typical research questions*	*Sampling implications*
Theory-building	Discovery and description	What is happening in a context, is it interesting? What are the major issues? What is the context of the phenomenon?	Typically, more focused on understanding a phenomenon in a specific group and context, via rich, deep data.
	Definition of variables and domain	What are key variables, assumptions, boundaries, or other features? Who, what, and initial when and where?	Knowledge of a phenomenon is codified, moving from context specific, ill-defined factors, to abstract information regarding key factor, variables, relationships, and predictions. Degree of generality (see Section 2.2.2) can be varied to examine the link between population and sample with respect to who, when, and where it applies.
	Relationship building	What are the links between variables? Are there patterns, or other interactions? What links could and should exist? Why, how?	
Theory-testing	Prediction, testing and validation	What is the predictive power? Do proposed relationships hold true in empirical data? Could something occur, should it, would it?	
	Extension and refinement	How widely applicable is the theory, how robust are predictions across contexts? What are the domain criteria? Refine when, where?	Typically, more focused on understanding scope of a theory across groups and contexts via selected, wide data.

Cash (2018) describes levels of theory development in a cycle from '*discovery and description*' to '*extension and refinement*'. In the early phases of this cycle, where specific factors are typically ill-defined, research focuses on rich, reflective description of phenomena and context. This places an emphasis on being able to explain the significance of a sample in context. For example, Carlson, Lewis, Maliakal, Gerber, and Easterday (2020) link their sampling decisions to their theory building aim, delineating the scope of their work and limiting the granularity of their sampling criteria. Subsequently, they use theoretical sampling to refine, develop, or refute insights from their initial rich data. Together, these support their qualitative description of metacognitive processes underpinning novice design work.

In the later phases of this cycle, where factors are typically well defined, research focuses on specific definition and control. This places an emphasis on being able to define and isolate the key factors linking a population and sample. Research in this context typically employs quantitative studies that can only function when the number of variables can be limited (Easterbrook et al., 2008; Melnyk & Handfield, 1998). For example, Vandevenne, Pieters, and Duflou (2016) detail a range of general and specific aspects of domain knowledge in their sample, which might impact the outcomes of their experiment. As such, they can limit the number of variables introduced by the prior knowledge of their sample participants, and thus more clearly elucidate their hypotheses. Table 1 summarises the maturity of theory in terms of this cycle (Cash, 2018).

Consideration 3: Theoretical framing: where in the theory-building/theory-testing research cycle is current knowledge?

2.2.2 Generalisability

One of the major challenges in discussing 'generalisability' is that the term has various meanings (Lee & Baskerville, 2003; Onwuegbuzie & Leech, 2007), and, as typically used, conflates two dimensions that more broadly define the *domain* of a contribution (Wacker, 2008):

- ***Generalisability***: The extent to which a contribution applies to existing populations ("who" factors such as nationality, organisation, or experience)
- ***Abstraction***: The extent to which a contribution is bounded by time and space ("when" and "where" factors such as location, timing, or scale)

Theory that is both general and abstract, applies to all people at all times in all places, and is often referred to as general or universalistic (Bello, Leung, Radebaugh, Tung, & Van Witteloostuijn, 2009; Gainsbury & Blaszczynski, 2011; Stevens, 2011). However, theory that applies to particular individuals in an explicit time and place is referred to as specific or particularistic (Stevens, 2011). When an intended contribution is universalistic, sampling becomes more about verifying the supposed universality when considering well defined factors. For example, Blizzard et al. (2015) use a stratified random sample across college students in the U.S., in order to evaluate universalistic design thinking traits. In contrast, the more *specific* a contribution, the greater the priority on being sensitive to emerging contextual factors that might

Table 2 A typology of generalisation approaches in terms of *generalisability* and *abstraction*

Generalisability/ Abstraction	*Generalisation approach*
Low/Low focused on within sample understanding	***Case-to-case transfer*** *(see also transferability)* Making generalisations from one case to another based on in-depth descriptions of each specific case that allow readers to make inferences about applicability to other contexts
Low/Low focused on within sample understanding	***Internal statistical generalisation*** Making generalisations from selected participants to the overall sample. This requires mature theoretical definition of variables relevant to within sample variation
Variable/Variable depending on the theory used	***Analytical generalisation*** *(see also theoretical generalisation)* Making generalisations from specific cases to other contexts via theory. This requires transforming specific understanding into general insights applicable to other cases where the theory is relevant
High/Variable depending on the theory used	***External statistical generalisation*** Making generalisations from sample to population based on data extracted from a representative sample of the population. This requires mature theoretical definition of variables relevant to the population/sample link

impact the findings. For example, Crilly and Moroşanu Firth (2019) focus on a small set of specific cases in order to develop rich contextualised insights. Thus, before considering the sample, it is necessary to scope the population. Hence, an initial consideration is definition of a population based on desired generalisability and abstraction (Lee & Baskerville, 2003; Onwuegbuzie & Leech, 2007; Robinson, 2014).

Consideration 4: Scope: how general and abstract is the intended contribution?

Given a population, the sample depends on the variables and relationships in focus as well as the planned generalisability and abstraction of the contribution (Bello et al., 2009; Lynch, 1999; Wacker, 2008). Numerous meta-analyses have demonstrated that this link must be evaluated on a variable-by-variable, relationship-by-relationship basis (Dasgupta & Hunsinger, 2008; Peterson, 2001). For example, in the design fixation literature, there are a wide range of variables and relationships that can vary substantially across studies (Crilly, 2019b; Vasconcelos & Crilly, 2016). A sample is thus defined based on the intended applicability of a contribution across the wider population (Robinson, 2014; Tuckett, 2004). This leads to a number of approaches, both qualitative and quantitative, each with implications for sampling (Onwuegbuzie & Leech, 2007). Thus, the second consideration here is the identification of an approach based on the desired generalisability and abstraction within the population, as outlined in Table 2 (Polit & Beck, 2010; Stevens, 2011; Wacker, 2008).

Consideration 5: Generalisation approach: what type of generalisability and abstraction is desired in the population?

With respect to Considerations 4 and 5 it is pertinent to contrast the generalisation approaches of qualitative and quantitative studies. For the qualitative, we highlight the work of Crilly and Moroşanu Firth (2019) who use thematic analysis to distil insights from select, in-depth cases. Although not explicit in their text, they employ analytical generalisation (Table 2), developing their insights by moving between their results, prior literature, and theory. In contrast, Graff et al. (2020) quantitatively examine a scale for evaluating perceived analogical communication. They test the robustness of this scale across three large samples to develop external statistical generalisation (Table 2), using a carefully defined population composed of graduate students.

2.2.3 *Sampling schema and size*

How the sample will be collected (sampling schema) and size link the prior considerations to research method (Onwuegbuzie & Collins, 2007). Schema are split into *probability* and *non-probability* (Daniel, 2012; Onwuegbuzie &

Collins, 2007), corresponding to statistical and analytical/cases-to-case generalisation (Table 2) (Onwuegbuzie & Leech, 2007; Polit & Beck, 2010). For example, Crilly and Moroşanu Firth (2019), use a number of criteria to develop a purposive sample of cases with specific characteristics of interest (e.g. the development of radically new, physical, commercial products, with potential for in-depth data access). This constrains the scope of their investigation and provides a basis for developing rich, in-depth insights. In contrast, Blizzard et al. (2015) use a probability sample in order to develop a general understanding of design thinking traits across universities, and student groups in the U.S. This serves to delineate the scope of their claims and provide a concrete basis for identifying sample participants. Probability samples use mathematical rules to ensure that everyone in a population has the same chance of being included in the sample, while non-probability samples include individuals based on a range of criteria. Non-probability can thus be further decomposed as (Creswell, 2012; Daniel, 2012):

- *purposive* (purposeful, judgemental, selective or subjective): based on the characteristics of the population and research purpose (considerations 3-5),
- *quota*: based on a stratified quota, and
- *convenience*: based on availability.

An overview of possible schemas is given in Figure 2 (Onwuegbuzie & Collins, 2007; Teddlie & Yu, 2007). Critically, each schema has implications for both the interpretation of the study and results as well as the practical identification of participants – as highlighted by Hay, Duffy, Gilbert, and Grealy (2022) in the design cognition context – yet few design research studies explicitly report the specific schema adopted.

Consideration 6: Sample schema: what schema fits your theoretical framing, generalisation approach, and research method?

Sample size is based on research purpose, generalisation approach, and research method. As such, both qualitative and quantitative research methods can draw on small or large samples depending on the other considerations in Figure 1, as well as the type of data collected and analysis approach used (Onwuegbuzie & Collins, 2007; Sandelowski, 1995). A number of authors provide minimum size guidelines for specific research methods, for example interview studies with >12 participants, one-tailed experiments with >21, and case studies with >4, summarised in Figure 3 (Onwuegbuzie & Collins, 2007). However, it is typical that qualitative samples should be large enough to support saturation (further data collection would only confirm the results already identified) and small enough to deliver rich insight (Onwuegbuzie & Leech, 2007; Sandelowski, 1995; Teddlie & Yu, 2007). This trade-off between breadth and depth is clearly illustrated by Crilly and Moroşanu Firth's (2019) deliberate constraint of their sample to only three cases, which they are able to

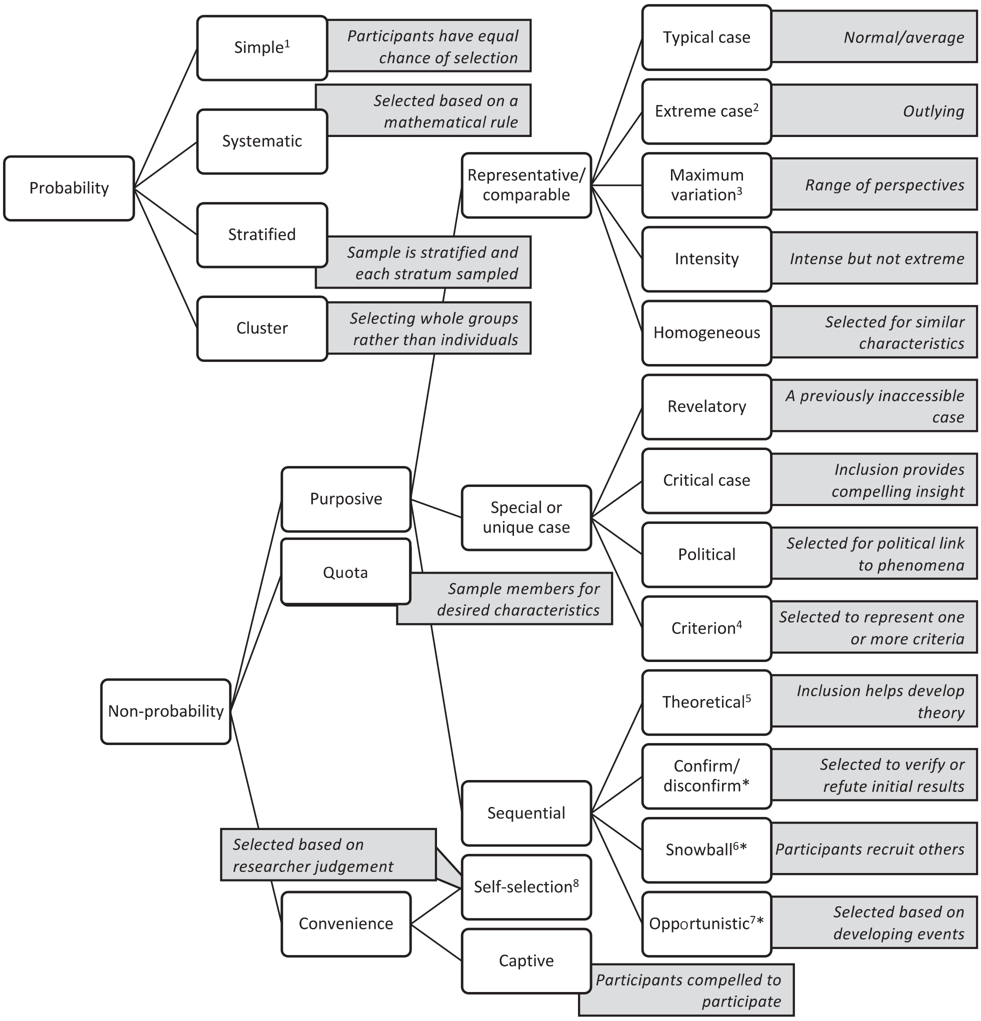

Figure 2 A typology of sampling schema split into probability and non-probability, with branching sub-types and summary definitions; numbering denotes alternative names: [1]Random, [2]Deviant, Outlier, [3]Heterogeneous, [4]Complete collection, [5]Theory-based, Concept, [6]Chain, Network, Reputational, [7]Volunteer; *∗denotes that the schema is implemented after data collection has begun.*

analyse in great depth. Similarly, quantitative samples should typically meet the statistical requirements of the generalisation approach, such as significance and statistical power. This is key when comparing experiments and their effects, as illustrated in the meta-analyses of design fixation by Sio, Kotovsky, and Cagan (2015) and Vasconcelos and Crilly (2016). While the numbers given for different approaches in Figure 3 are guidelines only, and sample size should always be justified with respect to the specific study, they provide an important point of reference and help normalise sample size discussions across studies within a field. For example, Nelius et al. (2020) use an experimental study in a small-scale theory building mode, which can lead to confusion if

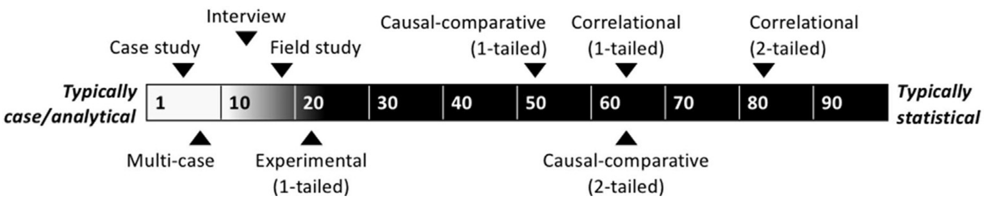

Figure 3 Sample size and research method

experimental and *quantitative* are conflated. Hence they explicitly contextualise their sample size discussion with respect to the guideline of Onwuegbuzie and Collins (2007). As such, this is key to the transparent positioning of the study in the wider research context, as well as in understanding the scope of the work.

Consideration 7: Sample size: what size fits with your generalisation approach and research method?

2.3 Refinement loop: sampling strategy

Key to understanding the value of studies in context – particularly with respect to their contribution to the wider qualitative, quantitative theory development cycle – is that no sample can fulfil all possible research demands. Therefore, when looking beyond a single study, it is necessary to consider how combinations of samples can be used to mitigate individual weaknesses and maximise collective strengths (Onwuegbuzie & Leech, 2007; Teddlie & Yu, 2007). This mirrors discussions that illustrate the strengths of combinatory mixed methods research designs (Cash & Snider, 2014; Hanson, Creswell, Plano Clark, Petska, & Creswell, 2005; Onwuegbuzie & Leech, 2006; Tashakkori & Teddlie, 2008). For example, when dealing with more constrained or quantitative studies it is often necessary to prioritise either *internal* or *external* validity (Gainsbury & Blaszczynski, 2011) (i.e. ***internal validity***: the extent to which evidence supports conclusions (usually causal) within the context and integrity of a specific study; ***external validity***: the extent to which conclusions from a specific study apply to other contexts with implications in the wider world). This leads to sampling strategies such as that by Cash, Hicks, and Culley (2013), who combine three studies: i) a qualitative study with a practitioner sample fully embedded in context, ii) a semi constrained study with a practitioner sample emphasising external validity, and iii) a semi constrained study with a student sample emphasising internal validity. While each study is limited in isolation, together they provide a rich picture linking real world practitioner and laboratory-based student practice. Another example of effective sample combination can be found in the work on Ventrella et al. (2020), who employ multiple samples in order to develop and evaluate a novel sensor system. Here, the combination of samples provides several insights and allows for a progression in research objectives from

Table 3 An overview of combinatory sampling strategies

Sampling strategy	*Description*
Schema level (Onwuegbuzie & Collins, 2007; Teddlie & Yu, 2007) i.e. combinations of schema from Figure 2	
Multiple probability (see also multi-stage random)	Using multiple probability schema in the same study
Multiple purposive	Using multiple purposive schema in the same study
Mixed purposive	Using more than one sampling schema to compare the results from the different approaches
Multi-stage purposive	Choosing samples in two or more stages, where all stages are purposive
Multi-stage purposive random	Choosing samples in two or more stages, where the first stage is probability, and subsequent stages are purposive
Multi-stage stratified purposive	Specific combination of first probability stratified sampling followed by purposive sampling
Overall level (Onwuegbuzie & Leech, 2007) i.e. combinations of research methods and studies	
Parallel	Using multiple schema to facilitate the comparison of cases (pairwise) or subgroups of cases (subgroup) to others in the sample
Concurrent mixed methods sampling	Using multiple studies employing probability and non-probability schema in parallel, where the studies are independent
Nested	Using multiple schema to facilitate the comparison of sub-samples within a subgroup
Multi-level mixed methods sampling	Using different schema at different levels of study, for example, using purposive sampling at the organisation level and probability sampling at the employee level, or using different schema at different levels of hierarchy amongst employees
Sequential mixed methods sampling	Using multiple studies employing probability and non-probability schema in sequence, where the first study informs the sampling used in the second

understanding the specific challenge to evaluating the final usability of the sensor. There are many recommendations for combinatory sampling, ranging from specific schema analogous to those outlined in Figure 2, to large-scale strategies that shape the whole research design, summarised in Table 3.

Consideration 8: Sampling strategy: (*when using multiple studies*) what combination of sample schema provide the best balance with respect to all prior considerations?

3 *Limitations and further work*

Before discussing implications, it is necessary to highlight two main limitations of this work. First, sampling is a broad topic, with common considerations but also many field-specific adaptations. While our work has focused on key considerations for one aspect of design research, further work is needed to i) examine the specific instantiation of these considerations across the design literature, ii) elaborate and detail their application and reporting, and iii) explore commonalities and differences across design research philosophies and approaches. Second, sampling discussions in related fields have built on a critical interplay between conceptual development used to synthesise knowledge, and critical literature review used to interrogate this. Due the lack of prior discussion or guidance in the design literature, our work necessarily focused on conceptual development i.e. synthesizing the considerations. This points to the subsequent need for i) systematic review of sampling practices across design research, ii) meta-analysis of interactions between sampling, method use, and knowledge outcomes, and iii) evaluation of potential areas requiring further adaptation of or new considerations. Thus, despite these limitations, our work provides an important point of departure and comparison for further work, and mirrors developments in related fields where guidelines are iteratively proposed, interrogated, and refined.

4 Implications and reporting

The considerations detailed in Section 2, draw a link between intended contribution to knowledge and specific research sample. In Table 4 we connect each consideration to recommended actions, relevant resources in this article, and suggested reporting, as well as providing examples of current good practice in design. Here, we reiterate the foundational nature of considerations 1 and 2 in establishing the design framing, and subsequently informing the interpretation and application of all other considerations. Reporting the considerations explicitly as exemplified in Tale 4 is key to supporting meta-analysis or literature review (Chai & Xiao, 2012; Vasconcelos & Crilly, 2016), and is also key to discussions of method appropriateness (Goldschmidt & Matthews, 2022), limitations, and research quality (Prochner & Godin, 2022). For example, Graff et al. (2020) provide a careful discussion of the strengths and weaknesses of their student sample, and reflect on the potential transferability of the insights to other populations, based on the theoretical variables at play and how they are expressed in different groups.

In addition to Table 4, the considerations also point to two implications related to typical sampling discussions in the theory development context. First, the strengths and weakness of samples employing, for example, students versus practitioners, can be potentially leveraged via variable-by-variable generalisation (Section 2.2.2), and sample combination with respect to the wider theory development cycle (Section 2.3). This aligns with results of multiple meta-analyses from across domains, including social science (Peterson, 2001), information and management (King & He, 2006; Schepers & Wetzels, 2007), and software engineering (Hannay, Dybå, Arisholm, & Sjøberg, 2009). This leads to a summary of student sample pros and cons (Table 5).

Table 4 Summary of actions, resources, reporting guidance, and design research examples for the eight considerations. References are omitted from the quoted examples for clarity

#	*Consideration*	*Actions*	*Resource*	*Reporting*	*Design research example*
1	Scientific good conduct: What ethical concerns are relevant?	Evaluate potential concerns and report their mitigation	Code of ethics in Section 2.1.1	In method introduction, sampling, and implications	"*The university's health research ethics review board approved this study and respondents provided written informed consent prior to their interview.*" (Rivard, Lehoux, & Hagemeister, 2021, p. 6)
2	Design framing: What type of impact on practice do you hope to achieve?	Concretise the design framing and evaluate potential mechanisms for involving practice with your sample	e.g. Sjøberg et al. (2002) in Section 2.1.2	In method introduction, sampling, and implications	"*this case study demonstrates how the co-operative work term experience could be enhanced by the addition of an academic tutor, and at the same time how the traditional and virtual studios could be enhanced by immersing students in a real world context*" (Nespoli, Hurst, & Gero, 2021, p. 24)
3	Theoretical framing: Where in the theory-building/theory-testing research cycle is current knowledge?	Establish research purpose; define variables, and their level of detail	Table 1	In research framework and method introduction	"*Previous studies revealed a potential influence of experience on confirmation bias. To evaluate both experienced and less experienced participants, the study was conducted with mechanical engineering students and experienced design engineers from industry*" (Nelius et al., 2020, p. 6)
4	Scope: How general and abstract is the intended contribution?	Establish scope (who, when and where); define inclusion/exclusion criteria	Bullets in Section 2.2.2 and Wacker (2008)	In method introduction and sample definition	"*To permit the study of creative work in the design process, we sought project examples where a change in the product was clearly evident … In identifying projects for inclusion in the study, four criteria were applied: …*" (Crilly & Moroşanu Firth, 2019, p. 170)

(*continued on next page*)

Table 4 (*continued*)

#	*Consideration*	*Actions*	*Resource*	*Reporting*	*Design research example*
5	Generalisation approach: What type of generalisability and abstraction is desired in the population?	Establish a generalisation approach and specific research objectives	Table 2	In method introduction and sample definition	"*Given our research question and aim to develop theory, a qualitative, multi-case research design was adopted. Here, a robust means of developing analytical and theoretical generalisability is selecting cases that exemplify theoretical opposites*" (Cash, Dekoninck, & Ahmed-Kristensen, 2020, p. 8)
6	Sample schema: What schema fits your theoretical framing, generalisation approach, and research method?	Establish and justify a specific schema and define its components	Figure 2	In sample definition and method selection	"*A random sample of institutions was recruited for participation using a comprehensive list of U.S. colleges and universities … The sample was stratified by institution type and the number of enrolled students: …*" (Blizzard et al., 2015, p. 94)
7	Sample size: What size fits with your generalisation approach and research method?	Establish and justify a specific sample size	Figure 3	In sample definition and method selection	"*samples were sized in order to provide a minimum conditional N of ∼30 following best practice sample size guidance for experimental studies*" (Cash & Maier, 2021, p. 10)
8	Sampling strategy: What combination of sample schema provide the best balance with all prior considerations?	Optionally form a sampling strategy and complete the definition loop for each sample	Table 3	In method introduction and sample definition	"*we used a design science approach that integrated rapid ethnographic and sensor-based methods in a multi-site case study and roughly followed the stages of the design process.* Table 2 *shows the overall progression of research phases, research goals, and methods used.*" (Ventrella et al., 2020, p. 88)

Second, scientific and practical concerns need to be combined to effectively balance the needs of impact on theory and practice (Section 2.1). This leads to three main insights: i) in most cases sampling decisions are driven by research purpose and theoretical framing; ii) mixed methods may be considered in any context; iii) practice may be engaged by design researchers to allow for feedback and to increase research credibility, transparency, and

Table 5 Typical pros and cons of a student sample in comparison to a practitioner sample

Student samples … (pros)
Are generally more ideal for internal validity due to homogeneity and ability to follow complex study designs
Are generally more ideal for statistical validity due to the large numbers of accessible participants
Are more susceptible to 'induced value' and are thus easier to induce preferences in for study purposes
Are more ideal for testing causal theory due to the above characteristics
Are more ideal for pre-testing, study design refinement and obtaining preliminary evidence
Present interesting variables in and of themselves e.g. they are often more open to learning and development
Student samples … (cons)
Are generally more limited in diversity/environment, but this needs to be checked on a variable-by-variable basis
Are generally less ideal for external validity, but this needs to be checked on a variable-by-variable basis, particularly if specific types of experience or practice-specific factors are relevant
Require greater explanation of moderating effects on the focus variable
Are less ideal for describing single effects or highly situated phenomena

comprehensibility, particularly if evidence is drawn from student samples (Sjøberg et al., 2002).

5 Conclusion

Effective sampling forms an essential element in developing design research that impacts both theory and practice. Here, we have brought together guidance from diverse research fields, to take a first step towards developing a terminology, literature, and framework for discussing key sampling considerations in design research. The double-loop sampling process and associated considerations (Figure 1) are a first of its kind in design research, and constitute a call to action, highlighting the need for further examination and theorising around the impact of sampling on design research claims. In doing so, we aim to contribute to a broader discussion of methodological and research synergies across design research approaches.

Ultimately, any study will inevitably have limited insight, and thus transparency in the reporting of methodological and sampling considerations is increasingly important, particularly in the diverse context of design research. Thus, we encourage researchers to report these considerations in their papers, to make justification, assumptions, and limitations explicit.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was funded by DTU, Technical University of Denmark. We would like to thank the organisers of the DESIGN conference for sparking this work, as well as Dr. Nathan Crilly and Dr. Laura Hay for their invaluable input during the development of this paper. I would also like to single out Dr. Nicholas Ciccone for his help in bringing a psychological perspective to the paper. Finally, I would like to thank the reviewers and editors for helping to further develop this manuscript.

References

American Educational Research Association. (2011). Code of ethics: American educational research association. *Educational Researcher, 40*(3), 145–156.

Ball, L. J., & Christensen, B. T. (2018). Designing in the wild. *Design Studies, 57*, 1–8, 1995.

Bello, D., Leung, K., Radebaugh, L., Tung, R. L., & Van Witteloostuijn, A. (2009). From the Editors: Student samples in international business research. *Journal of International Business Studies, 40*(3), 361–364.

Blessing, L. T. M., & Chakrabarti, A. (2009). *Drm: A design research methodology*. London: Springer.

Blizzard, J., Klotz, L., Potvin, G., Hazari, Z., Cribbs, J., & Godwin, A. (2015). Using survey questions to identify and learn more about those who exhibit design thinking traits. *Design Studies, 38*, 92–110.

Carlson, S. E., Lewis, D. G. R., Maliakal, L. V., Gerber, E. M., & Easterday, M. W. (2020). The design risks framework: Understanding metacognition for iteration. *Design Studies, 70*, 100961.

Cash, P. (2018). Developing theory-driven design research. *Design Studies, 56*(May), 84–119.

Cash, P. (2020). Where next for design research? Understanding research impact and theory building. *Design Studies, 68*, 113–141.

Cash, P., Daalhuizen, J., & Hay, L. (2022). Design research quality. *Design Studies*. (IN PRESS).

Cash, P., Dekoninck, E., & Ahmed-Kristensen, S. (2020). Work with the beat: How dynamic patterns in team processes affect shared understanding. *Design Studies, 69*(July), 1–39.

Cash, P., Hicks, B. J., & Culley, S. J. (2013). A comparison of designer activity using core design situations in the laboratory and practice. *Design Studies, 34*(5), 575–611.

Cash, P., & Maier, A. M. (2021). Understanding representation: Contrasting gesture and sketching in design through dual-process theory. *Design Studies, 73*, 100992.

Cash, P., & Snider, C. (2014). Investigating design: A comparison of manifest and latent approaches. *Design Studies, 35*(5), 441–472.

Chai, K. H., & Xiao, X. (2012). Understanding design research: A bibliometric analysis of design studies (1996 - 2010). *Design Studies, 33*(1), 24–43.

Chakrabarti, A., & Lindemann, U. (2016). In A. Chakrabarti, & U. Lindemann (Eds.), *Impact of design research on industrial practice*. Switzerland: Springer International Publishing.

Cooper, R. (2019). Design research - its 50-year transformation. *Design Studies, 65*, 6–17.

Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (4th ed.). Boston, US: Pearson.
Crilly, N. (2019a). Creativity and fixation in the real world : A literature review of case study research. *Design Studies, 64*, 154—168.
Crilly, N. (2019b). Methodological diversity and theoretical integration: Research in design fixation as an example of fixation in research design? *Design Studies, 65*, 78—106.
Crilly, N., & Moroşanu Firth, R. (2019). Creativity and fixation in the real world: Three case studies of invention, design and innovation. *Design Studies, 64*, 169—212.
Daniel, J. (2012). *Sampling essentials*. London, UK: Sage.
Dasgupta, N., & Hunsinger, M. (2008). The opposite of a great truth is also true: When do student samples help versus hurt the scientific study of prejudice? *Psychological Inquiry, 19*(2), 90—98.
Douglas, B. C., Noble, L. M., & Newman, P. S. (1999). Improving the accuracy of patients' expectations of the psychiatric out-patient consultation. *Psychiatric Bulletin, 23*(7), 425—427.
Easterbrook, S., Singer, J., Storey, M.-A., & Damian, D. (2008). Selecting empirical methods for software engineering research. In F. Shull, J. Singer, & D. I. K. Sjoberg (Eds.), *Guide to advanced empirical software engineering* (pp. 285—311). Springer.
Gainsbury, S., & Blaszczynski, A. (2011). The appropriateness of using laboratories and student participants in gambling research. *Journal of Gambling Studies, 27*(1), 83—97.
Goldschmidt, G., & Matthews, B. (2022). Formulating design research questions: A framework. *Design Studies*. (IN PRESS).
Graff, D., Meslec, N., & Clark, M. A. (2020). Perceived analogical communication in design teams: Development and validation of a scale. *Design Studies, 67*, 55—76.
Hannay, J. E., Dybå, T., Arisholm, E., & Sjøberg, D. I. K. (2009). The effectiveness of pair programming: A meta-analysis. *Information and Software Technology, 51*(7), 1110—1122.
Hanson, W. E., Creswell, J. W., Plano Clark, V. L., Petska, K. S., & Creswell, J. D. (2005). Mixed methods research designs in counseling psychology. *Journal of Counseling Psychology, 52*(2), 224—235.
Hay, L., Duffy, A., Gilbert, S., & Grealy, M. (2022). Functional magnetic resonance imaging (fMRI) in design studies: Methodological considerations, challenges, and recommendations. *Design Studies*. (IN PRESS).
Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems, 19*(2), 87—92.
Kavakli, M., & Gero, J. S. (2001). Sketching as mental imagery processing. *Design Studies, 22*(4), 347—364.
King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management, 43*(6), 740—755.
Kitchenham, B. A., Pfleeger, S. L., Pickard, L. M., Jones, P. W., Hoaglin, D. C., El-Emam, K., et al. (2002). Preliminary guidelines for empirical research in software engineering. *IEEE Transactions on Software Engineering, 28*(8), 721—734.
Le Dain, M., Blanco, E., & Summers, J. D. (2013). Assessing design research quality: Investigating verification and validation criteria. In *ICED 13 international conference on engineering design* (pp. 1—10), (Seoul, South Korea).
Lee, A. S., & Baskerville, R. L. (2003). Generalizing generalizability in information systems research. *Information Systems Research, 14*(3), 221—243.

Lynch, J. G. (1999). Theory and external validity. *Journal of the Academy of Marketing Science, 27*(3), 367–376.
McComb, C., & Jablokow, K. (2022). A conceptual framework for multidisciplinary design research with example application to agent-based modeling. *Design Studies*. (IN PRESS).
Melnyk, S. A., & Handfield, R. B. (1998). May you live in interesting times...the emergence of theory-driven empirical research. *Journal of Operations Management, 16*(4), 311–319.
Nelius, T., Doellken, M., Zimmerer, C., & Matthiesen, S. (2020). The impact of confirmation bias on reasoning and visual attention during analysis in engineering design: An eye tracking study. *Design Studies, 71*, 100963.
Nespoli, O. G., Hurst, A., & Gero, J. (2021). Exploring tutor-student interactions in a novel virtual design studio. *Design Studies, 75*, 101019.
Onwuegbuzie, A. J., & Collins, K. M. T. (2007). A typology of mixed methods sampling designs in social science research. *Qualitative Report, 12*(2), 281–316.
Onwuegbuzie, A. J., & Leech, N. L. (2006). Linking research questions to mixed methods data analysis procedures. *Qualitative Report, 11*(3), 474–498.
Onwuegbuzie, A. J., & Leech, N. L. (2007). Sampling designs in qualitative research: Making the sampling process more public. *Qualitative Report, 12*(2), 19–20.
Peterson, R. A. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of Consumer Research, 28*(3), 450–461.
Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies, 47*(11), 1451–1458.
Prochner, I., & Godin, D. (2022). Quality in research through design projects: Recommendations for evaluation and enhancement. *Design Studies*. (IN PRESS).
Rivard, L., Lehoux, P., & Hagemeister, N. (2021). Articulating care and responsibility in design: A study on the reasoning processes guiding health innovators' 'care-making' practices. *Design Studies, 72*, 100986.
Robinson, O. C. (2014). Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative Research in Psychology, 11*(1), 25–41.
Robson, C., & McCartan, K. (2011). *Real world research* (4th ed.). Chichester: Wiley.
Rosenthal, R. (1994). Science and ethics in conducting, analyzing, and reporting psychological research. *Psychological Science, 5*, 127–134.
Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing & Health, 18*(2), 179–183.
Schepers, J., & Wetzels, M. (2007). A meta-analysis of the technology acceptance model: Investigating subjective norm and moderation effects. *Information & Management, 44*(1), 90–103.
Sio, U. N., Kotovsky, K., & Cagan, J. (2015). Fixation or inspiration? A meta-analytic review of the role of examples on design processes. *Design Studies, 39*, 70–99.
Sjøberg, D. I. K., Anda, B., Arisholm, E., Dybå, T., Jørgensen, M., Karahasanovic, A., et al. (2002). Conducting realistic experiments in software engineering. In *Isese 2002 - proceedings, 2002 international symposium on empirical software engineering* (pp. 17–26).
Stevens, C. K. (2011). Questions to consider when selecting student samples. *Journal of Supply Chain Management, 47*(3), 19–21.

Tashakkori, A., & Teddlie, C. (2008). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage.
Teddlie, C., & Yu, F. (2007). Mixed methods sampling: A typology with examples. *Journal of Mixed Methods Research, 1*(1), 77–100.
Tuckett, a. G. (2004). Qualitative research sampling: The very real complexities. *Nurse Researcher, 12*(1), 47–61.
Vandevenne, D., Pieters, T., & Duflou, J. R. (2016). Enhancing novelty with knowledge-based support for Biologically-Inspired Design. *Design Studies, 46*, 152–173.
Vasconcelos, L. A., & Crilly, N. (2016). Inspiration and fixation: Questions, methods, findings, and challenges. *Design Studies, 42*(C), 1–32.
Ventrella, J., Zhang, S., & MacCarty, N. (2020). Study of the design of a sensor-based system for monitoring impacts of clean energy technologies. *Design Studies, 66*, 82–113.
Wacker, J. G. (2008). A conceptual understanding of requirements for theory-building research: Guidelines for scientific theory building. *Journal of Supply Chain Management, 44*(3), 5–15.
Wallace, W. L. (1971). *The logic of science in sociology*. Ann Arbor, Michigan: Aldine-Atherton.
Zielhuis, M., Sleeswijk Visser, F. S., Andriessen, D., & Stappers, P. J. (2022). Making design research relevant for design practice: what's in the way? *Design Studies*. (IN PRESS).