TU Delft

Delft University of Technology

Document Version
Final published version

Licence
CC BY

Citation (APA)
Cash, P., Daalhuizen, J., & Hekkert, P. (2023). Evaluating the efficacy and effectiveness of design methods: A systematic review and assessment framework. *Design Studies*, *88*, Article 101204. https://doi.org/10.1016/j.destud.2023.101204

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.

Evaluating the efficacy and effectiveness of design methods: A systematic review and assessment framework

Philip Cash, Northumbria University, UK

Jaap Daalhuizen, DTU Technical University of Denmark, Denmark

Paul Hekkert, Delft University of Technology (TU Delft), Netherlands

The increasingly transdisciplinary context of design, where designers collaborate with other disciplinary and domain experts, means there is a growing need to evidence the effectiveness of design methods. We address this need in two ways. First, we propose a 'chain of evidence', from motivation to claims, operationalising this in a systematic assessment framework. Second, we systematically review current design method research. Our results reveal that while all links in the chain of evidence are reported across the literature and best practices can be identified, no individual paper either reports all links or consistently achieves best practice. Our framework and results demonstrate the need for standards of evidence in this area, with implications for design method research, development, education, and practice.
© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Keywords: design methods, design tools, design methodology, design research, systematic review

The increasingly transdisciplinary context of design, where designers collaborate with other disciplinary and domain experts, means there is a growing need to evidence the effectiveness of design methods. Key questions are, what do design methods claim to achieve and what evidence are these claims based on? This comes in addition to method's traditionally central role in design; helping to shape, describe, teach, and explain our discipline (Roozenburg & Eekels, 1995; van Boeijen, Daalhuizen, & Zijlstra, 2020).

Design methods, and their associated claims of improved performance, have reached this status by translating research insights and best practices into real-world impact via education and practice (Blessing & Chakrabarti, 2009; Cantamessa, 2003; Daalhuizen, Person, & Gattol, 2014). The proposal and study of new methods arguably forms *the* central pillar of design research

Corresponding author:
Philip Cash
philip.cash@northumbria.ac.uk

www.elsevier.com/locate/destud
0142-694X *Design Studies* **88** (2023) 101204
https://doi.org/10.1016/j.destud.2023.101204
© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

(impact). Yet, in contrast with almost all similarly important research impact mechanisms (e.g., interventions in education (Levin & O'Donnell, 1999) or health (Gottfredson et al., 2015; Grimes & Schulz, 2002)), there is no consolidated procedure for assessing design methods and their supporting evidence (Gray, 2022). This undermines the very foundations of design research impact and design practice's credibility in contributing to social transformation.

Design methods capture key procedural knowledge to provide '*a formalised representation of a design activity, which functions as a mental tool to* support *designers in achieving a goal, in relation to the circumstances and resources available*' (Daalhuizen, Timmer, van der Welie, & Gardien, 2019). In this form, methods significantly contributed to the emergence of design research as a field in the 1960s and have continued to grow in importance across domains (Jones, 1977; Roozenburg & Eekels, 1995; van Boeijen et al., 2020). However, they have also formed the subject of heated debate, exacerbated by conflicting accounts of how methods impact practice (if at all) (Daalhuizen & Cash, 2021; Dorst, 2008; Jones, 1992; Wallace, 2011), leading to a critical gap between the importance and credibility of design methods. This gap is rooted in the multifaceted nature of rigorous method evaluation (Daalhuizen & Cash, 2021, Figure 3; Gericke, Eckert, & Stacey, 2017, sec. 4.4) and poses a grave challenge to design research (Lloyd, 2019; Meyer & Norman, 2020), as well as leaving numerous disciplines — ranging from health to engineering — reliant on methods built on ambiguous evidence and lacking transparency. Thus, there is a vital need to better understand how design methods and their associated claims can be assessed.

To address this need we build on the recent work of Daalhuizen and Cash (2021), who defined a basic understanding of 'good' method content,[1] and thus provide a foundation for assessing methods more generally. Taking this as a starting point, we first develop a framework for assessing method development, reporting, claims, and supporting evidence, before using this to systematically review and analyse current method research. Throughout, we focus on methods that involve human activities or interaction and thus cannot be evaluated in isolation, in contrast to more technical and/or computational tools whose efficacy can be directly assessed. Our assessment framework and review provide a basis for grounding the current debate on method credibility. This substantially extends research towards a wider *Theory of Design Methods* and has significant implications for method research, development, education, and practice.

1 Background

To assess design methods, it is first necessary to clarify a common understanding of what methods in design are and subsequently how research proposing methods and reporting their associated claims can be evaluated.

1.1 Methods in design

Design methods broadly serve to embody understanding of design work (Bucciarelli, 1994; Stappers & Sanders, 2005), direct design outcomes (Araujo, 2001), and shape design education, skill development, and practices (Andreasen, 2011; Kunrath, Cash, & Kleinsmann, 2020). They do this by supporting belief formation and cognition via an interaction between method content and method user, which we call 'method use' (Daalhuizen & Cash, 2021). This interaction can range from almost complete offloading (e.g., via an algorithmic, computer supported method such as a design structure matrix (Pektaş & Pultar, 2006)) to cognitively intensive (e.g., via a heuristic principle such as satisficing (Simon, 2019)), or any combination thereof (Daalhuizen, 2014). Further, the specific focus of this support and the associated claims is as varied as design itself (Cash, Valles Gamundi, Echstrom, & Daalhuizen, 2022; Kumar, 2013; van Boeijen et al., 2020). For example, claims can include statements about efficacy (e.g., the brainstorming method leads to a greater number of and more diverse ideas compared to using competing methods or no method at all) or effectiveness (e.g. the brainstorming method can be successfully used by multidisciplinary teams in corporate organizations to generate ideas), as well as other direct and indirect outcomes (e.g. the brainstorming method leads to better ideas and also shared understanding in multidisciplinary teams). Thus, the interaction between method content and method user (for simplicity, and to avoid confusion with product user, we refer to the method user throughout as 'designer') provides a common foundation for understanding method impact.

In this context, it is possible to trace a logical link between a method's support for individual belief formation and cognition to at least: i) taskwork processes and effects on outputs and artefacts (e.g. by fostering creativity (Chulvi, Mulet, Chakrabarti, López-Mesa, & González-Cruz, 2012)), and simultaneously ii) teamwork processes and effects on team and environment (e.g. by fostering shared understanding or affect (Vaajakallio & Mattelmäki, 2014)). As highlighted above, methods have been claimed to impact all aspects of the design process, with the ultimate contribution to more successful design work and outcomes (Lewrick, Link, & Leifer, 2018, 2020). However, current approaches to evaluating method impact typically focus on method's effects at the artefact (outcome) or organisation level (Andreasen, Thorp Frey & Dym, 2006; Hansen, & Cash, 2015), and thus provide little insight into the relationship between the basic method/designer interaction and the ultimate impact. This is particularly problematic because design methods are typically heuristic in nature, and thereby enhance chances of producing desirable outcomes, rather than guaranteeing them. This means that good methods that are used properly might be discarded based on undesirable outcomes.

This disconnect between method function and impact evaluation lies at the heart of methodological debate; perhaps best illustrated by Christopher Alexander's dual role as both a founding father of the methods movement and one of its greatest critics (Alexander, 1971). Central to his criticism was the overemphasis on method development at the expense of understanding of how methods work and — most importantly — how they contribute to better design practices and real-world impact. As illustrated in this section, this challenge is as vital today as it was in Alexander's time. Hence, there is a pressing need to better align understanding of method function and impact in the discussion and assessment of methods. This is essential to fostering a more mature and robust research culture around method development, testing, and dissemination.

1.2 Evaluating methods

Current evaluation efforts largely focus on criteria related to consistency and overall impact in terms of design outcomes (Frey & Dym, 2006; Vermaas, 2016). This neglects the key method/designer interaction (see Section 1.1) (Daalhuizen & Cash, 2021; Daalhuizen et al., 2019). For example, a prototypical evaluation framework in the current design literature is the validation square (Seepersad et al., 2006; Vermaas, 2016). This deals with the internal coherence and consistency of method content in terms of evaluating algorithm-like procedures that require information and resources to be processed to produce design outputs, independent of both designer and purpose. While internal consistency is essential to methods' function, the recent work of Daalhuizen and Cash (2021) demonstrates that content is processed by designers with respect to a method's purpose or goal, and therefore content, purpose, and designer response are related. Hence, Daalhuizen and Cash (2021) expand the scope of method evaluation by treating both the internal conceptual coherence of a method and the interaction between the method and the designer.

Daalhuizen and Cash (2021) operationalise this understanding in four key properties of 'good' method content. These broadly align with conceptualisations of 'good' theory and efficacy (Flay et al., 2005; Wacker, 2008), and 'good' artefacts and effectiveness (Araujo, 2001; Daalhuizen, 2014), and include (from Daalhuizen and Cash (2021, Figure 3)).

- **Defined**: The major content variables are logically complete, coherent (i.e. not conflicting), and unambiguously described; and the domain of operation is clear i.e. designers understand in what context(s) the method will perform as described;
- **Predictable**: The internal structure of the method is understandable and predictable i.e. designers can predict how altering one variable will impact the

other variables; and the outcome(s) of interacting with the content is understandable and predictable within the domain of operation;
- **Useable**: The method is accessible, understandable, and credible to the designer;
- **Desirable**: The outcome(s) of interacting with the method is appropriate and valuable.

Based on these properties, Daalhuizen and Cash (2021) predict that method performance will be negatively impacted by incomplete reporting of the content, conflict between the content elements, and conflict between the content and designer/use context. Some of these properties have been operationalized, for example by Tromp and Hekkert (2016) who define process quality, process efficiency, and design quality as key measures to evaluate effect-driven design methods. The four properties of 'good' method content broadly complement prior work on method development by elaborating Daalhuizen et al.'s (2019) discussion of types of methodological elements and their internal hierarchy, as well as Gericke et al.'s (2017) distinction between method evaluation and validation.

More generally, these properties provide an initial basis for evaluating the content of design methods, which broadly mirrors assessment frameworks in related fields such as software engineering (Kitchenham et al., 2002; Kitchenham, Dyba, & Jorgensen, 2004) and prevention science (Flay et al., 2005; Gottfredson et al., 2015). However, this only treats the reporting of the method itself and thus neglects the other major aspect affecting the credibility of methods i.e. the robustness of the supporting evidence for methodological claims (Cash, Daalhuizen, & Hay, 2022; Gottfredson et al., 2015; Kitchenham et al., 2004; Prochner & Godin, 2022). For example, Tromp and Hekkert (2016) argue for the need to combine qualitative and quantitative evidence to credibly evaluate a method, while Vermaas (2016) critiques expert justifications and highlights robust empirical evidence as essential to credible method validation. Similarly, Olewnik and Lewis (2005) discuss how valid decision support methods should be logical, use reliable information, and not bias the designer, and use this as a basis for evaluating the evidence used to support method claims. Thus, while Daalhuizen and Cash's (2021) framework complements prior design research on method development, method content and evaluation (Olewnik & Lewis, 2005; Tromp & Hekkert, 2016; Vermaas, 2016), this must be contextualised with respect to its supporting research claims.

1.3 Evaluating research claims

Fundamental to evaluating research claims, as well as communicating their significance to practitioners, is understanding how they are supported by evidence derived from the research and development process (Flay et al., 2005).

However, while the importance of assessing evidence has been highlighted by several authors (Cash, 2018; Cross, 2012; Reich, 2010) there are few current standards for evidence evaluation in design research (Cash, 2018). This further contrasts with fields developing interventions akin to design methods, including software engineering (Kitchenham et al., 2004), management (Bansal & Corley, 2011), education (Levin & O'Donnell, 1999), and all areas of health (Glasgow & Emmons, 2007; Grimes & Schulz, 2002). Despite this deficit in design research, there is a remarkable degree of commonality across fields, especially in relation to evaluating interventions influenced by interaction with a method user (as with design methods; Levin & O'Donnell, 1999). Hence, it is possible to build on these standards in developing an understanding of claims supporting design methods.

In this context, two works are particularly relevant. First, Gottfredson et al. (2015) describe standards for evaluating the research and development process, from theory to 'effective' recommendations for practice. Second, Grimes and Schulz (2002) describe standards for evaluating evidence for causal research claims (e.g. relating an intervention to increased performance). Together these highlight the need to not only evaluate the intervention itself (as in Section 1.2), but also: i) the rationale for the specific intervention based on prior research and real-world practice; ii) the development of the intervention, including the key decisions taken and how these are grounded in research; and iii) the nature of the claims being made, their extent and scope of relevance (e.g. contextual bounding), and the evidence that is being used to support these. Coupled with the properties of 'good' method content outlined in Section 1.2, these provide a basis for developing an assessment framework for design methods.

2 *Building a systematic assessment framework for design methods*

Bringing together the literature from Section 1 it is possible to build a basic lens for understanding the assessment of design methods from conception to application. This forms a logical chain of evidence from initial insights regarding the need, through development, to claimed impact, anchored around the method content (Daalhuizen & Cash, 2021). We conceptualise this chain with respect to five major links as illustrated in Figure 1: method motivation (Section 2.1), method nature (Section 2.2), method development (Section 2.3), method content (Section 2.4), and method claims (Section 2.5). Weaknesses in any of these links diminishes the strength of the whole chain, and ultimately the design method itself. This forms the conceptual basis for our assessment framework, which we operationalise in Table 1 (and detail in Sections 2.1–2.5).

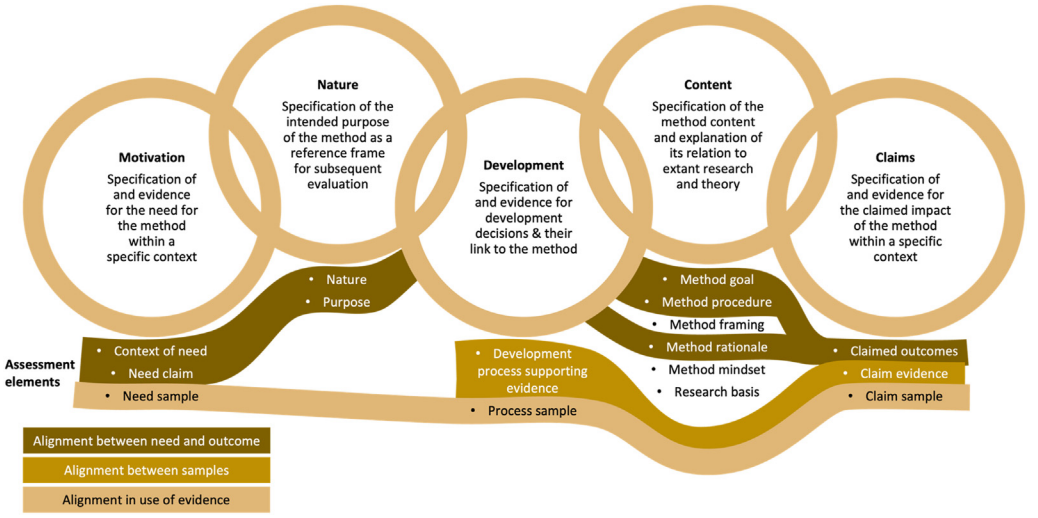


Figure 1 The chain of evidence supporting the proposal of a design method with assessment elements listed below each link, and alignments between elements indicated by the coloured threads

Important to note here, is that no assessment framework can exhaustively capture every aspect of method research and the associated ecosystem of method producers, promoters, users, and wider 'market' (for lack of a better term) of methods available in the wild. Therefore, we explicitly limit our focus to evaluating the robustness of the research supporting the proposal of a method and hence an understanding of its likely reliability in fulfilling its claims. This focus on the endogenous elements of method assessment implies two important exclusions relevant to the broader discussion of design methods.

First, the 'good' artefact property of desirability goes far beyond the scope of endogenous assessment criteria. Just as with products, desirability is also linked to users' perceptions of value or novelty, marketing, branding, and many other exogenous factors beyond the content of a method and its supporting evidence (many of which are little acknowledged in academic reporting, including in our sample of papers). Second, and following the same logic, we also consider method adoption, popularity, or other similar assessment criteria beyond the scope of endogenous assessment. Again, such outcomes are highly dependent on exogenous factors including marketing as well as the producers and promoters of a method and can often completely overshadow endogenous factors in user decision making (just as with products and their marketing). Hence, we exclude exogenous factors related to the method ecosystem or adoption unless there are specific claims being made about this in the reporting of the method itself.

With these exogenous exclusions in mind, Figure 1 illustrates the chain of evidence together with the endogenous assessment elements associated with each

link. Further, it highlights alignments between these elements that can strengthen or weaken the overall chain. For example, alignment between the need sample and the claim sample ensures that claims are being recontextualised in the same general group from which the motivation for the method emerged. For details of each assessment element see Table 1.

Table 1 A proposed systematic assessment framework for design methods, from need to impact (see Figure 1), building on method content theory (Daalhuizen & Cash, 2021)

Element	*Operationalisation*	*Description based on literature*
Method motivation		
Context of need	Practice Education Research Not Reported	*Multiple choice*: Where does the need for the method originate? This includes the major sources of insight in design research following Cash (2018, 2020)
Need claims	Efficacy Effectiveness Dissemination Not reported	*Multiple choice*: What is the extent of the need claim? This includes efficacy (validity), effectiveness (scale of impact), and dissemination (scope of applicability) following Daalhuizen and Cash (2021) and Gottfredson et al. (2015)
Need sample	Students Practitioners Both Not reported	*Selection*: From what sample is evidence for the need derived? This includes the major sub-samples typically recognised in design research following Atman et al. (2007)
Method nature		
Nature	Principle Approach Strict method Tool Template	*Multiple choice*: What is the nature of the method? This ranges from generic principles that do not necessarily say what to do to templates describing the structure of a specific output (Lewrick et al., 2020; van Boeijen et al., 2020)
Purpose	Support practice Support education Support both Not reported	*Selection*: What is the general purpose of the method? This includes the major impact areas typically recognised in design research (Meyer & Norman, 2020; Zielhuis et al., 2022)
Method development		
Development process supporting evidence	Expert practitioner opinion Research through design Single case Multi-case Experiment RCT Not reported	*Selection*: What approach has been used to generate evidence to inform method development? This ranges from expert opinion to Randomised Controlled Trails (RCTs)) adapted from the generic evidence hierarchy by Grimes and Schulz (2002). While RCTs are not expected for development they are included for completeness in terms of levels of evidence.
Process sample	Students Practitioners Both	*Selection*: What sample has been used to inform method development e.g. via prototyping with students? Again, following Atman et al. (2007)
Method content		
Method goal	Specific goal(s) Prioritisation or hierarchy Not reported	*Selection*: What is the goal(s) of the method? This includes the specific goal the method is to contribute to, its scope, and degree of flexibility building on Lee, Bobko, Earley, and Locke (1991)
Method procedure	Specific steps How to complete steps Purpose of steps	*Selection*: What are the steps in the method? This includes the structural knowledge about a specific way to reach a goal building on Roozenburg and Eekels (1995)

(*continued on next page*)

Table 1 (*continued*)

Element	*Operationalisation*	*Description based on literature*
Method framing: i) Context	Description … with explicit boundaries Not reported	*Selection*: In what context can the method be applied? This includes several dimensions: i) organisational and environmental context of use, ii) task or type of action involved, and iii) positioning or relation to the wider design process, building on various descriptions of method staging (Andreasen, Thorp Hansen, & Cash, 2015; Badke-Schaub, Daalhuizen, & Roozenburg, 2011; Gericke et al., 2017)
Method framing: ii) Task	Description … with explicit boundaries Not reported	
Method framing: iii) Positioning	Description … with explicit boundaries Not reported	
Method framing: Prerequisites	Required competences Required materials Required resources Required knowledge	*Multiple choice*: What is needed to stage the method? This includes the prerequisites or resources necessary for successful use (Andreasen et al., 2015; Badke-Schaub et al., 2011; Gericke et al., 2017)
Method rationale	Goal(s) success criteria Goal(s) end conditions Rationale for above criteria Support to reflect on progress/goal completion Not reported	*Multiple choice*: What is the performance-goal relationship for the method? This includes how to evaluate if you have succeeded, when to end, and how to reflect on progress with respect to the method's goal(s) in its specific domain and context of use. Again, building on Lee et al. (1991)
Method mindset	Values and beliefs Working principles Not reported	*Multiple choice*: What is the required method mindset? This includes descriptions of underlying values and beliefs and basic working principles building on Andreasen (2003)
Research basis	Logical speculation or inductive reasoning References to past findings Existing conceptual arguments Existing models, diagrams, or figures Existing theory	*Selection*: What is the basis for the components incorporated in the method? This includes the logic behind translation from grounding research to method content, and the associated implicit or explicit expectations for its performance, building on Cash (2020) and Colquitt and Zapata-Phelan (2007)
Method claims		
Claimed outcomes	Efficacy Effectiveness Dissemination	*Multiple choice*: What is the extent of the outcome claim? Again, following Daalhuizen and Cash (2021) and Gottfredson et al. (2015)
Claim evidence *	Expert practitioner opinion Research through design Single case Multi-case Experiment RCT Not reported	*Selection*: What approach has been used to generate evidence to inform method claims? Again, following Grimes and Schulz (2002) * Iterated for each claim
Claim sample *	Students Practitioners Both	*Selection*: From what sample is evidence for the claim derived? Again, following Atman et al. (2007) * Iterated for each claim

Evaluating design methods

2.1 Method motivation: why is the method needed?

Evaluation of the validity and impact of a method logically builds on a foundational understanding of the need(s) that it addresses. Without this, success claims become detached from context. Thus, specification and evidence for need(s) forms the first link in our chain (Figure 1).

Three elements are required to understand need. First, where does the need for the method originate: practice, education, and/or research (following the major sources of insight typically found in design research (Cash, 2018, 2020))? This provides an initial context essential to understanding the later validity of causal claims (Wacker, 2008).

Second, what is the claimed extent of the need: efficacy (i.e., the validity of the need), effectiveness (i.e., the scale of the need), and/or dissemination (i.e., the generalisability of the need) (adapted from Gottfredson et al. (2015) and Daalhuizen and Cash (2021))? This provides the basis for understanding the extent of causal claims (Wacker, 2008).

Third, from what sample is evidence for the need derived: students and/or practitioners (following the major sub-samples typically found in design research (Atman et al., 2007; Cash, Isaksson, Maier, & Summers, 2022))? This again informs understanding of the context of the need and provides a first point of alignment across the chain. For example, if a need is based in practice but evidence is drawn from student samples during method development or testing then potential weaknesses in the chain could emerge (Figure 1). Together, these three elements serve to evaluate the specificity and evidence for the method need. While these do imply a gap in the wider method market (i.e., the need is at least relevant to the sample reported in the motivation) they primarily deal with the claims made in the research itself and thus generalisation to the whole method ecosystem and systematic positioning of originality and market research are not the focus here.

2.2 Method nature: what type of method is it?

Evaluation of method performance requires a reference frame for its intended purpose. For example, some methods are intended to provide general, abstract guidelines that are contextually adapted and applied by the designer (e.g., some versions of design thinking (Brown, 2008)), while others intended to provide a more constrained framework that will lead to repeatable outcomes across designers (e.g., many versions of the design structure matrix (Pektaş & Pultar, 2006)). Such differences can substantially impact how method claims should be evaluated in terms of their generalisability (Wacker, 2008). Thus, specification of method type forms the second link in our chain (Figure 1).

Two elements are required to understand type. First, what is the nature of the method? Despite the widespread use of the term 'method', this has substantially different meanings across the design research literature. Cutting across this variation, the key criteria affecting evaluation of causal claims is generalisability i.e., their sensitivity to method user and context (Wacker, 2008). Therefore, we build upon Daalhuizen et al.'s (2019) logic for method categorisation based on the scale of the cognitive support offered. Here, we contend that the greater the support the less the (particularities of the) designer and context will impact method outcomes. Hence, we identify five distinct types of methods commonly found in the literature (Lewrick, Link, & Leifer, 2020; van Boeijen et al., 2020): i) values and principles that guide overall work (e.g., principles of sustainable design or user-centered design), ii) approaches that structure a whole process (e.g. the Vision in Design method; van Dijk & Hekkert (2011), or Product Development Process; Ulrich and Eppinger (2008), iii) strict methods that structure sequences of tasks (e.g. mind mapping or Quality Function Deployment (QFD)), iv) tools that support tasks (e.g. PrEmo; Desmet (2018), or CAD software), and v) templates that structure the output of actions (e.g. business model canvas; Osterwalder, Pigneur, Oliveira, and Ferreira (2011) or the eco-design strategy wheel; van Boeijen et al. (2020)). This provides the basis for understanding the generality, abstraction, and contextual bounding of a method with respect to design work (Wacker, 2008).

Second, what is the general purpose of the method (as opposed to the specific goal, see Section 2.4): supporting practice and/or education (following the major impact areas in design research (Meyer & Norman, 2020; Zielhuis, SleeswijkVisser, Andriessen, & Stappers, 2022))? This provides a frame of reference for testing of causal claims (Wacker, 2008), and a second point of alignment across the chain. For example, a need may emerge from practice, but a method may be directed towards education, and then tested in practice and/or education. As such, clarification of alignment between these contexts and samples is needed to avoid weakness in the chain. Together, these two elements serve to evaluate the specific method type, providing a reference frame for the rest of the chain (Figure 1).

2.3 Method development: how is the method designed?

Evaluation of developmental robustness builds on an understanding of the evidence used to support key choices during the development of a method (Gottfredson et al., 2015; Vermaas, 2016). Thus, specification and evidence for development decisions forms the third link in our chain (Figure 1).

Two elements are required to understand development. First, what approach has been used to generate evidence to inform method development decisions? As this typically deals with the use of evidence to support causal claims, we

follow a widely accepted hierarchy of evidence in this context (Grimes & Schulz, 2002). This provides the basis for understanding the support for key decisions that might impact method content or claims (Gottfredson et al., 2015).

Second, what sample has been used to inform method development: students and/or practitioners (again following Atman et al. (2007))? This again informs understanding of the context of the development and provides a third point of alignment across the chain. While there are many other details required for replicability, these two elements serve to evaluate specificity and evidence for method development.

2.4 Method content: how does the method work?

Evaluation of method efficacy is underpinned by understanding of method content (Daalhuizen & Cash, 2021; Seepersad et al., 2006), as well as how this has been translated from basic research (Gottfredson et al., 2015). This is a key feature of design methods as an output of design *research* — as opposed to design *practice*. Thus, specification of method content and its basis in design research insights forms the fourth link in our chain (Figure 1).

Two elements are required to understand content. First, what does the method comprise? Here, we build directly on the model of method content proposed by Daalhuizen and Cash (2021), which includes the five main components listed below. This provides the basis for understanding the key concepts and relationships that define how a method functions, as a foundation for explanation, prediction, and causal claims (Wacker, 2008). Notably, these reflect the reported content of the method itself and thus provide several points of alignment with the wider supporting research across the chain, as highlighted and exemplified with the Contextmapping method (listed in van Boeijen et al., 2020) below.

- **Method Goal**: the explicit description of the goals and their prioritization a method aims to help achieve through method use (related to Sections 2.1 and 2.2). For example, the contextmapping method describes its goal as helping designers create solutions that fit people's needs.
- **Method Procedure**: the explicit description of the structural activities involved in the proper use of the method and their relative chronological and logical ordering (related to Section 2.2). For example, the contextmapping method describes its procedure in three steps for collecting and communicating user insights: preparation and sensitizing, generative assignments, and analysis and ideation.
- **Method Framing**: the explicit description of the scope of use setting (including context, task, and positioning) and its implications and prerequisites for method use. For example, the contextmapping method described

that it is to be preferably used in the pre-concept stage and as part of co-design or co-creation processes.

- **Method Rationale**: the explicit description of the performance-goal relationship and motivations underlying the goals of the method (related to Section 2.1). For example, the contextmapping method describes that it helps uncover users' latent knowledge that helps to empathize with intended users.
- **Method Mindset**: the explicit description of the set of values, principles, underlying beliefs, and logic that inform method use. For example, the contextmapping method describes that a core value is that the people designers design for are the expert of their own experiences and that designers should respect this.

Second, what is the basis for the components incorporated in the method? This follows the idea that methods embody research-based understanding of design work (Daalhuizen, 2014; Gray, 2022; Stappers & Sanders, 2005). Hence, the translation from research insight to a functional method constitutes some form of test (even if only implicitly). Therefore, it is possible to follow the acknowledged schema for assessing such translation provided by Colquitt and Zapata-Phelan (2007), and recently applied in design research by Cash (2020). This provides the basis for understanding the robustness of the research logic underpinning method content. Together, these two elements serve to evaluate the specificity and research grounding of the method content.

2.5 Method claims: what is the Method's impact?

Evaluation of claimed impact builds on the robustness of evidence for both the basic functionality of a method, as well as its ability to address its intended need/purpose in context (Daalhuizen & Cash, 2021; Gottfredson et al., 2015). These form a logical specification of and counter point to the claimed needs (Section 2.1). Thus, specification and evidence for claims forms the final link in our chain (Figure 1).

Three elements are required to understand claims. First, what is the extent of the outcome claim(s): efficacy (i.e., the functional predictability and validity of the method), effectiveness (i.e., the usability and scale of its impact in context), and/or dissemination (i.e. the desirability and general uptake) (adapted from Gottfredson et al. (2015) and Daalhuizen and Cash (2021))? This provides the basis for understanding the specific nature of the claims being made, providing a reference frame for subsequent testing methods and evidence (Gottfredson et al., 2015).

Second, what approach has been used to generate evidence to inform each method claim? Due to the typical focus on causal claims (i.e., the method improves …) we again follow the hierarchy of evidence provided by Grimes and

Schulz (2002). This provides the basis for understanding the robustness of the support for each claim (Gottfredson et al., 2015).

Third, from what sample is evidence for the claim derived: students and/or practitioners (again following Atman et al. (2007))? This again informs understanding of the context of the claim and provides a final point of alignment across the chain. Together, these three elements serve to evaluate the specificity and evidence for the method claims. As with motivation, while these do imply a contribution to the wider method market (i.e., the method is at least relevant to the degree reported in the claim) they primarily deal with the claims made in the research itself and thus generalisation to the whole method ecosystem and evaluation of potential uptake or market buy-in are beyond the scope of these elements.

Bringing together the criteria discussed in this section, we can propose an assessment framework for design methods, which systematically evaluates each link in the chain of evidence (Figure 1). Table 1 provides an overview of this assessment framework, detailing each link, their major elements, operationalisation, and description based on literature.

3 *Method*

To provide a foundation for field development and ground current debate, we conducted a systematic review of recent research proposing a design method, following the updated PRISMA guidelines (Page et al., 2021). Here, our intention was not to characterise the whole history of methods research, but rather establish the current state of design methods research as a basis for moving forward as a field. Given this aim the review comprised three main phases: (1) identification of design research proposing a method via search of major design journals, (2) screening the records and eligibility assessment according to pre-defined inclusion and exclusion criteria, and (3) inclusion of eligible records in the final sample for further review and analysis. The PRISMA flow diagram summarising our process is presented in Figure 2.

While the PRISMA approach was developed primarily as a guide for meta-analyses and systematic reviews of health interventions, such as clinical trials, the 27-item checklist and flow diagram are applicable for other types of systematic reviews (Page et al., 2021). This provides a standard approach, which has been applied in several recent reviews of design research (e.g., see Hay et al. (2017)). In this paper, we focus on current research proposing new methods bounded by two main criteria. First, we only consider methods published in recognised design research journals for two main reasons: i) these represent the 'best' of the field with more stringent peer review and acceptance criteria than either conferences or books (especially with respect to the provision of

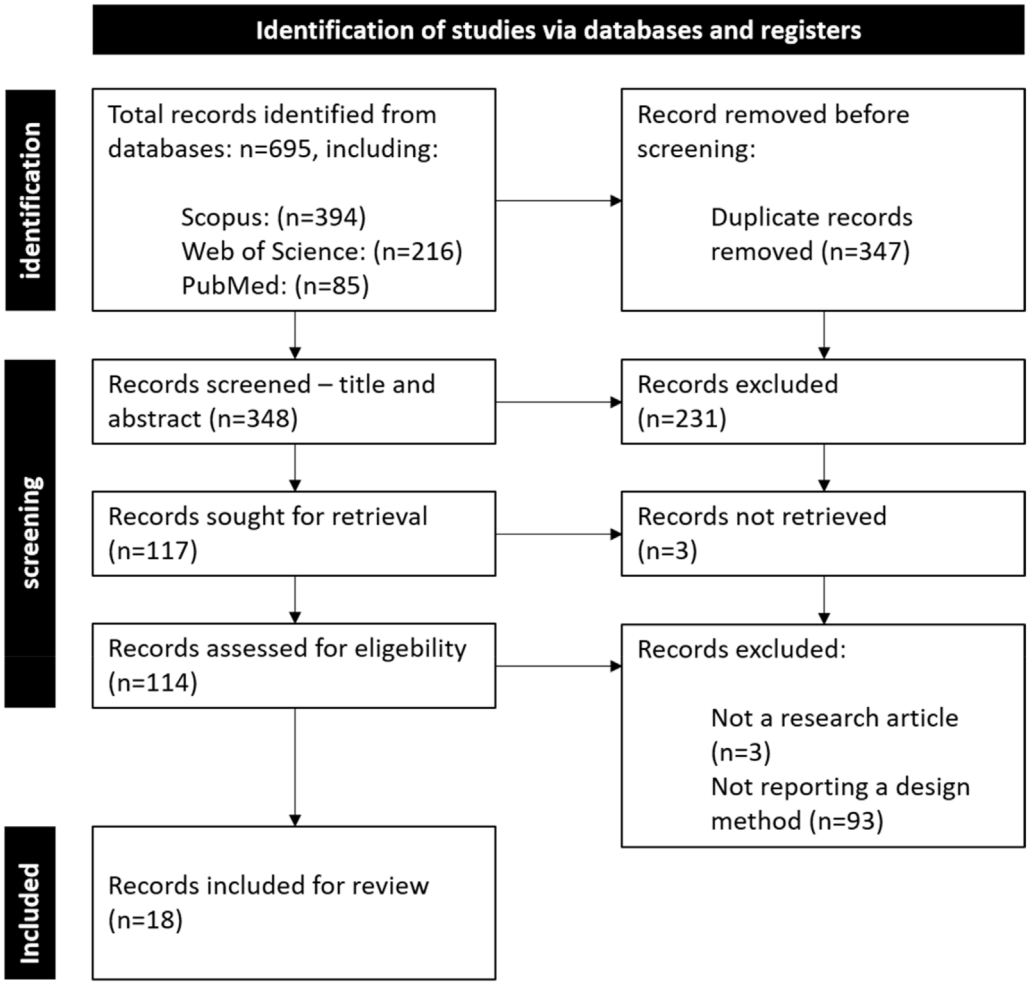


Figure 2 The PRISMA flow diagram specific to the literature review presented in this paper, based on Page et al. (2021)

supporting evidence) and ii) these reflect a bounding of the design research domain recognised within the field itself (Cash, 2018; Gemser, de Bont, Hekkert, & Friedman, 2012). Notably, we exclude textbooks and method repositories (where most method descriptions are found) because these primarily focus on dissemination and hence do not typically report underlying research and evidence (Figure 1). Second, we only considered methods published in 2020, due to these representing the most current still unaffected by the potential limitations of the COVID-19 pandemic (e.g., limiting data access).

3.1 Article selection process

The search was conducted based on the journal list developed by Gemser et al. (2012), which is widely recognised in the community as identifying top design

research journals. Again, as our focus is on establishing the current best case, we deliberately focus on a limited set of journals recognised as leaders in the field. Due to this focus, we were able to review every paper published in 2020 across all six journals summarised in Table 2. This resulted in a total of 695 records identified.

Following this, records were screened by a trained research assistant based on the title and abstract. Here, 231 records were excluded as immediately out of scope (e.g., reporting general design research and not including the proposal of a new method), not retrievable (n = 3) or not being a research article (n = 8). A second round of screening examined the remaining 106 records, which were screened based on the full text. Again, records were excluded if they did not propose a specific method that fell within the basic definition summarised in Table 2 (e.g., reporting a conceptual framework that could be applied by a designer but was not claimed as a specific method or 'methods' that were purely computational and thus eliminated interaction between method and designer). This resulted in a final selection of 18 records. The full list of inclusion and exclusion criteria applied can be found in Table 2.

3.2 Article analysis process

To analyse the 18 records selected for the review, we used a multi-part coding process. First, we coded the records for background information like application area, research context, etc. Second, we analysed selected records according to the elements of our proposed assessment framework (Table 1). The first round of analysis was performed by a trained research assistant. Training happened iteratively, starting with an explanation of the assessment framework after which the research assistant analysed three records, noting down any doubts that arose during analysis. After this, one of the authors and the research assistant discussed the analysed records and resolved all instances of doubt. After training, the research assistant analysed all records, again marking any instances of doubt. The second round of analysis was performed

Table 2 Review inclusion and exclusion criteria

Criteria	*Explanation*
Quality	Publication in top general design research outlets as recognised by the design research community (Gemser et al., 2012): *Design Studies, Design Issues, Journal of Engineering Design, International Journal of Design, The Design Journal, Journal of Design Research*
Publication date (2020)	Most recent publication date unambiguously unaffected by potential data access limitations introduced by the COVID-19 pandemic
Proposal of a method	Proposal of a contribution corresponding to the definition: '*a formalised representation of a design activity, which functions as a mental tool to support designers in achieving a goal, in relation to the circumstances and resources available*' (Daalhuizen et al., 2019).

by one of the authors during which all instances marked in the previous round were analysed and resolved.

4 Results

In this section we first provide an overview of the reviewed records before detailing the outcomes of our analysis.

4.1 Descriptive summary

Of the initial set of 695 records 18 were found to contain a method proposal. These were distributed across the surveyed journals: *Design Studies* (3), *Design Issues* (0), *Journal of Engineering Design* (8), *International Journal of Design* (3), *The Design Journal* (4), *Journal of Design Research* (0). Further, they targeted a wide range of design practices, including, ideation (3), development (10), modelling (2), and evaluation (3). Similarly, they covered an array of contexts from manufacturing and engineering design to participatory and co-design applied to everything from products to food or shared spaces. An overview of all the reviewed records is provided in Appendix Table A. Given the spread of the reviewed records and distribution across outlets we conclude that this sample provides a credible foundation for evaluating current research across the field.

4.2 Assessing current design method research

Analysing the selected records in detail revealed the proposal of 18 methods, with some methods containing multiple elements (e.g., a method and a tool) resulting in a total of 25 methodological elements. Upon inspection, the evidence presented within a record typically addressed all included elements, as such we take the 18 records as our primary level of analysis. Evaluating these provided important findings across the chain of evidence (Figure 1).

4.2.1 Method motivation

In this link of the chain (Figure 1) we evaluated the context of the need, the need claim, and the supporting sample. The overall results for each of these evaluation elements is summarised in Figure 3a–c. Taken together, the major insight in this link was that while 14 records claimed a valid or specific scale of need (Figure 3b) emerging from a specific context (practice, research, or both) (Figure 3a), only 3 reported the sample supporting these claims (Figure 3c). Further, there was a general lack of maturity in need claims with only 3 records claiming anything beyond validity, neglecting the scale and applicability of the need. Together these leave significant gaps in the initial link in most records and point to the potential for much clearer articulation of the scope of need and the evidence supporting this.

4.2.2 Method nature

In this link we evaluated the nature and purpose of the method type, as summarised in Figures 3d and 3e. The reviewed records presented a total of 25 methodological artefacts, which included 2 principles, 7 approaches, 13 strict methods, 1 tool, and 2 templates (Figure 3d). Notably 7 records proposed two methodological artefacts as part of a wider proposal (e.g. a strict method and a supporting tool). While all artefacts could be consistently characterised based on Table 1 the nomenclature varied significantly across records. Further, 2 records did not report the intended target of the proposal. Thus, while most records were clear in their articulation of a specific purpose these results reveal a critical need for alignment in terminology across the design literature.

4.2.3 Method development

In this link we evaluated the development process, supporting evidence, and its associated sample, as summarised in Figures 3f and 3g. While we would not expect the full scale to be utilised in the context of method development (e.g., RCTs are rarely—if ever in the design context—used for developmental purposes) our results reveal a critical deficit in the chain of evidence. Specifically, 12 records did not report on the evidence used as the basis for their development process (Figure 3f). For those records that did report on evidence used, one reported a research through design study, two a single case study, two a multi-case study, and one an experiment. This again reveals a critical lack of maturity in reporting practices with regards to the development process. Thus, despite the acknowledged importance of understanding the basis for developmental decision making in traditional design, this is currently not reported in the context of method development.

4.2.4 Method content

In this link we evaluated the various aspects of method content as summarised in Figure 3h–p. Here, while all records provided some explanation of their research basis there was significant inconsistency and incompleteness with respect to the reporting of the method content itself. Specifically, no single record described all aspects of method content, with several elements notably neglected. For example, 10 records did not provide goal rationale (Figure 3n) and 7 did not provide specific goals (Figure 3h). Further, 13 records did not describe the underlying method mindset (Figure 3o) and 7 did not even report the prerequisites needed to complete the proposed method (Figure 3m). Again, this reveals a lack of maturity in reporting and highlights the need for further examination of what is needed to understand and use method proposals to increase consistency in this critical link. It is surprising that even for perhaps the most basic element of methodological proposals – their core content – reporting is so often lacking.

4.2.5 Method claims and evidence

In this link we evaluated the claimed outcomes, their supporting evidence, and associated samples, as summarised in Figure 3q–s. These results revealed two important insights. First, while 13 records reported on the efficacy of their proposals and 5 reported effectiveness, none reported on dissemination (Figure 3q). Second, of the 13 records making efficacy claims, 12 provided some form of evidence (Figure 3r) but only 6 explicitly defined the sample from which this was derived (Figure 3s). However, it is notable that 2 records reported claims on both efficacy and effectiveness, including the evidence and samples used. Hence, while examples of more complete reporting can be found in the reviewed literature there is a critical lack of clarity in reporting the supporting evidence and sample for method claims.

4.3 Summarising current design method research

Figure 3 provides a summary of the results from across the chain of evidence; however, one critical insight should be highlighted based on the detailed results reported in this section. When the reviewed literature is taken as a whole, there are major deficits in reporting in almost every link in the chain of evidence (Figure 3). Yet, for all links at least one record was identified that clearly reported the assessment elements (Table 1), with individual records often providing complete and clear reporting of one or more specific links (however, no record covered all links). Hence, within the reviewed records there is overall recognition and reporting of all links in the chain of evidence, and evidence for the emergence of possible best practices in each specific link. Thus, while there is an evident lack of consistency and maturity across the reviewed records there is also evidence for the possibility of improvement and the overall potential value to be derived from structuring the reporting of method proposals.

5 Discussion

We set out to better understand how design methods and their associated claims can be assessed and subsequently ground current debate on method credibility in design research. In answer to this, our results revealed that even though some papers are quite complete in reporting specific links in the chain of evidence (Figure 1), overall reporting is quite incomplete and never complete (in a single paper) across all links in the chain. Taking two of the most complete works as good examples, Stylidis, Wickman, and Söderberg (2020) propose a method for ranking attributes of perceived product quality while Y. Lee, Breuer, and Schifferstein (2020) propose a set of food design tools. Both works report to a moderate/high degree of maturity on the method itself, its development, evidence supporting its efficacy and the need, yet do not report the sample for the need nor on the dissemination of the method. Yet these examples are exceptions with most cases offering relatively incomplete reporting, reflected in the high numbers of 'not reported' results (red bars)

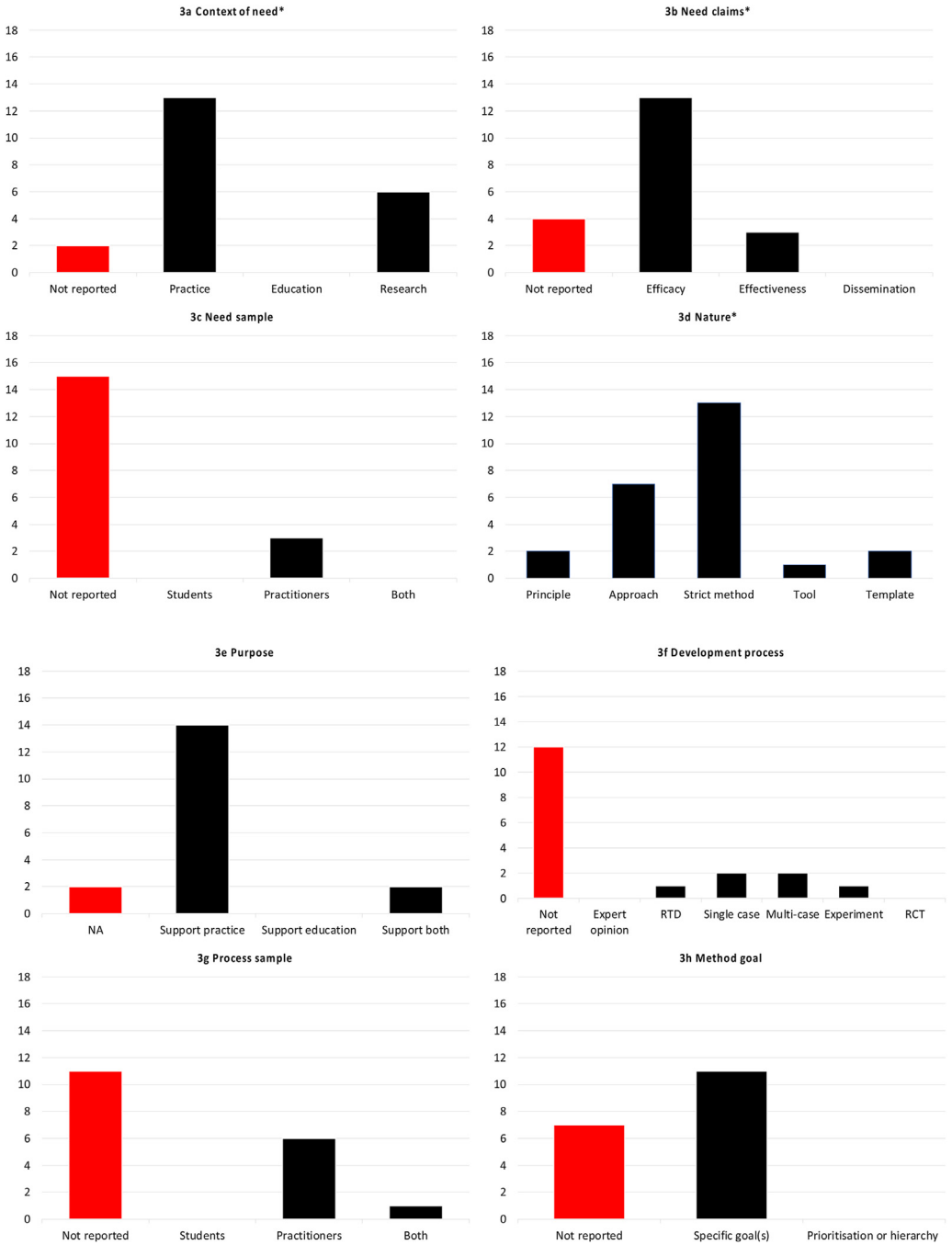


Figure 3 (a–s) Summary of results across the chain of evidence. Note: all y-axis reflect number of records and ∗ denotes multiple choice assessment

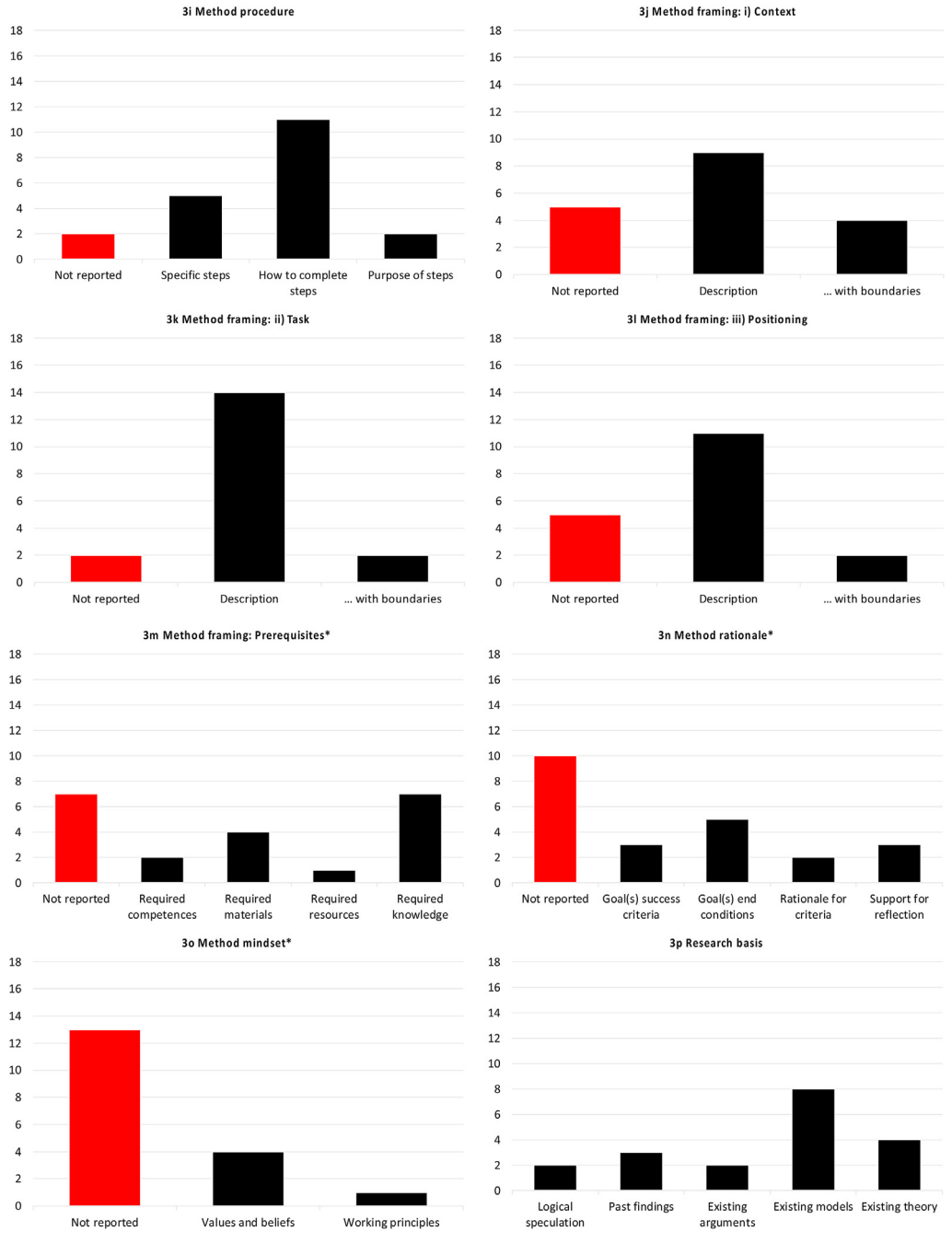


Figure 3 (continued)

Evaluating design methods

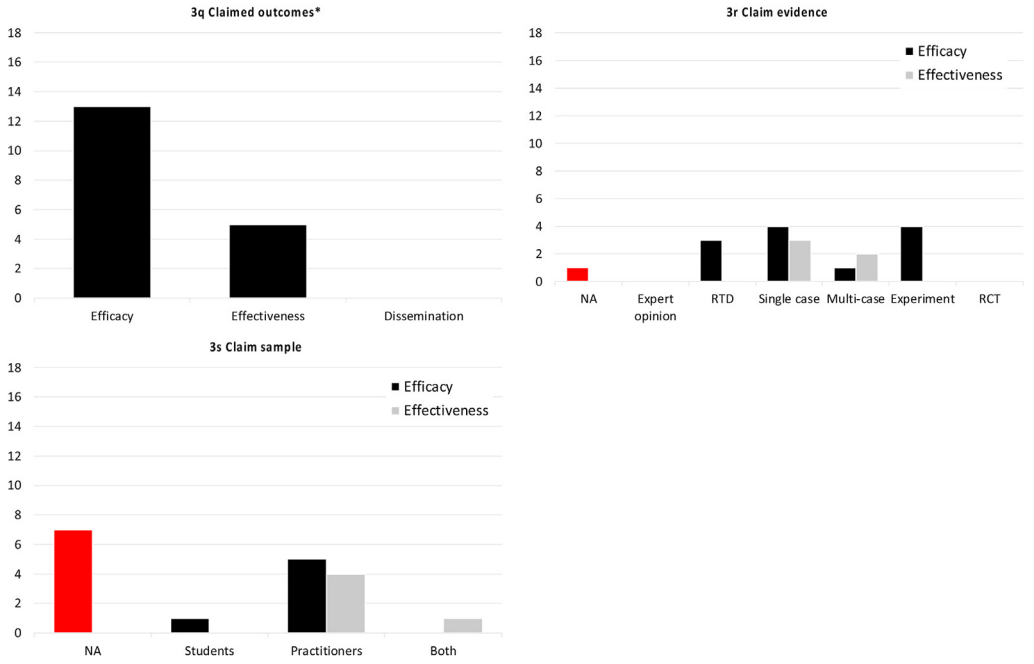


Figure 3 (continued)

in Figure 3a–s. Overall, most of the reviewed papers do not provide the information and/or evidence crucial to understanding the content and potential impact of the method proposals and their development. In short, the chain of evidence is often broken, with key links missing in most papers. Further, these missing links are also typically not recognised in the limitations of these papers.

Our findings reveal three major shortcomings highlighting the importance and potential value of the proposed assessment framework for supporting self-reflection, method assessment, and research evaluation: i) all links were found when looking across the whole set of reviewed papers, but never complete in individual papers; ii) for every link, potential 'best' practices were found, yet most papers did not report or only superficially addressed each link; and iii) where links were missing or otherwise incompletely addressed such deficits were typically not reported in the limitations. However, it is important to recognise that these results are perhaps not surprising given the lack of standards for method reporting in design research. Therefore, the positive exceptions highlighted (in completing most of the chain, in fully addressing specific links, and in reflecting on relevant limitations) should be applauded rather than the less complete examples criticized (see Table A for examples). Thus, our results demonstrate the need for clear and complete standards of

evidence in method research, as well as a more general need to take concerted action towards developing the maturity of the field.

5.1 Limitations

Before discussing implications, it is important to consider the main limitations of this work. First, the extent of the conceptual framework. The proposed framework reflects a first step in assessing the whole evidence chain from need to claimed impact; but does not deal with many other secondary factors that can impact methods, such as embodiment, facilitation, staging, context of deployment etc. and explicitly excludes interactions with exogenous factors that can impact users' perceptions of novelty, value, or wider uptake, such as the method producer/promotor, its branding/marketing, and other positioning and promotion within the method market and wider ecosystem. However, given the lack of maturity in this area this still provides an important contribution as evidenced by the findings reported in Figure 3, and a foundation on which market facing and other exogenous assessment elements could be built.

Second, the extent of the review. The review reflects current research in core design research journals. This serves its purpose as we aimed to assess the 'best' current practices, which are fostered by rigorous peer review and editorial oversight provided by the journal outlets. However, the reality is that most methods are only proposed in books or conferences, which vary wildly in peer review, editorial oversight, and ultimately quality and content. Thus, while our work provides an important foundation for moving the field forward, we need more work to really understand to what degree the 'best' represents the rest as well as how this might have changed over time, and how standards can be effectively deployed in the face of such varied outlets.

5.2 Implications and future research

Our work has several implications for method research, development, education, and practice. First, in terms of method research and theory, our evaluation, particularly of method content, highlights the potential for cross-cutting theory to scaffold maturation of the research field. Specifically, our findings show how increasing our understanding of the phenomenon underlying design methods and their use can directly contribute to improvements in development and reporting. While some elements were more consistently reported (such as aspects of method content), many elements are still treated implicitly or not reported at all (Figure 3). This inconsistency in reporting—even in leading design research journals—also emphasises the need for more systematic and consistent peer review. In this context, our work could provide a general guide for reviewers confronted with manuscripts involving method development, and points to the need for further investigation of quality criteria in this context, linking to the call for action in the recent Design Research Notes

initiative (Cash, Isaksson, et al., 2022). Hence, key future research questions include how method content interacts with the designer (as user of a method) in context, how to understand the adaption necessary in translating and embodying method content in practice, how the staging of methods might impact outcomes, and how method research can be consistently reviewed.

Second, in terms of wider method research, our work highlights the need for focused study of exogeneous factors affecting perceptions of methods and their uptake. Specifically, there is a need to better understand how users perceive methods (including evaluation of their novelty, relevance, and value in context) and how this relates to method content, its reporting, dissemination, and marketing. Further, there is a need to examine what other exogeneous factors impact method adoption and how these relate to the elements in Figure 1. For example, little is known about the impact of the method producer, promotor, or branding and marketing efforts on user uptake. Further, designers don't typically have time to try out multiple methods for a given project/task, and hence need support in identifying and applying the most relevant methods to their context, yet research on this meta level is, to the authors' knowledge, lacking. As such, a logical follow-up to this work is an examination of the evidence and reasons for method adoption amongst method users across contexts, including the impact of third-party actors in the method ecosystem, such as design consultancies or university marketing teams. Hence, key future research questions include how to understand exogenous factors impacting method perception and adoption and how method research can be tailored to engage with these without compromising quality.

Third, in terms of method development, there is currently no framework that establishes what good methods are and how to develop and report them. As such, our proposal for a chain of evidence (Figure 1) could form a foundation for developing good practice in this context. A second aspect of this is to acknowledge the contextual nature of method use, and to report what skills are required to properly use the method and in what contexts the method is best applied – and in what contexts we better refrain. Further, the systematic assessment framework can be used 'in reverse' as a checklist when planning, developing, and reporting methods. However, it is important to acknowledge that this is not a meta-method for method development, and that such an approach is an area for further research. Hence, key future research questions include how to understand and report on method development processes balancing generic and context specific elements and how best practices might be further developed in this area.

Finally, in terms of societal, practice, and educational impact, methods often target key challenges and include corresponding large-scale impact claims. Our findings highlight the need for a more in-depth discussion of what constitutes quality of methods and evidence in this context, and what types of research

infrastructure are needed to be able to achieve such quality standards reliably. Both policy makers and industry, for example, are more and more asking for evidence-based methods (Alonso et al., 2020; Design Council, 2020), which are proven to be actionable and effective. This is critical if design methods are to be held up alongside methods developed in related fields such as engineering or health. We need to show non-design sectors that deal with major societal challenges that our methods can—and do—make a difference and that they deliver the promises we make. It is therefore crucial that future research continues to examine how methods are adapted and applied, how implementation can be understood in and across contexts, and how we can offer compelling evidence of both methodological rigour and impact to diverse stakeholders.

6 *Conclusions*

We set out to better understand how design methods and their associated claims can be assessed and subsequently ground current debate on method credibility in design research. In doing so, we first developed a systematic assessment framework for design methods (Table 1) built on a logical chain of evidence from initial insights regarding the need to claimed impact (Figure 1). Based on this, we reviewed all papers published in 2020 in leading design research journals. Specifically, we examined whether papers that report new methods provide the information necessary to define and evaluate the proposed method and its development process, as well as support the claims associated with this.

Our results revealed that while all links in the chain of evidence are reported across the literature and best practices can be identified for each link, no individual paper either reports all links or consistently achieves best practice. While these findings might not be surprising—due to the lack of current standards of evidence in this area—they highlight the potential value of our proposed assessment framework and point to critical implications for maturing this central pillar of design research (impact).

Ultimately, we started with the question, 'what's in a claim', and can conclude with the realisation, that while the answer is complex and multifaceted (Table 1) it is also tractable in design research (with many positive examples of good—yet patchy—practice across the field). Our work thus provides a foundation for evaluating method research, demonstrates the need for clear and complete standards of evidence in this area, and highlights directions for future method research.

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.

Data availability

Data will be made available on request.

Appendix.

Table A Overview of papers included in the systematic review

Journal	*Citation*	*Title*	*Example of complete reporting in links:*
Design Studies	Hoolohan and Browne (2020)	Design thinking for practice-based intervention: Co-producing the change points toolkit to unlock (un)sustainable practices	• Motivation • Development • Content • Claims: effectiveness
	Strömberg, Pettersson, and Ju (2020)	Enacting metaphors to explore relations and interactions with automated driving systems	• Development • Claims: efficacy
	Pirinen and Tervo (2020)	What can we share? A design game for developing the shared spaces in housing	• Development
International Journal of Design	Karana, Barati, and Giaccardi (2020)	Living Artefacts: Conceptualizing Livingness as a Material Quality in Everyday Artefacts	
	Woo and Lim (2020)	Routinoscope: Collaborative Routine Reflection for Routine-Driven Do-It-Yourself Smart Homes	• Content • Claims: efficacy • Claims: effectiveness
	(Y. Lee, Breuer, & Schifferstein, 2020)	Supporting Food Design Processes: Development of Food Design Cards	• Development • Claims: efficacy • Claims: effectiveness
Journal of Engineering Design	Paparistodimou, Duffy, Whitfield, Knight, and Robb (2020)	A network science-based assessment methodology for robust modular system architectures during early conceptual design	• Claims: effectiveness
	Rigger, Vosgien, Shea, and Stankovic (2020)	A top-down method for the derivation of metrics for the assessment of design automation potential	• Claims: effectiveness
	Bashir and Ojiako (2020)	An integrated ISM-MICMAC approach for modelling and analysing dependencies among engineering parameters in the early design phase	• Motivation
	Eddy, Krishnamurty, Grosse, and Steudel (2020)	Early design stage selection of best manufacturing process	
	(D. Lee, Pan, & Fang, 2020)	Improving early stage system design under the uncertainty in reliability-wise structure	• Motivation
	Stylidis et al. (2020)	Perceived quality of products: a framework and attributes ranking method	• Development • Content • Claims: efficacy
	Stief, Dantan, Etienne, Siadat, and Burgat (2020)	Product design improvement by a new similarity-index-based approach in the context of reconfigurable assembly processes	
	Barravecchia, Mastrogiacomo, and Franceschini (2020)	The player-interface method: a structured approach to support product-service systems concept generation	

(*continued on next page*)

Table A (*continued*)

Journal	*Citation*	*Title*	*Example of complete reporting in links:*
The Design Journal	Güneş (2020)	Extracting Online Product Review Patterns and Causes: A New Aspect/Cause Based Heuristic for Designers	
	Rodgers, Mazzarella, and Conerney (2020)	Interrogating the Value of Design Research for Change	
	Celikoglu, Krippendorff, and Ogut (2020)	Inviting Ethnographic Conversations to Inspire Design: Towards a Design Research Method	• Development
	Jacobs et al. (2020)	Made-Up Rubbish: Design Fiction as a Tool for Participatory Internet of Things Research	• Claims: efficacy

Notes

1. 'Method content' is here deliberately separated from 'method use', to stress that methods can have all the proper ingredients yet fail to deliver when used improperly or under the wrong circumstances.

References

Alexander, C. (1971). The state of the art in design methods. *DMG Newsletter, 5*(3), 3–7.

Alonso, M. B., van der Bijl-Brouwer, M., Hekkert, P., Hummels, C., Kraal, J., Krul, K., et al. (2020). *Sleutelmethodologieën (KEM's) voor missiegedreven innovatie*.

Andreasen, M. M. (2003). Improving design methods usability by a mindset approach. In U. Lindemann (Ed.), *Human behaviour in design*. (209–218). Berlin Heidelberg: Springer.

Andreasen, M. M. (2011). 45 Years with design methodology. *Journal of Engineering Design, 22*(5), 293–332.

Andreasen, M. M., Thorp Hansen, C., & Cash, P. (2015). *Conceptual design: Interpretations, mindset and models*. Springer.

Araujo, C. S. (2001). *Acquisition of product development tools in industry: A theoretical contribution*. Denmark: Lyngby.

Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., & Saleem, J. (2007). Engineering design processes: A comparison of students and expert practitioners. *Journal of Engineering Education, 96*(4), 359–379.

Badke-Schaub, P., Daalhuizen, J., & Roozenburg, N. (2011). Towards a Designer-Centred Methodology: Descriptive considerations and prescriptive reflections. In B. Herbert (Ed.), *The future of design methodology* (pp. 181–197). Springer.

Bansal, P., & Corley, K. (2011). From the editors: The coming of age for qualitative research. *Academy of Management Journal, 54*(2), 233–237.

Barravecchia, F., Mastrogiacomo, L., & Franceschini, F. (2020). The player-interface method: A structured approach to support product-service systems concept generation. *Journal of Engineering Design, 31*(6), 331–348.

Bashir, H., & Ojiako, U. (2020). An integrated ISM-MICMAC approach for modelling and analysing dependencies among engineering parameters in the early design phase. *Journal of Engineering Design, 31*(8–9), 461–483.

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

van Boeijen, A., Daalhuizen, J., & Zijlstra, J. (2020). *Delft design guide: Perspectives - models - approaches - methods* (2nd ed.). BIS Publishers BV.
Brown, T. (2008). Design thinking. *Harvard Business Review, 2008*(June), 1–9.
Bucciarelli, L. L. (1994). *Designing engineers*. Cambridge, Massachusetts: MIT Press.
Cantamessa, M. (2003). An empirical perspective upon design research. *Journal of Engineering Design, 14*(1), 1–15.
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). Editorial: Design research notes. *Design Studies, 78*, 101079.
Cash, P., Isaksson, O., Maier, A. M., & Summers, J. D. (2022). Sampling in design research: Eight key considerations. *Design Studies, 78*, 101077.
Cash, P., Valles Gamundi, X., Echstrom, I., & Daalhuizen, J. (2022). Method use in behavioural design: What, how, and why? *International Journal of Design, 16*(1), 1–21.
Celikoglu, O. M., Krippendorff, K., & Ogut, S. T. (2020). Inviting ethnographic conversations to inspire design: Towards a design research method. *The Design Journal, 23*(1), 133–152.
Chulvi, V., Mulet, E., Chakrabarti, A., López-Mesa, B., & González-Cruz, C. (2012). Comparison of the degree of creativity in the design outcomes using different design methods. *Journal of Engineering Design, 23*(4), 241–269.
Colquitt, J. A., & Zapata-Phelan, C. P. (2007). Trends in theory building and theory testing: A five-decade study of the academy of management journal. *Academy of Management Journal, 50*(6), 1281–1303.
Cross, N. (2012). Editorial. *Design Studies, 33*(1), 1–3.
Daalhuizen, J. (2014). Method Usage in design: How methods function as mental tools for designers *(PhD)*. TU Delft.
Daalhuizen, J., & Cash, P. (2021). Method content theory: Towards a new understanding of methods in design. *Design Studies, 75*, 101018.
Daalhuizen, J., Person, O., & Gattol, V. (2014). A personal matter? An investigation of students' design process experiences when using a heuristic or a systematic method. *Design Studies, 35*(2), 133–159.
Daalhuizen, J., Timmer, R., van der Welie, M., & Gardien, P. (2019). An architecture of design doing: A framework for capturing the ever-evolving practice of design to drive organizational learning. *International Journal of Design, 13*(1), 37–52.
Design Council. (2020). *Making life better by design: Design Council's story of impact*.
Dorst, K. (2008). Design research: A revolution-waiting-to-happen. *Design Studies, 29*(1), 4–11.
Eddy, D. C., Krishnamurty, S., Grosse, I. R., & Steudel, M. (2020). Early design stage selection of best manufacturing process. *Journal of Engineering Design, 31*(1), 1–36.
Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., et al. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*(3), 151–175.
Frey, D. D., & Dym, C. L. (2006). Validation of design methods: Lessons from medicine. *Research in Engineering Design, 17*(1), 45–57.

Gemser, G., de Bont, C., Hekkert, P., & Friedman, K. (2012). Quality perceptions of design journals: The design scholars' perspective. *Design Studies, 33*(1), 4–23.
Gericke, K., Eckert, C., & Stacey, M. (2017). What do we need to say about a design method? *Proceedings of the International Conference on Engineering Design, ICED'17*, 101–110.
Glasgow, R. E., & Emmons, K. M. (2007). How can we increase translation of research into practice? Types of evidence needed. *Annual Review of Public Health, 28*(1), 413–433.
Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., Sandler, I. N., et al. (2015). Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: Next generation. *Prevention Science, 16*(7), 893–926.
Gray, C. (2022). Languaging design methods. *Design Studies, 78*, 101076.
Grimes, D. A., & Schulz, K. F. (2002). An overview of clinical research: The lay of the land. *The Lancet, 359*(9300), 57–61.
Güneş, S. (2020). Extracting online product review patterns and causes: A new aspect/cause based heuristic for designers. *The Design Journal, 23*(3), 375–393.
Hay, L., Duffy, A., McTeague, C., Pidgeon, L., Vuletic, T., & Grealy, M. (2017). A systematic review of protocol studies on conceptual design cognition: Design as search and exploration. *Design Science, 3*, e10.
Hoolohan, C., & Browne, A. L. (2020). Design thinking for practice-based intervention: Co-Producing the change points toolkit to unlock (un)sustainable practices. *Design Studies, 67*, 102–132.
Jacobs, N., Markovic, M., Cottrill, C. D., Edwards, P., Corsar, D., & Salt, K. (2020). Made-up rubbish: Design fiction as a tool for participatory internet of things research. *The Design Journal, 23*(3), 419–440.
Jones, J. C. (1977). How my thoughts about design methods have changed during the years. *Design Methods and Theories: Journal of DMG and DRS, 11*(1).
Jones, J. C. (1992). *Design methods*. John Wiley and Sons.
Karana, E., Barati, B., & Giaccardi, E. (2020). Living artefacts: Conceptualizing livingness as a material quality in everyday artefacts. *International Journal of Design, 14*(3), 37–53.
Kitchenham, B. A., Dyba, T., & Jorgensen, M. (2004). Evidence-based software engineering. In *26th international conference on software engineering* (pp. 273–281).
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.
Kumar, V. (2013). *101 design methods - a structured approach for driving innovation in your organization*. New Jersey: John Wiley & Sons.
Kunrath, K., Cash, P., & Kleinsmann, M. (2020). Designers' professional identity: Personal attributes and design skills. *Journal of Engineering Design, 31*(6), 297–330.
Lee, C., Bobko, P., Earley, C. P., & Locke, E. A. (1991). An empirical analysis of a goal setting questionnaire. *Journal of Organizational Behavior, 12*(6), 467–482.
Lee, Y., Breuer, C., & Schifferstein, H. N. J. (2020). Supporting food design processes: Development of food design cards. *International Journal of Design, 14*(2), 51–64.

Lee, D., Pan, R., & Fang, G. (2020). Improving early stage system design under the uncertainty in reliability-wise structure. *Journal of Engineering Design, 31*(10), 485–507.
Levin, J. R., & O'Donnell, A. M. (1999). What to do about educational research's credibility gaps? *Issues in Education, 5*(2), 177–229.
Lewrick, M., Link, P., & Leifer, L. (2018). *The design thinking playbook: Mindful digital transformation of teams, products, services, businesses and ecosystems*. John Wiley & Sons.
Lewrick, M., Link, P., & Leifer, L. (2020). *The design thinking toolbox: A guide to mastering the most popular and valuable innovation methods*. John Wiley & Sons.
Lloyd, P. (2019). You make it and you try it out: Seeds of design discipline futures. *Design Studies, 65*, 167–181.
Meyer, M. W., & Norman, D. (2020). Changing design education for the 21st century. *She Ji, 6*(1), 13–49.
Osterwalder, A., Pigneur, Y., Oliveira, M. A. Y., & Ferreira, J. J. P. (2011). Business model generation: A handbook for visionaries, game changers and challengers. *African Journal of Business Management, 5*(7), 22–30.
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). *The PRISMA 2020 statement: An updated guideline for reporting systematic reviews*. *The BMJ*. BMJ Publishing Group.
Paparistodimou, G., Duffy, A., Whitfield, R. I., Knight, P., & Robb, M. (2020). A network science-based assessment methodology for robust modular system architectures during early conceptual design. *Journal of Engineering Design, 31*(4), 179–218.
Pektaş, S. T., & Pultar, M. (2006). Modelling detailed information flows in building design with the parameter-based design structure matrix. *Design Studies, 27*(1), 99–122.
Pirinen, A., & Tervo, A. (2020). What can we share? A design game for developing the shared spaces in housing. *Design Studies, 69*, 100941.
Prochner, I., & Godin, D. (2022). Quality in research through design projects: Recommendations for evaluation and enhancement. *Design Studies, 78*, 101061.
Reich, Y. (2010). My method is better. *Research in Engineering Design, 21*(3), 137–142.
Rigger, E., Vosgien, T., Shea, K., & Stankovic, T. (2020). A top-down method for the derivation of metrics for the assessment of design automation potential. *Journal of Engineering Design, 31*(2), 69–99.
Rodgers, P. A., Mazzarella, F., & Conerney, L. (2020). Interrogating the value of design research for change. *The Design Journal, 23*(4), 491–514.
Roozenburg, N. F. M., & Eekels, J. (1995). *Product design: Fundamentals and methods*. Chichester: John Wiley Sons.
Seepersad, C. C., Pedersen, K., Emblemsvåg, J., Bailey, R., Allen, J. K., & Mistree, F. (2006). The validation square: How does one verify and validate a design method?. In *Decision making in engineering design* (pp. 305–316) ASME.
Simon, H. A. (2019). *The sciences of the artificial*. Cambridge: [M.I.T. Press.
Stappers, P. J., & Sanders, E. B. N. (2005). Tools for designers, Products for Users: The role of creative design techniques in a squeezed-in design process. In *The proceedings of the international conference on planning and design: Creative interaction and sustainable development* (pp. 1–16).
Stief, P., Dantan, J.-Y., Etienne, A., Siadat, A., & Burgat, G. (2020). Product design improvement by a new similarity-index-based approach in the context

of reconfigurable assembly processes. *Journal of Engineering Design, 31*(6), 349–377.
Strömberg, H., Pettersson, I., & Ju, W. (2020). Enacting metaphors to explore relations and interactions with automated driving systems. *Design Studies, 67*, 77–101.
Stylidis, K., Wickman, C., & Söderberg, R. (2020). Perceived quality of products: A framework and attributes ranking method. *Journal of Engineering Design, 31*(1), 37–67.
Vaajakallio, K., & Mattelmäki, T. (2014). Design games in codesign: As a tool, a mindset and a structure. *CoDesign, 10*(1), 63–77.
Vermaas, P. E. (2016). A logical critique of the expert position in design research: Beyond expert justification of design methods and towards empirical validation. *Design Science, 2*(May), e7.
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, K. (2011). Transferring design methods into practice. In H. Birkhofer (Ed.), *The future of design methodology*. (239–248). London: Springer.
Woo, J., & Lim, Y. (2020). Routinoscope: Collaborative routine reflection for routine-driven do-it-yourself smart homes. *International Journal of Design, 14*(3), 19–36.
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, 78*, 101063.