Journal of Engineering Design

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cjen20

Design method validation – an investigation of the current practice in design research

Matthias Eisenmann, Patric Grauberger, Selin Üreten, Dieter Krause & Sven Matthiesen

To cite this article: Matthias Eisenmann, Patric Grauberger, Selin Üreten, Dieter Krause & Sven Matthiesen (2021): Design method validation – an investigation of the current practice in design research, Journal of Engineering Design, DOI: 10.1080/09544828.2021.1950655

To link to this article: https://doi.org/10.1080/09544828.2021.1950655

JOURNAL OF ENGINEERING DESIGN
https://doi.org/10.1080/09544828.2021.1950655

Design method validation – an investigation of the current practice in design research

Matthias Eisenmann [a], Patric Grauberger [a], Selin Üreten [b], Dieter Krause [b] and Sven Matthiesen [a]

[a]IPEK – Institute of Product Engineering, Karlsruhe Institute of Technology (KIT), Kaiserstr. 10, 76131 Karlsruhe, Germany; [b]PKT – Institute of Product Development and Mechanical Engineering Design, Hamburg University of Technology (TUHH), Hamburg, Germany

ABSTRACT

Validation of design methods is a challenge in design research, as there is a lack of common methodology. This contribution investigates how validation of design methods is currently conducted in order to clarify this issue and identify approaches that might be suitable as best practices. A mixed-methods literature study including the years 2010–2020 is conducted. Systematic mapping structures the identified literature in an overview. The following state-of-the-art review focuses on challenges in validation and on how researchers address them. The overview of 54 identified studies shows a preponderance of non-comparative studies conducted in laboratory environment. Challenges arising in conducting experiments and field studies are caused by a lack of common metrics and established study designs. In success evaluation, the challenge is to objectively measure the effects of method application on the design outcome. Based on the identified examples, possible strategies to overcome the challenges are proposed.

ARTICLE HISTORY
Received 29 September 2020
Accepted 29 June 2021

KEYWORDS
Design methods; design research; research methods; evaluation; design method validation

1. Introduction

The application of design methods is an important factor in successful product development (Pahl et al. 2007; Cross 2008; Ehrlenspiel and Meerkamm 2017). Consequently, it is an essential goal of design research to provide appropriate methods (Blessing and Chakrabarti 2009). Validation is one important success factor to demonstrate the scientific contribution of the developed methods and ensure the adoption by companies (Jagtap et al. 2014).

However, the investigation of design methods is particularly difficult, as design is influenced by multiple factors, including the three following areas deemed central by the authors:

(1) The designer as a human being and, therefore, his/her cognitive and emotional interaction with the design problem at hand (Milojevic and Jin 2019) and his/her interaction with the team members (Cash, Škec, and Štorga 2019).

CONTACT Matthias Eisenmann matthias.eisenmann@kit.edu IPEK – Institute of Product Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

(2) The design problem, which is usually ill-structured (Simon 1977) and complex and unique in its appearance. Additionally, the problem may change during the process due to new insights gained by designing a solution (Dorst and Cross 2001) or due to environmental influences, such as new requirements of manufacturing (Sudin and Ahmed 2009).
(3) The environment in which the design problem is solved: From small tasks solved by single designers in a few minutes to extensive design projects conducted by whole companies over several years. This environment might also be subject to changes, e.g. through the availability of new technologies (see Zhang et al. 2020).

The influences arising from these areas cause interferences that need to be considered in the validation procedure, as they cannot be eliminated completely. One major difficulty in design method validation therefore lies in achieving the proper degree of simplification to reach a scientifically sound conclusion on the method's effects without leaving out crucial aspects emerging in design practice. In order to identify these crucial aspects, studies need to be conducted on design method application in practice.

To reach scientifically sound conclusions on relationships of cause and effect, study designs with a high internal validity are necessary, such as human subject experiments, where control of variables is possible (Hussy, Schreier, and Echterhoff 2013). However, conducting experiments in design practice is nearly impossible due to a multitude of influencing factors (Vermaas 2016) and their interrelations. This is affirmed by a review of articles in Research in Engineering Design (Barth, Caillaud, and Rose 2011) which identifies only four experimental studies out of 71 publications concerned with validation in design research.

Summing up the requirements for research, method validation is difficult, as it needs to consider two aspects at the same time: On the one hand, studies on method application in an industrial context are necessary to gain insights into design methods' effects on design practice. On the other hand, investigation in controlled contexts is important to test underlying assumptions and theories of cause and effect.

Multiple research approaches that support method validation have been developed (Wallace 2011). However, those approaches differ in their descriptions of necessary stages and appropriate research methods for validation (see also Section 2.2). Gericke, Eckert, and Stacey (2017) also conclude that there is a lack of consensus as to the research methods and evidence necessary for carrying out a successful validation of design methods. The state of the art does not provide an overview of how design methods are currently validated. We therefore conclude on the first research question:

RQ1: How are design methods currently validated in design research practice?

Building on the overview obtained from RQ1, the second research question aims to shed light on how researchers deal with challenges in design method validation:

RQ2: How are challenges in design method validation approached?

This paper answers both questions through a two-stage literature review structured as follows: Section 2 substantiates the focus by setting a frame for the investigation of design method validation. Section 3 describes the review procedure, which follows a mixed-methods approach containing systematic mapping and a state-of-the-art review. Section 4 discusses the review results to give an overview of current design method validation practice and to examine challenges as well as approaches to overcome them. Section 5 concludes with implications for research by outlining strategies to address the challenges in design method validation.

2. What is meant by design method validation? Clarification of terms and approaches

Design method validation describes investigations into whether a proposed design method fulfils its goal. There is, however, an incoherent use and understanding of the term *design method* (Gericke, Eckert, and Stacey 2017) and an ongoing discussion on the understanding and necessity of *method validation* (Vermaas 2016) in the design research community. To prevent misunderstandings, these terms are defined for further use in this paper. Also, two dimensions for categorising studies regarding design method validation are presented with the aim of structuring the identified papers: *Evaluation types* to address the objectives and scope of design method validation, and *levels of evidence* to address classification with respect to a possible internal validity of study designs.

2.1. Design methods and related concepts

Design methods can be seen as one type of design support (Blessing and Chakrabarti 2009). A *design method* can also be described as a concept that is part of or includes other types of support like methodologies, guidelines, processes, and tools (Gericke, Eckert, and Stacey 2017). The description of the term *design method* by Gericke, Eckert, and Stacey (2017, 105) is used here to differentiate this type of design support from other types of support (see Table 1):

> A specification of how a specified result is to be achieved. This may include specifications of how information is to be shown, what information is to be used as inputs to the method, what tools are to be used, what actions are to be performed and how, and how the task should be decomposed and how actions should be sequenced.

Tools, processes, and guidelines give strict prescriptions on how to perform tasks and interpret results. The authors chose to focus their investigation on design methods because of their recommendatory character, which leaves more room for adaption and interpretation. This makes design method validation particularly difficult. Therefore, in the authors' view, design methods pose unique challenges as compared to other types of design support. As methodologies contain multiple methods, it seems sensible to look at single methods first before expanding the view to whole methodologies.

2.2. Objectives and scope in the process of design method validation

As described before, design method validation includes various research activities ranging from lab experiments to case studies in practice. Recommendations on these activities and the corresponding research methods are described in approaches to design method validation. Definitions as to what constitutes *method validation* vary within these approaches depending on the activities described as being necessary for a comprehensive validation.

Tromp and Hekkert (2016) suggest two stages for the validation of effect-driven design methods. The first stage should include an in-depth study of one or multiple cases of

Table 1. Differentiation of the term *design method* from other *types of design support* based on Gericke, Eckert, and Stacey (2017).

Design support	Differentiation from *design method*
Tool	An object, artefact, or software that is used to perform an action.A tool can be used for multiple purposes whereas a design method is specifically developed to achieve a defined goal within the design process. A design method can include a tool by giving instructions of how to use it and when to implement it in the process.
Guideline	A strict prescription of activities to achieve a certain goal.A guideline is only to be violated for a good reason, which makes it more binding than a method that rather gives recommendations.
Process	A formally specified sequence of activities to be carried out in developing a particular design, or a class of designs.Methods rather specify how single results are to be achieved and do not always follow a strict sequence like processes do. Also, methods usually do not aim at developing a whole design but rather support steps or phases in the process.
Methodology	A methodology contains multiple methods and paradigms for thinking about the design problem and the priorities given to particular decisions or aspects of the design, or ways of thinking about the design.It is substantially more extensive than a single method.

method use to investigate the structure and logic of a preliminary method. This means that this stage is directly connected to method development and focusses on the applicability of the design method under development. The second stage should then consist of a comparative validation using test and control groups to validate the effect of the design method. The two stages are presented as a conclusion to a study which is later defined as a stage-one study. This is because no detailed advice is given on the necessary research methods and possible intermediate steps in stage two.

Marxen and Albers (2012) present a framework for the validation of design methods which builds on categories of design science by Cantamessa (2003). They suggest two steps of method validation:

(1) *Experimental research* in a controlled environment to test hypotheses on how the method supports designers.
(2) *Implementation* and *education studies* concerned with making the method available for either industrial application or for being taught to students.

These steps are then linked to the evaluation types described within DRM (Blessing and Chakrabarti 2009).

In the validation square (Pedersen et al. 2000), the usefulness of design methods is split up into qualitative and quantitative evaluation each at a theoretical and an empirical level. Pedersen et al. (2000) suggest two validation steps:

(1) *Structural validation* is concerned with the effectiveness of the method. This step includes the theoretical investigation of the logic and consistency of the method and of the suitability of example problems to be used in the subsequent empirical validation.
(2) *Performance validation* targets the efficiency of the method. This includes empirical investigations to quantify the effect of method application using the example problems. Further activities are concerned with establishing a connection to problems in practice by argumentation.

Hence, the validation square aims at establishing relationships by empirical investigations in a controlled context. A connection to practice is established by argumentation.

Empirical investigations of method application in practice are not part of this approach.

DRM (Blessing and Chakrabarti 2009) provides a description of design method validation activities by three different evaluation types for design support:

(1) *Support evaluation* aims at checking the design method for internal consistency and completeness and includes argumentation of its possible usefulness.
(2) *Application evaluation* aims at assessing the applicability and usability of the design method, i.e. at finding out whether designers can actually use the method and whether it addresses the intended factors directly related to the application.
(3) *Success evaluation* deals with the usefulness of a method in the intended context, i.e. the effect of method application on a success factor that is not directly related to the application. Success evaluation therefore aims at investigating the effect of the design method in practice.

The scope of validation increases throughout the stages from support evaluation to application evaluation through to success evaluation, and each of these evaluation types represents different objectives.

Summing up, the presented approaches have in common that they describe different stages necessary for a comprehensive design method validation. They start with an evaluation of the design method itself as regards consistency and applicability. This is followed by an empirical validation comprising the application of the method and the investigation of its effects in a successively wider context. Of the presented approaches, the DRM description of evaluation types is the only approach giving detailed advice on the steps to be conducted as it is designed as a guide to design research. It includes more aspects of design method validation than other approaches, which are limited to the presentation of frameworks within single papers or even sections of papers. A detailed description of the validation process including three stages concerned with different success criteria and a distinction between investigations of proximal and distal method effects (Blessing and Chakrabarti 2009) makes it possible to structure design method validation activities. This is because the DRM evaluation types propose a possibility to categorise method validation activities concerning objectives and scope over the whole process of design method validation.

Looking at different approaches revealed multiple stages of design method validation concerned with different criteria. To include all possible validation activities in all stages of validation in our review, we used the following definition: *Method validation includes all research activities that investigate whether a design method can fulfil its purpose for an intended context*.

2.3. Study designs and levels of evidence

When validating design methods, defining the requirements of the evidence and study design suitable for achieving scientifically sound results remains a challenge. Based on this challenge, Frey and Dym (2006) suggest to apply well-documented, objective, and evidence-based validation procedures like those established in the research on medical treatments. This seems sensible as both design methods and medical treatments aim at influencing human beings. Medical treatments achieve this by directly influencing the human body to effect positive changes in health. Design methods aim at changing designers' behaviour and thinking to influence the result of a specific activity during the design process. In both cases, researchers need to be sure that the 'treatment' they choose is the cause of a measurable effect. Since this effect may also be a negative one and unwanted negative effects are to be strictly avoided when applying methods in practice, it is especially relevant to reach a high validity of results.

Burns, Rohrich, and Chung (2011) describe how in evidence-based medicine, the type of study used to investigate the effectiveness of a medical treatment influences the rating of the evidence obtained. This rating is described in terms of *levels of evidence* that enable researchers to assess whether research results can be believed to be reliable. Sets of levels of evidence can be modified with respect to specialties (Burns, Rohrich, and Chung 2011) of investigations. In design method validation, it is also necessary to gain valid insights into the effects of the investigated method. In the light of the medical treatment – design method analogy described by Frey and Dym (2006), *levels of evidence* seem a fitting way of relating a study's design to a rating of evidence that can be obtained. *Levels of evidence* are therefore used in the mapping review (Section 3.1) to classify studies of design method validation based on their study design.

Summing up, a differentiation of *design method* from related concepts can help to focus on specific aspects of design method validation. Based on a discussion of existing approaches for method validation, it is apparent that there are multiple stages and objectives to be considered for a comprehensive design method validation. Different types of study designs are needed to achieve these objectives during validation. The stages can best be described by the DRM *evaluation types*. By looking at research approaches from other fields, an analogy to *levels of evidence* in evidence-based medicine can be drawn. To structure literature on design method validation in relation to stages and objectives, *evaluation types* and *levels of evidence* are thus selected as categories for review.

3. Method – review procedure

To answer the research questions, a systematic literature review is conducted. The review follows a two-step review procedure using a mixed-methods approach (Grant and Booth 2009) which is shown in Figure 1. To answer RQ1, systematic mapping is used to create an overview of the current practice in design method validation. DRM *evaluation types* and *levels of evidence* are used as categories. In order to answer RQ2, a more detailed state-of-the-art review of the categorised literature is conducted to identify implications for design method validation.

3.1. Systematic mapping review

The mapping review procedure is based on the systematic mapping process described by Petersen et al. (2008). Instead of identifying categories during the review, *evaluation types* and *levels of evidence* are used for classification. To identify relevant literature, several journals addressing design research have been selected for review: Design Studies, Research in Engineering Design (RIED), International Journal of Design (IJD), Journal of Engineering Design (JED), and Design Science. Design Studies, RIED, IJD, and JED have been selected

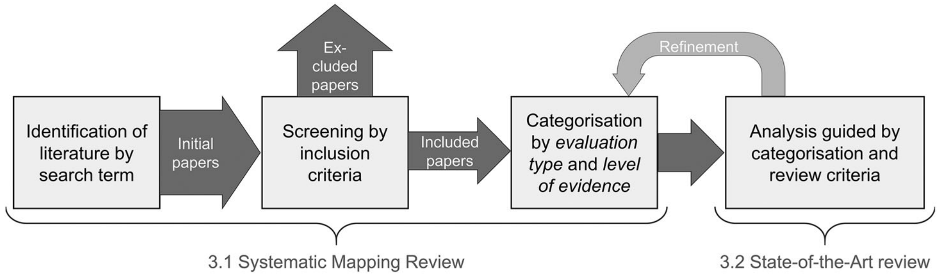

Figure 1. Visualisation of the two-step review procedure of the mixed-methods approach used.

as they are among the top ten research journals in the field of design research (Gemser et al. 2012) that explicitly include research on design methods and design methodology in their scope or in frequently used keywords. Design Science was included as a relatively new journal that might offer different approaches to design method validation. This selection of journals is to be seen as a sample used to gain insight into research activities concerning design method validation rather than comprehensively identifying studies introducing and validating design methods in the investigated period.

Identification of relevant literature. To review current practice, papers from the years 2010–2020 were included. Scopus was chosen as a research platform as it includes all journals to be examined. The search string was developed in order to identify as many validation activities as possible. As outlined before, there are many different notions of *design method* and *method validation* in design research. Therefore, an analysis of the journals' keywords related to design methods and validation was used to identify initial search terms. The initial terms were then used for a screening of articles from the different journals which resulted in the following search string:

('design method∗' OR 'development method∗' OR 'design technique∗' OR 'design strateg∗') AND ('validat∗' OR 'evaluat∗' OR 'assess∗' OR 'verif∗')

Subsequently, the identified literature was screened. By using the following inclusion criteria, literature out of the scope of 'studies of design method validation' was excluded from further investigation. Papers within scope were to fulfil two criteria:

(1) describe research on design methods
(2) describe validation activities (clarification of what is to be considered a validation activity see Section 2.2)

To distinguish design methods from other support in inclusion criterion (1), the differentiations shown in Table 1 were used together with the method description by Gericke, Eckert, and Stacey (2017).

Categorisation. After applying the inclusion criteria, the remaining papers were each categorised by *type of evaluation* and *levels of evidence*. Categorisation was supported by the descriptions given in Tables 2 and 3. A further distinction was made concerning success evaluation in order to enable a view on investigations that go further than application evaluation but still use a controlled context to do so (see Table 2).

Table 2. Categorisation of evaluation types, based on Blessing and Chakrabarti (2009).

Evaluation type	Description
Support evaluation	• Investigation concerning applicability or usability only • Application is illustrated by examples or argumentation
Application evaluation	Direct influence of method is investigated, i.e. on the specified result of the method (e.g. for brainstorming: number of ideas generated)
Success evaluation	Influence of the method on success factors is investigated, i.e. influence of the specified result on the design (e.g. in design for manufacturing, reduced assembly costs of the actual product)
• Lab	Investigation is conducted in a controlled context
• Field	Investigation is conducted into a real design process in a company

Table 3. Levels of evidence – adaption compared to levels of evidence for therapeutic studies from the Centre for Evidence-based Medicine as shown in Burns, Rohrich, and Chung (2011).

Original		Adaption	
Level	Type of evidence	Level	Type of evidence
1A	Systematic review of RCTs	I	Multiple randomised controlled trials (RCTs) or meta-analysis of RCTs
1B	Individual RCT		
1C	All-or-none study		
2A	Systematic review of cohort studies	II	Experiment (including RCT)
2B	Individual cohort study(incl. low-quality RCT)		
2C	‘Outcomes’ research; Ecological studies		
3A	Systematic review of case control studies	III	Comparative or correlation study
3B	Individual case control study		
4	Case series	IV	Descriptive case study
5	Expert opinion	V	Expert opinion or illustrative case

Categorisation according to the *level of evidence* was based on the terms used in the articles’ titles and abstracts to describe the study design. This categorisation was revised later in the state-of-the-art review, when descriptions in the full paper indicated another category. A modification of levels of evidence to suit the corresponding context is also customary in evidence-based medicine (Burns, Rohrich, and Chung 2011). Levels of evidence for therapeutic studies were chosen as the basis for modification (shown in Table 3) because therapy shows similarities to the support of design methods (see Section 2.3). In a first step of modification, sub-levels of the described levels 1–5 were summarised, because such fine granularity was not considered useful for the research goal of gaining an overview. These summarised levels were then specified in the context of designing. This was done in two ways:

(1) Identification of corresponding research methods in design research comparable to those in medicine to replace them (e.g. *cohort studies* in evidence-based medicine are similar to human subject *experiments* in design research, and *case control studies* are similar to *comparative case studies*)
(2) Adding research methods to levels which are customary in design research, such as the use of *illustrative cases* to show the potential value of a design method.

This resulted in a simplified and adapted version of levels of evidence that was used for categorisation (see Table 3).

The characteristics of each level of evidence concerning design method validation are explained in the following, starting with the lowest level:

Level V describes expert opinions or illustrative cases, which do not include actual application of a design method by a designer. This results in a lack of empirical evidence concerning the design method's usefulness. Illustrative cases implemented by researchers can only illustrate the application of design methods and therefore enable discussion on usability and applicability. In *Level IV,* design methods are actually applied by designers and investigated in a descriptive way. However, without comparing the observed effects during design method application with a similar situation without design method application, the effects cannot be directly attributed to the design method under investigation. This comparison is established in *Level III* by introducing a control group that performs a similar design task without using a design method or that uses a benchmark or placebo method. However, study designs on this level remain descriptive without formulating hypotheses prior to conducting the study. This results in correlations between method application and observed effects. These correlations form the basis for the formulation of hypotheses on design methods' positive effects. Causal relationships can only be obtained in *Level II* upwards by conducting experiments that aim at testing hypotheses on the design methods' desired effects. This includes controlling possible disturbances to focus the experiments' results solely on the method and its desired effects. *Level I* extends *Level II* by a multiple replication. This enables a reflection of the results in meta-analyses, which allow further insights into, e.g. influences of subject characteristics or the identification of possible weaknesses.

The categorisation into levels of evidence does *not* rate quality of research in general or the academic contribution of the investigated studies. Descriptive case studies, for example, make a significant contribution to understanding design processes and gaining detailed information on the application of methods, because they do not focus on single aspect. Case studies are especially helpful to identify phenomena only observable within a practical context (Teegavarapu, Summers, and Mocko 2008). This is because they form an important component in comprehensive design method validation. However, to validate a design method, comparative studies as described by Tromp and Hekkert (2016) and reliable relationships between method and outcome (Blessing and Chakrabarti 2009) are also necessary. Therefore, in order to obtain reliable results on causal relationships of method and effect, focusing investigations as in Levels II and I should be aimed for.

The proposed categorisation into *levels of evidence* enables the inclusion of necessary components in design method validation which can be attributed to different stages and objectives of validation by using the *DRM evaluation types*.

3.2. State-of-the-art review

To identify how researchers address challenges in design method validation, the second review step focussed on the validation activities in more detail. The categorisation obtained through systematic mapping was used as a basis in order to focus the review. By looking at each *evaluation type* separately, challenges corresponding to different stages of validation were to be identified. Categorisation into levels of evidence additionally allowed an analysis of how and why the corresponding research methods were used.

To establish interrelationships between method application and success of the design process, operationalisation of the desired outcome into criteria is necessary. Operationalisation also sets a scope for validation as different criteria can address results on different levels from a single step to success of the whole design process. Known criteria of method success are applicability, usability, and usefulness (Blessing and Chakrabarti 2009), or acceptance by the method user (Reiß, Albers, and Bursac 2017). In order to carry out design method validation in a targeted manner, these high-level criteria need to be operationalised further to enable an objective observation and measurement.

Thus, the state-of-the-art review was guided by criteria focussing on connections between the goals of the method under investigation and the characteristics of the validation activities:

(1) *Scope of method validation*
This criterion aims at understanding the scope in more detail. This includes the identification of criteria defined for method success within the different evaluation types and the studies' contexts.
(2) *Design method objectives*
This criterion allows the identification of similarities between design methods and therefore a comparison of research approaches used for investigation.
(3) *Validation metrics*
Metrics represent the most detailed level of operationalisation when investigating design methods and their effects. By analysing their connection to *design method objectives* and *scope of method validation*, the identification of challenges concerning operationalisation becomes possible.
(4) *Connections to related research* (own results, results from other researchers, or connection to theory)
This criterion specifically aims at identifying which challenges arise from the availability of knowledge concerning method validation for the chosen context. This includes insights into design available in the state of the art and into the availability of established research methodology.

The insights gained were used to derive the challenges arising in the validation of design methods and to point out possible approaches and strategies that are applied by researchers to overcome them.

4. Review results and discussion

Results addressing RQ1 are presented and discussed in Section 4.1 to give a first overview. Section 4.2 presents a deeper discussion building on results of a state-of-the-art review that aims at answering RQ2.

4.1. Overview of the current practice in design method validation

Identification of relevant literature. The literature search on Scopus in March 2020 resulted in a total of 456 papers. An overview of the number of papers found allocated to the journals examined is shown in Table 4. 52 papers (see Table A1 in the Appendix for a separate

Table 4. Resulting number of papers after literature search and after application of exclusion criteria.

	Design Science	Design Studies	IJD	JED	RIED	Total
After search	33	114	37	147	125	**456**
After screening	2	14	5	12	19	**52**

Table 5. Results of the categorisation of method validation studies.

	Evaluation type				
Levels of evidence	Support evaluation	Application evaluation	Success evaluation (lab)	Success evaluation (field)	Total
(I) Meta-analysis	-	-	-	-	**0**
(II) Experiment	-	7	1	-	**8**
(III) Comparative study	1	6	-	1	**8**
(IV) Descriptive study	-	9	3	2	**14**
(V) Expert opinion/illustrative example	14	5	3	2	**24**
Total	**15**	**27**	**7**	**5**	**54**

reference list) remained for categorisation after screening using the inclusion criteria. Most papers were excluded due to not fulfilling the criterion 'describing research on design methods' (382). Another 22 studies were excluded because they described research on a design method without validation.

Categorisation. The 52 resulting papers described 54 studies of design method validation, as two of the papers contained two separate validation studies that were conducted one after the other. A summary of the results of categorisation into the categories of *evaluation type* and *levels of evidence* is shown in Table 5 (for the full categorisation results, see Table A1 in the Appendix).

It is evident that the scope of validation regarding the *evaluation type* is often limited to application evaluation (27) or support evaluation (15). Success evaluation was addressed by 12 studies, seven in a laboratory environment, and five in the field.

When the scope of validation is limited to support evaluation, validation activities are usually carried out along with the method's development. During method development, comparative studies are not beneficial because it first needs to be investigated whether the design method's applicability is sufficient. To investigate applicability, it is necessary to know how the design method should achieve its objective. Experts on the activity that is supported by the design method can give insights specific to the approaches by participants. Nevertheless, human subject experiments can be useful at this stage when there is an existing benchmark, e.g. when comparing different variants of a design method or different methods concerning their applicability.

Application evaluation has the objective of investigating direct effects of method application on the results of design activities. Compared to success evaluation, the influence of possible disturbances is limited in application evaluation. This makes investigations in a controlled environment possible. Therefore, an experimental study design seems a reasonable choice, because it enables to establish causal relations between method application and outcomes of design activities. Indeed, most (7/8) of the identified experiments target application evaluation. However, the results also show that only seven out of 27 studies targeting application evaluation are experiments. A deeper discussion of this aspect is given in Section 4.2 as it is connected to the goals of the method that is investigated. In success evaluation, only one experiment is identified.

Regarding the *levels of evidence*, non-comparative studies of the evidence levels V (24) and IV (14) outweigh the analytical studies in evidence levels II (8) and III (8). It is also apparent that no studies could be identified reaching the meta-analysis level. For meta-analysis, multiple experiments using the same metrics and validating the same design methods are necessary. There are, however, only eight experiments in the sample that could possibly have been analysed in a meta-analysis study. In addition, not all of these studies investigate the same methods. The lack of meta-analysis in design method validation aligns with the findings by Cash (2018) who states that meta-analysis is severely limited in design research. The results presented here indicate that this lack is also apparent for the validation of design methods. Common standards for research methods are needed to make meta-analysis possible (Glass 1976; Cash 2018). The lack of meta-analysis therefore indicates a possible *lack of standardisation* in the field.

Summing up for answering RQ1. The presented overview provides an answer to *RQ1: How are design methods currently validated in design research practice?* A multitude of different studies spanning all *evaluation types* and *evidence levels* apart from meta-analysis are identified in the overview. However, there is a preponderance of studies focussing on the early stages of method validation in *support* and *application evaluation*. In *support evaluation*, the findings indicate that expert opinion and illustrative cases are common practice in the validation of design methods. Whereas *application evaluation* shows a broad scatter of different research methods applied for validation. Also, there are two areas crucial for method validation that seem to be underrepresented: Experimental validation targeting *application evaluation* as well as studies targeting *success evaluation* in the field in general. This indicates challenges and further research potential concerning both areas. Arising challenges are investigated further to answer RQ 2 in 4.2.

4.2. Challenges and solution approaches in design method validation

To identify challenges regarding standardisation, the identified studies were analysed to find design methods with common features. This was followed by an analysis of studies with the same *evaluation type* to identify challenges at different stages of validation. It became apparent that validation of design methods focussing on ideation forms the biggest group with 15 studies spanning all evidence levels apart from meta-analysis. Also, all of the eight identified experimental studies are investigations of methods for ideation. These 15 studies are used as a starting point to identify how challenges in design method validation could be approached. This is done by contrasting them to examples of the 39 studies validating other types of methods.

4.2.1. Challenges in application evaluation

Challenge 1: Lack of standardisation. The research on ideation methods shows a certain degree of standardisation connected to common method goals. However, a look at the remaining identified validation studies reveals a multitude of different method goals which hinder a differentiation of method classes.

Apart from several ideation methods, only the *persona method* was investigated more than once by different groups of researchers but with differing foci of investigation. The persona method aims at creating a fictitious character – the persona – to represent a whole category of future customers by assigning a set of typical attributes to it (Brangier and Bornet 2011). The method is used to foster user-centred product design by summarising traits relevant to future customers. Miaskiewicz and Kozar (2011) investigate this method's possible usefulness whereas Turner and Turner (2011) focus on the mitigation of possible negative effects of the method in design. Additionally, both studies perform a support evaluation without actually applying the method and therefore do not define metrics.

By searching for a field that is addressed by multiple methods, four studies targeting the modularisation of products and development of product families were identified. The methods presented in these papers all address different goals concerning product family development and therefore measure different metrics. Baylis, Zhang, and McAdams (2018) aim for finding a trade-off between *commonality* and *quality of modular architecture*, whereas Pakkanen, Juuti, and Lehtonen (2016) evaluate their method by *costs*. Jung and Simpson (2016) especially develop metrics to measure the *modularity* and *sparsity* to investigate *connectivity in the platform architecture*, and Koh et al. (2015) build on a metric that is part of a design method developed in earlier work to support the prioritisation of component modularisation. All four contributions try to quantify the characteristics of modular product structures by operationalising them through metrics. The field of modular product families has developed a large number of such metrics. Several authors analysed success factors in this field before in order to reduce their variety and to form an overarching understanding of the relevant concepts (Gershenson, Prasad, and Zhang 2003; Salvador 2007). However, by looking at the high variety of metrics, one can assume that this is a still ongoing discussion in the field.

The analysis of current studies shows that most design method developers set goals for their own methods and therefore develop a separate operationalisation resulting in a multitude of metrics. This makes it very challenging to compare methods with each other and hinders researchers to build a common standard for similar methods. The effort for operationalisation increases dramatically for higher levels of evidence, making it difficult for researchers to reach them without a common standard. Because of the multifaceted character of design and a lot of possible areas being addressed by various design methods, complete standardisation of the field is not purposeful. However, the value of standardisation needs to be discussed in each particular sub-area. There are multiple examples for areas in design which could benefit from the further development of common standards. For example, research into design fixation is challenged by missing standards, as Vasconcelos and Crilly (2016) argue convincingly through a review of 25 experimental studies which use 14 different metrics to investigate the same phenomenon. The need for standardisation is also supported by the advice by Cash (2018, 108) who encourages researchers to 'ensure that, wherever possible, the measures and methods used include standard elements to facilitate meta-analytics.'

Solution approach A: Common goals and metrics. Ideation methods follow the common goal of fostering creativity. This results in a very similar operationalisation of goals and use of metrics in all of the 15 identified studies validating ideation methods. Eight of these studies explicitly build on the four metrics proposed by Shah, Smith, and Vargas-Hernandez (2003), which are *quantity, quality, novelty,* and *variety*. The remaining studies found their operationalisation on other work but also measure at least one of the four aspects. Additionally, other metrics such as *idea quality* (Wierenga and van Bruggen 1998) composed of *originality* and *appropriateness* (Massetti 1996), or *unobviousness* (Howard, Culley, and Dekoninck 2006) as used by Howard, Culley, and Dekoninck (2011) show a strong connection to the other concepts. This similarity in operationalisation of common method goals enables a comparison of results of different researchers and different ideation methods. Besides, the similarity in metrics enables a critical comparison like it is done by Chulvi et al. (2012), who studied different types of ideation methods to discuss the applicability of different metrics. The use of common metrics also fosters theoretical discussion as is shown in current reflections on Shah's novelty metric (Fiorineschi, Frillici, and Rotini 2020) and enables researchers to further develop new approaches to analyse ideation outcomes on a more detailed level, as in Hay et al. (2020).

The example of ideation methods shows that the definition of common goals and metrics for similar design methods leads to the comparability of results and thus the further development of both the design methods themselves and the metrics for validation. This enables researchers to reach higher levels of evidence and is the basis for applying meta-analyses that enable identification of reliable insights.

The presented examples indicate that future research in application evaluation should focus on identifying common goals and metrics of similar design methods where appropriate. This might intensify exchange between researchers and foster the discussion of results on common grounds.

Discussion of Solution Approach A. The presented approach builds on best practices identified within application evaluation of design ideation methods. However, ideation targets an early stage of design. It might therefore benefit from a small number of influencing factors compared to later stages of design, where many additional requirements must be considered in method development. This might foster standardisation in ideation. Indeed the majority of methods within the 27 identified studies concerning application evaluation focus on ideation (15) or on the concept stage (7). This raises the question, whether the design stage a method aims for limits standardisation within application evaluation.

The five remaining studies in our sample can yield further insight in this regard. Two of those studies target modularisation in application evaluation. As already described before, modularisation as a research area in design is in the process of building common goals and metrics despite addressing a later stage in the design process. An additional example in our sample is the validation study of Moultrie and Maier (2014) which introduces a method for design for assembly, targeting the stage of detail design. Moultrie and Maier (2014) refer to common standards and concepts in design for assembly to define goals and metrics for their method. Both examples show that it is possible to establish common goals and metrics for methods targeting different stages of the design process.

Application evaluation deals with a defined and controlled situation to establish direct effects of design methods. This should be especially difficult for design methods which address multiple stages in the design process or the process as a whole. Two such methods are described in the remaining two studies in our sample. Ćatić and Malmqvist (2013) present and evaluate a method addressing knowledge management throughout the whole design process. Consequently, they study a single case within a company to learn about the effects of their method through the whole process. While this process and the occurring effects might be unique in their nature, the authors nevertheless define goals which might be used to study the effects of their method in single stages of design. Another example of how to conduct validation of design methods spanning multiple design stages is given by Ahmad, Wynn, and Clarkson (2013). They conduct a laboratory experiment on their method for engineering change management by modelling the design process in a laboratory context.

The presented examples from modularisation, design for assembly and knowledge as well as change management illustrate, that is indeed possible to define concise goals to validate design methods addressing different or even multiple design stages. However, it might take a lot of effort to do so when including influences and requirements arising in later stages of design like restrictions of manufacturing.

Solution Approach B: Using common tools and models to identify similarities. The identified studies on method validation, except for studies on ideation methods, lack common goals even if they show similarities. Similar elements within different methods can contribute to the discussion of direct effects of those elements. Like this, apart from common goals, other links between the validated methods can be identified.

Similar elements can be well-established models and tools such as the Design Structure Matrix (DSM). Nine methods identified in the review included DSM, three of them (Hamraz et al. 2015; Loureiro, Ferreira, and Messerschmidt 2020; Tilstra, Seepersad, and Wood 2012) as a central element that was developed further. This illustrates that there are common elements which multiple design methods build upon.

Tools and models help to document knowledge on the product and the design process in a simplified and standardised manner, which makes this knowledge comparable between different design contexts. By comparing the documented knowledge between multiple studies using the same tools or models, criteria for successful design related to this knowledge might be identified. The tools and models used can then become part of either methods which have the goal to achieve those criteria or become research tools to assess those criteria.

An example on how to possibly define goals from the use of tools can be derived from the study of Ćatić and Malmqvist (2013). The design method investigated in their study addresses the generation of engineering checklists as a means for knowledge management during design. By applying their method in a case study, Ćatić and Malmqvist (2013) illustrate in which design stages their proposed method was deemed useful by users and why. By using engineering checklists, the knowledge of the method users was documented in a simplified form. This enabled the identification of the most relevant knowledge areas for different contexts in the design process. Future studies on design knowledge management could use these insights by assessing knowledge with engineering checklists to (1) reach comparable results on relevant knowledge areas in industrial application or (2) benchmark their own methods for knowledge management against the design method of Ćatić and Malmqvist (2013). That means the design tool *engineering checklist* could also be used as a research tool for the assessment of relevant knowledge.

Discussion of Solution approach B. Models and tools are used frequently in design what makes them a fruitful starting point to search for similarities. However, every use of a tool as well as every modelling process includes abstraction of the initial information. On the one hand, this abstraction supports in focusing on a smaller number of aspects and therefore reduces complexity. On the other hand, it bears the risk of over-simplification, which is also discovered by Ćatić and Malmqvist (2013) through the feedback of method users. Additionally, the use of common models and tools might foster a fixation on criteria related to them rather than identifying the real issues. This concern is also raised by Eckert, Stacey, and Clarkson (2003) who see the danger of selecting over-specific methods for research.

However, current research might be over-specific in its own way as a lot of researchers define goals and metrics only suitable for their own design method. Common models and tools might at least contribute a common ground for initial discussion of metrics which are suitable for a group of methods that integrate those models and tools. Additionally, an initial reduction of complexity is needed to make investigation of single effects possible, especially in the multi-faceted field of design research. The resulting simplification still needs to be checked for its external validity within real practice after showing its value in a controlled environment.

4.2.2. Challenges in success evaluation

Challenge 2: Transition from the lab to the field. *Support* and *application evaluation* are necessary for comprehensive method validation and are also helpful to discuss newly developed design methods in the community. However, they are not sufficient for research into relationships between method application and its effects in practice. It seems challenging for researchers to take the next step after introducing a new method. However, there are multiple examples in the sample of continuous development and validation of design methods. For example, Baek, Meroni, and Manzini (2015) present a method to analyse communities to form design goals. The method is initially validated through a support evaluation using an illustrative example. In a follow-up study to further investigate the developed design method, they conducted a success evaluation in a descriptive case study (Baek et al. 2018). What stands out is a concise operationalisation of method success.

Challenge 3: High effort for success evaluation in the field. With the goal of design methods to support design practice, it is surprising that only five of the identified studies investigate the effects of methods on success in the field (see Table 5). This might be due to the high effort required for studies in design practice. In the identified studies, only one comparative study addresses success evaluation in the field (Snelders, Morel, and Havermans 2011). However, in this study, results of the same method are compared as regards different cultural contexts rather than comparing method application with unguided development or a benchmark.

A challenge that arises in success evaluation is objectively measuring the effects of method application on the design outcome. Additionally, a high effort is required to conduct studies in the field, where success of the design method should be measured. Criteria on method success for design practice are often unclear. Even in the case of established criteria, the information necessary for measurement might be unavailable in practice. In a real-world design process, studies to clearly establish effects of method application therefore generally are a challenge.

In ideation, requirements are known to researchers. Here, researchers use common method goals and validation metrics (see Section 4.2.1). This should enable research to target success evaluation in the field. However, the identified studies on ideation methods mostly (13/15) target support or application evaluation. This might originate from common goals that are connected to direct, proximal effects of method application. To enable a comprehensive design method validation, there is a need to define long-term goals, which are linked to objectives in design practice.

Solution approaches to success evaluation. A possible strategy to connect measurement of direct effects with objectives in design practice is shown by the investigation by Cardin et al. (2013), which is the only experimental success evaluation in the investigated papers (see Table 5). Cardin et al. (2013) connected direct effects of an ideation method on the flexibility of solutions with a rating of the lifecycle performance of the created concepts. They operationalised a design practice objective through the metric 'anticipated lifecycle performance' and connected it with the metric 'flexibility' which can be directly affected by the ideas generated. This enables experimental success evaluation in a laboratory study. The developed metrics can then be used to investigate the influence of other design methods or supports on flexibility and lifecycle performance. This is shown in multiple following studies by Cardin (Cardin, Jiang, and Lim 2017; Cardin, de Neufville, and Geltner 2015; Cardin, Ranjbar-Bourani, and de Neufville 2015; Mak et al. 2018; Hu and Cardin 2015). The strategy shows operationalisation of multiple stages of goals addressing direct proximal effects (e.g. *quantity of ideas*) over intermediate (e.g. *flexibility*) to long-term distal goals (e.g. *lifecycle performance*). This seems to be a strategy to establish common goals not only for application but also for success evaluation of design methods.

Approaches to establishing such goals are also visible in the work on product families and modularisation as presented in Section 4.2.1. This can be seen as an intermediate step on the way to validate design methods in practice as it connects issues from practice with the validation activities in the lab.

4.2.3. Overarching solution approach: using potentials of theory

The identified research on ideation methods is strongly connected to each other, which is visible in mutual citations of research groups in the field. Also, well-established ideation methods are investigated multiple times. These are, e.g. TRIZ (Chulvi et al. 2013; Fiorineschi et al. 2018), the 6-3-5 method (Wodehouse and Ion 2012; Petersson and Lundberg 2018), and, most of all, brainstorming, which is used as a basis for research in three ways: Firstly, to enable replication of results (Chulvi et al. 2012, 2013), secondly, to be used as a benchmark (Cardin et al. 2013; Hatcher et al. 2018), and thirdly, as a basis for further development by, for example, adding new stimuli (Vandevenne, Pieters, and Duflou 2016; Keshwani et al. 2017; Howard, Culley, and Dekoninck 2011).

In the research on methods supporting ideation, a common body of knowledge can be used for the development and validation of methods. Especially, validation against a benchmark produces valuable insights and enables a deeper understanding of relations between the core mechanisms of the effects and design methods. This fosters insights that can be used to build and test theories. For example, Chulvi et al. (2013) compare different types of ideation methods to gain insights into the influence of the core ideas of different method types on ideation, which they use in their subsequent work (Mulet et al. 2016; Chulvi et al. 2017). This enables them to form contributions to theory that link the type of information used to creativity. Subsequently, these contributions can lead to the development of new ideation methods that can be validated again.

Design models can also enable a connection to theory and therefore to established and well-known concepts within design. For example, Hamraz et al. (2015) connect DSM to Gero's function-behaviour structure (FBS) ontology of design (Gero and Kannengiesser 2014). He and Gu (2016) develop their method using FBS, whereas Weisbrod and Kroll (2018) develop their idea configuration evaluation (ICE) method by implementing the steps in the knowledge and concept spaces of C-K design theory (Hatchuel and Weil 2009). By connecting design methods to elements of theory, these elements provide a commonly understood starting point for discussing method core mechanisms and their effects on design.

By connecting research and building on existing results, design method validation has the potential to become part of the theory building/testing cycle as proposed by Cash (2018).

Summing up for answering RQ2. By analysing the categorised studies in more detail, RQ2: How are challenges in design method validation approached? can be answered. The challenges arising by a lack of standardisation in application and success evaluation can be approached by defining common goals and metrics for similar design methods. The field of ideation methods provides an example of how to define such goals and metrics by a strong connection with theory of creativity. Tools and models implemented in design methods can provide a common ground to form such goals and help to develop fitting research tools. Also, a comprehensive validation including multiple stages from support evaluation to success evaluation in the field poses a challenge to researchers. The identified approaches to address this challenge use different perspectives: One approach (see Section 4.2.2) uses design practice objectives and breaks them down into measurable criteria to bring practice to the laboratory. This reduces the effort required for validation and helps to compare different methods targeting the same objectives of design practice. The second approach (see Section 4.2.3) uses the established theory as a starting point. Either by connecting a design method under development explicitly with theoretical models or by using criteria from theory for validation of already developed methods. This approach facilitates the discussion of results for all researchers who are familiar with the corresponding theory. Additionally, it enables the application of theoretical concepts and a further transfer to practice.

The presented findings are subject to certain limitations. By limiting the scope of the review to a certain number of journals and certain keywords, relevant literature using other keywords or published in other journals may not have been identified. This was mitigated by selecting high-rated journals with a fitting scope. Additionally, the search string used was developed iteratively by screening papers and keywords to include synonyms and differing descriptions of validation activities. Besides, the identified approaches originate from a small number of papers which were selected by the authors. This prevents a comprehensive view of approaches and includes a possible bias in the selection of papers. However, the aim of this contribution is to point out possible ways of current practice to address challenges and put them into perspective of the used categories rather than giving an exhaustive list of possible approaches.

5. Conclusion

The presented investigation of the current practice shows that the validation of design methods remains a demanding and evolving task in design research. The investigation of the different stages of design method validation shows a preponderance of studies focussing on the early stages of *support* and *application evaluation*. While there seems to be a standard for research methods in using expert opinion or illustrative cases in *support evaluation*, standardisation for further stages is missing. A deeper reflection on studies on validating ideation methods shows that common design method objectives can help to define standards for method validation and enable exchange as well as a target-oriented development of design methods. By developing common metrics for validation, experiments become possible that enable establishment of causal relationships and comparisons of methods with suitable benchmarks. With a connection of these relationships to theory, a common understanding of the effects of the core concepts of design methods can be built. We therefore argue that validation as part of application evaluation should be done using human subject experiments.

Additionally, the presented review shows a lack of success evaluation of design methods in general. We argue that the lack of studies in this context originates from the high effort required and from missing strategies for the transition from the laboratory to the field. Examples from ideation and other fields show that the first step to success evaluation in the field can be taken by systematically breaking down factors from practice and operationalising them for studies in the laboratory. In this way, the gap between lab and field becomes smaller and easier to bridge in a next step.

To enable comprehensive design method validation in the future, we suggest the use of established strategies for validation from other research areas. Validation in mechanical design starts in the laboratory, and each technical system has to fulfil requirements on functionality as well as security before being integrated in a product which can be sold. Medical treatments also have to be thoroughly tested in the lab for their efficacy, followed by large-scale investigations to show that their value is greater than possible side effects. The authors believe the same to be true for design methods. For being accepted in practice, design methods should prove their value in the laboratory before being brought into industrial practice. We therefore suggest to put research effort into the discussion of appropriate operationalisations connecting different method goals to aim at a certain degree of standardisation. In this way, comparable human subject experiments on method usefulness become possible. Only by reaching an agreement on what is necessary for a design method to be valid at this early stage of method validation, we can proceed to further validation aiming at success in the field.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Matthias Eisenmann http://orcid.org/0000-0002-1007-4260
Patric Grauberger http://orcid.org/0000-0002-8367-3889
Selin Üreten http://orcid.org/0000-0003-1996-411X
Dieter Krause http://orcid.org/0000-0002-1253-1699
Sven Matthiesen http://orcid.org/0000-0001-5978-694X

References

Ahmad, N., D. C. Wynn, and P. J. Clarkson. 2013. "Change Impact on a Product and its Redesign Process: A Tool for Knowledge Capture and Reuse." *Research in Engineering Design* 24 (3): 219–244. doi:10.1007/s00163-012-0139-8.

Augustine, M., O. P. Yadav, R. Jain, and A. Rathore. 2012. "Cognitive map-Based System Modeling for Identifying Interaction Failure Modes." *Research in Engineering Design* 23 (2): 105–124. doi:10.1007/s00163-011-0117-6.

Bacciotti, D., Y. Borgianni, G. Cascini, and F. Rotini. 2016. "Product Planning Techniques: Investigating the Differences Between Research Trajectories and Industry Expectations." *Research in Engineering Design* 27 (4): 367–389. doi:10.1007/s00163-016-0223-6.

Baek, J. S., S. Kim, Y. Pahk, and E. Manzini. 2018. "A Sociotechnical Framework for the Design of Collaborative Services." *Design Studies* 55: 54–78. doi:10.1016/j.destud.2017.01.001.

Baek, J. S., A. Meroni, and E. Manzini. 2015. "A Socio-Technical Approach to Design for Community Resilience: A Framework for Analysis and Design Goal Forming." *Design Studies* 40: 60–84. doi:10.1016/j.destud.2015.06.004.

Barth, Alex, Emmanuel Caillaud, and Bertrand Rose. 2011. "How to Validate Research in Engineering Design?" In ICED11 international conference on engineering design, Lyngby/Copenhagen, Denmark.

Baylis, K., G. Zhang, and D. A. McAdams. 2018. "Product Family Platform Selection Using a Pareto Front of Maximum Commonality and Strategic Modularity." *Research in Engineering Design* 29 (4): 547–563. doi:10.1007/s00163-018-0288-5.

Blessing, Lucienne T.M., and Amaresh Chakrabarti. 2009. *DRM, a Design Research Methodology*. London: Springer.

Brahma, A., and D. C. Wynn. 2020. "Margin Value Method for Engineering Design Improvement." *Research in Engineering Design* 31: 353–381. doi:10.1007/s00163-020-00335-8.

Brangier, Eric, and Corinne Bornet. 2011. "Persona: A Method to Produce Representations Focused on Consumers' Needs." In *Human Factors and Ergonomics in Consumer Product Design*, edited by Waldemar Karwowski, 37–61. Ergonomics Design and Management: Theory and Applications. Boca Raton, FL: CRC Press.

Burns, Patricia B., Rod J. Rohrich, and Kevin C. Chung. 2011. "The Levels of Evidence and Their Role in Evidence-Based Medicine." *Plastic and Reconstructive Surgery* 128 (1): 305–310. doi:10.1097/PRS.0b013e318219c171.

Camere, S., H. N. J. Schifferstein, and M. Bordegoni. 2018. "From Abstract to Tangible: Supporting the Materialization of Experiential Visions with the Experience map." *International Journal of Design* 12 (2): 51–73.

Cantamessa, Marco. 2003. "An Empirical Perspective upon Design Research." *Journal of Engineering Design* 14 (1): 1–15. doi:10.1080/0954482031000078126.

Cardin, Michel-Alexandre, Richard de Neufville, and David M. Geltner. 2015. "Design Catalogs: A Systematic Approach to Design and Value Flexibility in Engineering Systems." *Systems Engineering* 18 (5): 453–471. doi:10.1002/sys.21323.

Cardin, Michel-Alexandre, Yixin Jiang, and Terence Lim. 2017. "Empirical Studies in Decision Rule-Based Flexibility Analysis for Complex Systems Design and Management." In *Complex Systems Design & Management*, edited by Gauthier Fanmuy, Eric Goubault, Daniel Krob, and François Stephan, 171–185. Cham: Springer.

Cardin, Michel-Alexandre, G. L. Kolfschoten, Daniel D. Frey, Richard de Neufville, O. L. de Weck, and David M. Geltner. 2013. "Empirical Evaluation of Procedures to Generate Flexibility in Engineering Systems and Improve Lifecycle Performance." *Research in Engineering Design* 24 (3): 277–295. doi:10.1007/s00163-012-0145-x.

Cardin, Michel-Alexandre, Mehdi Ranjbar-Bourani, and Richard de Neufville. 2015. "Improving the Lifecycle Performance of Engineering Projects with Flexible Strategies: Example of On-Shore LNG Production Design." *Systems Engineering* 18 (3): 253–268. doi:10.1002/sys.21301.

Cash, P. J. 2018. "Developing Theory-Driven Design Research." *Design Studies* 56: 84–119. doi:10.1016/j.destud.2018.03.002.

Cash, P. J., S. Škec, and M. Štorga. 2019. "The Dynamics of Design: Exploring Heterogeneity in Meso-Scale Team Processes." *Design Studies* 64: 124–153. doi:10.1016/j.destud.2019.08.001.
Ćatić, A., and J. Malmqvist. 2013. "Effective Method for Creating Engineering Checklists." *Journal of Engineering Design* 24 (6): 453–475. doi:10.1080/09544828.2013.766824.
Chulvi, V., M. C. González-Cruz, E. Mulet, and J. Aguilar-Zambrano. 2013. "Influence of the Type of Idea-Generation Method on the Creativity of Solutions." *Research in Engineering Design* 24 (1): 33–41. doi:10.1007/s00163-012-0134-0.
Chulvi, V., E. Mulet, Amaresh Chakrabarti, B. López-Mesa, and C. González-Cruz. 2012. "Comparison of the Degree of Creativity in the Design Outcomes Using Different Design Methods." *Journal of Engineering Design* 23 (4): 241–269. doi:10.1080/09544828.2011.624501.
Chulvi, V., E. Mulet, F. Felip, and C. García-García. 2017. "The Effect of Information and Communication Technologies on Creativity in Collaborative Design." *Research in Engineering Design* 28 (1): 7–23. doi:10.1007/s00163-016-0227-2.
Corremans, J. A. M. 2011. "Measuring the Effectiveness of a Design Method to Generate Form Alternatives: An Experiment Performed with Freshmen Students Product Development." *Journal of Engineering Design* 22 (4): 259–274. doi:10.1080/09544820903312416.
Cross, Nigel. 2008. *Engineering Design Methods: Strategies for Product Design*. 4th ed. Chichester: Wiley.
da Cunha Barbosa, G. E., and G. F. M. de Souza. 2017. "A Risk-Based Framework with Design Structure Matrix to Select Alternatives of Product Modernisation." *Journal of Engineering Design* 28 (1): 23–46. doi:10.1080/09544828.2016.1258458.
Dorst, Kees, and Nigel Cross. 2001. "Creativity in the Design Process: co-Evolution of Problem–Solution." *Design Studies* 22 (5): 425–437. doi:10.1016/S0142-694X(01)00009-6.
Eckert, C., M. K. Stacey, and P. J. Clarkson. 2003. "The Spiral of Applied Research: A Methodological View on Integrated Design Research." In ICED03 international conference on engineering design, Stockholm, Sweden.
Ehrlenspiel, Klaus, and Harald Meerkamm. 2017. *Integrierte Produktentwicklung: Denkabläufe, Methodeneinsatz, Zusammenarbeit*. 6th ed. Munich: Carl Hanser Verlag GmbH & Co. KG.
Eisenbart, B., and M. Kleinsmann. 2017. "Implementing Shared Function Modelling in Practice: Experiences in six Companies Developing Mechatronic Products and PSS." *Journal of Engineering Design* 28 (10-12): 765–798. doi:10.1080/09544828.2017.1395395.
Fiorineschi, L., F. S. Frillici, and F. Rotini. 2020. "Impact of Missing Attributes on the Novelty Metric of Shah et al." *Research in Engineering Design* 31: 221–234. doi:10.1007/s00163-020-00332-x.
Fiorineschi, L., F. S. Frillici, F. Rotini, and M. Tomassini. 2018. "Exploiting TRIZ Tools for Enhancing Systematic Conceptual Design Activities." *Journal of Engineering Design* 29 (6): 259–290. doi:10.1080/09544828.2018.1473558.
Frey, Daniel D., and Clive L. Dym. 2006. "Validation of Design Methods: Lessons from Medicine." *Research in Engineering Design* 17 (1): 45–57. doi:10.1007/s00163-006-0016-4.
Gemser, Gerda, Cees de Bont, P. Hekkert, and Ken Friedman. 2012. "Quality Perceptions of Design Journals: The Design Scholars' Perspective." *Design Studies* 33 (1): 4–23. doi:10.1016/j.destud.2011.09.001.
Gericke, Kilian, Claudia Eckert, and Martin Stacey. 2017. "What Do We Need to Say about a Design Method." In ICED17 international conference on engineering design, 101–110. Vancouver, Canada.
Gero, John S., and Udo Kannengiesser. 2014. "The Function-Behaviour-Structure Ontology of Design." In *An Anthology of Theories and Models of Design*, edited by Amaresh Chakrabarti and Lucienne T. M. Blessing, 263–283. London: Springer.
Gershenson, J. K., G. J. Prasad, and Y. Zhang. 2003. "Product Modularity: Definitions and Benefits." *Journal of Engineering Design* 14 (3): 295–313. doi:10.1080/0954482031000091068.
Glass, Gene V. 1976. "Primary, Secondary, and Meta-Analysis of Research." *Educational Researcher* 5 (10): 3–8. doi:10.3102/0013189X005010003.
Grant, Maria J., and Andrew Booth. 2009. "A Typology of Reviews: an Analysis of 14 Review Types and Associated Methodologies." *Health Information and Libraries Journal* 26 (2): 91–108. doi:10.1111/j.1471-1842.2009.00848.x.

Hamraz, B., N. H. M. Caldwell, T. W. Ridgman, and P. J. Clarkson. 2015. "FBS Linkage Ontology and Technique to Support Engineering Change Management." *Research in Engineering Design* 26 (1): 3–35. doi:10.1007/s00163-014-0181-9.

Hatcher, G., W. Ion, R. Maclachlan, M. Marlow, B. Simpson, N. Wilson, and A. Wodehouse. 2018. "Using Linkography to Compare Creative Methods for Group Ideation." *Design Studies* 58: 127–152. doi:10.1016/j.destud.2018.05.002.

Hatchuel, Armand, and Benoit Weil. 2009. "C-K Design Theory: an Advanced Formulation." *Research in Engineering Design* 19 (4): 181–192. doi:10.1007/s00163-008-0043-4.

Hay, L., A. H. B. Duffy, M. Grealy, M. Tahsiri, C. McTeague, and T. Vuletic. 2020. "A Novel Systematic Approach for Analysing Exploratory Design Ideation." *Journal of Engineering Design* 31 (3): 127–149. doi:10.1080/09544828.2019.1662381.

He, B., and Z. Gu. 2016. "Sustainable Design Synthesis for Product Environmental Footprints." *Design Studies* 45: 159–186. doi:10.1016/j.destud.2016.04.001.

Hofstetter, W. K., and E. F. Crawley. 2013. "A Methodology for Portfolio-Level Analysis of System Commonality." *Research in Engineering Design* 24 (4): 349–373. doi:10.1007/s00163-012-0151-z.

Howard, T. J., S. J. Culley, and E. A. Dekoninck. 2006. "Information as an Input into the Creative Process." In DESIGN2006 – international design conference, Vol. 9.

Howard, T. J., S. Culley, and E. A. Dekoninck. 2011. "Reuse of Ideas and Concepts for Creative Stimuli in Engineering Design." *Journal of Engineering Design* 22 (8): 565–581. doi:10.1080/09544821003598573.

Hu, J., and Michel-Alexandre Cardin. 2015. "Generating Flexibility in the Design of Engineering Systems to Enable Better Sustainability and Lifecycle Performance." *Research in Engineering Design* 26 (2): 121–143. doi:10.1007/s00163-015-0189-9.

Hussy, Walter, Margrit Schreier, and Gerald Echterhoff. 2013. *Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor*. Berlin: Springer.

Hwang, D., and W. Park. 2018. "Design Heuristics set for X: A Design aid for Assistive Product Concept Generation." *Design Studies* 58: 89–126. doi:10.1016/j.destud.2018.04.003.

Jagtap, S., A. Warell, V. Hiort, D. Motte, and A. Larsson. 2014. "Design Methods and Factors Influencing Their Uptake in Product Development Companies: A Review." In DESIGN2014 – International design conference, 231–240.

Jung, E.-C., and K. Sato. 2010. "Methodology for Context-Sensitive System Design by Mapping Internal Contexts Into Visualization Mechanisms." *Design Studies* 31 (1): 26–45. doi:10.1016/j.destud.2009.07.001.

Jung, S., and T. W. Simpson. 2016. "An Integrated Approach to Product Family Redesign Using Commonality and Variety Metrics." *Research in Engineering Design* 27 (4): 391–412. doi:10.1007/s00163-016-0224-5.

Karana, E., B. Barati, V. Rognoli, and A. Zeeuw van der Laan. 2015. "Material Driven Design (MDD): A Method to Design for Material Experiences." *International Journal of Design* 9 (2): 35–54.

Keshwani, S., T. A. Lenau, S. Ahmed-Kristensen, and Amaresh Chakrabarti. 2017. "Comparing Novelty of Designs from Biological-Inspiration with Those from Brainstorming." *Journal of Engineering Design* 28 (10-12): 654–680. doi:10.1080/09544828.2017.1393504.

Kimita, K., T. Sakao, and Y. Shimomura. 2018. "A Failure Analysis Method for Designing Highly Reliable Product-Service Systems." *Research in Engineering Design* 29 (2): 143–160. doi:10.1007/s00163-017-0261-8.

Koh, E. C. Y., N. H. M. Caldwell, and P. J. Clarkson. 2013. "A Technique to Assess the Changeability of Complex Engineering Systems." *Journal of Engineering Design* 24 (7): 477–498. doi:10.1080/09544828.2013.769207.

Koh, E. C. Y., A. Förg, M. Kreimeyer, and M. Lienkamp. 2015. "Using Engineering Change Forecast to Prioritise Component Modularisation." *Research in Engineering Design* 26 (4): 337–353. doi:10.1007/s00163-015-0200-5.

Kroll, E., and G. Weisbrod. 2020. "Testing and Evaluating the Applicability and Effectiveness of the new Idea-Configuration-Evaluation (ICE) Method of Conceptual Design." *Research in Engineering Design* 31 (1): 103–122. doi:10.1007/s00163-019-00324-6.

López-Mesa, B., and N. Bylund. 2011. "A Study of the use of Concept Selection Methods from Inside a Company." *Research in Engineering Design* 22 (1): 7–27. doi:10.1007/s00163-010-0093-2.

Loureiro, G. B., J. C. E. Ferreira, and P. H. Z. Messerschmidt. 2020. "Design Structure Network (DSN): a Method to Make Explicit the Product Design Specification Process for Mass Customization." *Research in Engineering Design* 31: 197–220. doi:10.1007/s00163-020-00331-y.

Mak, W. H. J., Michel-Alexandre Cardin, Liu Ziqi, and P. J. Clarkson. 2018. "Towards the Design of Resilient Waste-to-Energy Systems Using Bayesian Networks." In 44th Design automation conference, ASME.

Marxen, L., and Albert Albers. 2012. "Supporting Validation in the Development of Design Methods." In DESIGN2012 – international design conference.

Massetti, Brenda. 1996. "An Empirical Examination of the Value of Creativity Support Systems on Idea Generation." *MIS Quarterly* 20 (1): 83–97.

Miaskiewicz, T., and K. A. Kozar. 2011. "Personas and User-Centered Design: How Can Personas Benefit Product Design Processes?" *Design Studies* 32 (5): 417–430. doi:10.1016/j.destud.2011.03.003.

Milojevic, Hristina, and Yan Jin. 2019. "Building a Social-Cognitive Framework for Design: Personality and Design Self-Efficacy Effects on Pro-Design Behaviors." In *Design Computing and Cognition '18*, edited by John S. Gero, 323–339. Cham: Springer.

Moreno, D. P., A. A. Hernández, M. C. Yang, K. N. Otto, K. Hölttä-Otto, J. S. Linsey, K. L. Wood, and A. Linden. 2014. "Fundamental Studies in Design-by-Analogy: A Focus on Domain-Knowledge Experts and Applications to Transactional Design Problems." *Design Studies* 35 (3): 232–272. doi:10.1016/j.destud.2013.11.002.

Moultrie, J., and A. M. Maier. 2014. "A Simplified Approach to Design for Assembly." *Journal of Engineering Design* 25 (1-3): 44–63. doi:10.1080/09544828.2014.887059.

Mulet, E., V. Chulvi, M. Royo, and J. Galán. 2016. "Influence of the Dominant Thinking Style in the Degree of Novelty of Designs in Virtual and Traditional Working Environments." *Journal of Engineering Design* 27 (7): 413–437. doi:10.1080/09544828.2016.1155697.

Nagel, R. L., R. Hutcheson, D. A. McAdams, and R. Stone. 2011. "Process and Event Modelling for Conceptual Design." *Journal of Engineering Design* 22 (3): 145–164. doi:10.1080/09544820903099575.

Nam, T.-J., and C. Kim. 2011. "Design by Tangible Stories: Enriching Interactive Everyday Products with Ludic Value." *International Journal of Design* 5 (1): 85–98.

Pahl, Gerhard, Wolfgang Beitz, Lucienne T. M. Blessing, Jörg Feldhusen, Karl-Heinrich Grote, and K. Wallace. 2007. *Engineering Design: A Systematic Approach*. 3rd ed. London: Springer.

Pakkanen, J., T. Juuti, and T. Lehtonen. 2016. "Brownfield Process: A Method for Modular Product Family Development Aiming for Product Configuration." *Design Studies* 45: 210–241. doi:10.1016/j.destud.2016.04.004.

Pedersen, K., Jan Emblemsvag, R. Bailey, J. K. Allen, and F. Mistree. 2000. "The "Validation Square" – Validating Design Methods." Proceedings of DETC '00, 2000 ASME design engineering technical conferences, Baltimore, MD, USA.

Petersen, Kai, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. 2008. "Systematic Mapping Studies in Software Engineering." In EASE: Electronic workshops in computing. Electronic Workshops in Computing: BCS Learning & Development.

Petersson, A. M., and J. Lundberg. 2018. "Developing an Ideation Method to be Used in Cross-Functional Inter-Organizational Teams by Means of Action Design Research." *Research in Engineering Design* 29 (3): 433–457. doi:10.1007/s00163-018-0283-x.

Petersson, A. M., J. Lundberg, and M. Rantatalo. 2017. "Ideation Methods Applied in a Cross-Functional Inter-Organizational Group: an Exploratory Case Study from the Railway Sector." *Research in Engineering Design* 28 (1): 71–97. doi:10.1007/s00163-016-0238-z.

Reiß, Nicolas, Albert Albers, and Nikola Bursac. 2017. "Approaches to Increasing Method Acceptance in Agile Product Development Processes." In ICED17 international conference on engineering design, Vancouver, Canada.

Salvador, Fabrizio. 2007. "Toward a Product System Modularity Construct: Literature Review and Reconceptualization." *IEEE Transactions on Engineering Management* 54 (2): 219–240. doi:10.1109/TEM.2007.893996.

Shah, Jami J., Steve M. Smith, and Noe Vargas-Hernandez. 2003. “Metrics for Measuring Ideation Effectiveness.” *Design Studies* 24 (2): 111–134. doi:10.1016/S0142-694X(02)00034-0.

Simon, Herbert A. 1977. “The Structure of Ill-Structured Problems.” In *Models of Discovery*. Vol. 54, edited by R. S. Cohen, M. W. Wartofsky, and Herbert A. Simon, 304–225. Boston Studies in the Philosophy of Science. Dordrecht: Springer.

Snelders, D., K. P. N. Morel, and P. Havermans. 2011. “The Cultural Adaptation of web Design to Local Industry Styles: A Comparative Study.” *Design Studies* 32 (5): 457–481. doi:10.1016/j.destud.2011.03.001.

Sohn, M., and T.-J. Nam. 2015. “Understanding the Attributes of Product Intervention for the Promotion of pro-Environmental Behavior: A Framework and its Effect on Immediate User Reactions.” *International Journal of Design* 9 (2): 55–77.

Stark, R., T. Buchert, S. Neugebauer, J. Bonvoisin, and M. Finkbeiner. 2017. “Benefits and Obstacles of Sustainable Product Development Methods: A Case Study in the Field of Urban Mobility.” *Design Science* 3: e17. doi:10.1017/dsj.2017.20.

Strömberg, H., I. Pettersson, J. Andersson, A. Rydström, D. Dey, M. Klingegård, and J. Forlizzi. 2018. “Designing for Social Experiences with and Within Autonomous Vehicles-Exploring Methodological Directions.” *Design Science* 4: e13. doi:10.1017/dsj.2018.9.

Sudin, Mohd N., and Saeema Ahmed. 2009. “Investigation of Change in Specifications During a Product’s Lifecycle.” In ICED09 international conference on engineering design, 371–380. Vancouver, Canada.

Teegavarapu, Sudhakar, Joshua D. Summers, and Gregory M. Mocko. 2008. “Case Study Method for Design Research: A Justification.” In Volume 4: 20th international conference on design theory and methodology; second international conference on micro- and nanosystems, 495–503: ASMEDC.

Tilstra, A. H., C. C. Seepersad, and K. L. Wood. 2012. “A High-Definition Design Structure Matrix (HDDSM) for the Quantitative Assessment of Product Architecture.” *Journal of Engineering Design* 23 (10-11): 767–789. doi:10.1080/09544828.2012.706748.

Tromp, N., and P. Hekkert. 2016. “Assessing Methods for Effect-Driven Design: Evaluation of a Social Design Method.” *Design Studies* 43: 24–47. doi:10.1016/j.destud.2015.12.002.

Turner, P., and S. Turner. 2011. “Is Stereotyping Inevitable When Designing with Personas?” *Design Studies* 32 (1): 30–44. doi:10.1016/j.destud.2010.06.002.

van der Bijl-Brouwer, M., and K. Dorst. 2017. “Advancing the Strategic Impact of Human-Centred Design.” *Design Studies* 53: 1–23. doi:10.1016/j.destud.2017.06.003.

Vandevenne, D., T. Pieters, and J. R. Duflou. 2016. “Enhancing Novelty with Knowledge-Based Support for Biologically-Inspired Design.” *Design Studies* 46: 152–173. doi:10.1016/j.destud.2016.05.003.

Vasconcelos, Luis A., and Nathan Crilly. 2016. “Inspiration and Fixation: Questions, Methods, Findings, and Challenges.” *Design Studies* 42: 1–32. doi:10.1016/j.destud.2015.11.001.

Vermaas, P. 2016. “A Logical Critique of the Expert Position in Design Research: Beyond Expert Justification of Design Methods and Towards Empirical Validation.” *Design Science* 2: e7. doi:10.1017/dsj.2016.6.

Wallace, K. 2011. “Transferring Design Methods Into Practice.” In *The Future of Design Methodology*, edited by Herbert Birkhofer, 239–248. London: Springer.

Weisbrod, G., and E. Kroll. 2018. “Idea-configuration-evaluation (ICE): Development and Demonstration of a new Prescriptive Model of the Conceptual Engineering Design Process Based on Parameter Analysis and C–K Theory.” *Research in Engineering Design* 29 (2): 203–225. doi:10.1007/s00163-017-0263-6.

Wierenga, Berende, and Gerrit H. van Bruggen. 1998. “The Dependent Variable in Research Into the Effects of Creativity Support Systems: Quality and Quantity of Ideas.” *MIS Quarterly* 22 (1): 81–87.

Wodehouse, A. J., and W. J. Ion. 2010. “Information use in Conceptual Design: Existing Taxonomies and new Approaches.” *International Journal of Design* 4 (3): 53–65.

Wodehouse, A., and W. Ion. 2012. “Augmenting the 6-3-5 Method with Design Information.” *Research in Engineering Design* 23 (1): 5–15. doi:10.1007/s00163-011-0110-0.

Zhang, Guanglu, Elissa Morris, Douglas Allaire, and Daniel A. McAdams. 2020. “Research Opportunities and Challenges in Engineering System Evolution.” *Journal of Mechanical Design* 142 (8): 081401-1–081401-14. doi:10.1115/1.4045908.

Appendix. Identified papers describing design method validation studies

Table A1. Categorisation of design method validation studies identified in systematic mapping.

Evaluation type	Level of evidence	Identified papers describing design method validation studies
Support	III	Koh et al. (2015) [Study 2]
Support	V	Augustine et al. (2012)
Support	V	Bacciotti et al. (2016)
Support	V	Baek, Meroni, and Manzini (2015)
Support	V	Brahma and Wynn (2020)
Support	V	da Cunha Barbosa and de Souza (2017)
Support	V	He and Gu (2016)
Support	V	Jung and Simpson (2016)
Support	V	Koh, Caldwell, and Clarkson (2013)
Support	V	Loureiro, Ferreira, and Messerschmidt (2020)
Support	V	Miaskiewicz and Kozar (2011)
Support	V	Tilstra, Seepersad, and Wood (2012)
Support	V	Turner and Turner (2011)
Support	V	van der Bijl-Brouwer and Dorst (2017)
Support	V	Weisbrod and Kroll (2018)
Application	II	Chulvi et al. (2013)
Application	II	Chulvi et al. (2012)
Application	II	Corremans (2011)
Application	II	Howard, Culley, and Dekoninck (2011)
Application	II	Hwang and Park (2018)
Application	II	Keshwani et al. (2017)
Application	II	Vandevenne, Pieters, and Duflou (2016)
Application	III	Hatcher et al. (2018)
Application	III	Moreno et al. (2014)
Application	III	Petersson and Lundberg (2018)
Application	III	Petersson, Lundberg, and Rantatalo (2017)
Application	III	Sohn and Nam (2015)
Application	III	Wodehouse and Ion (2012)
Application	IV	Ahmad, Wynn, and Clarkson (2013)
Application	IV	Camere, Schifferstein, and Bordegoni (2018) [Study 1]
Application	IV	Camere, Schifferstein, and Bordegoni (2018) [Study 2]
Application	IV	Fiorineschi et al. (2018)
Application	IV	López-Mesa and Bylund (2011)
Application	IV	Moultrie and Maier (2014)
Application	IV	Nagel et al. (2011)
Application	IV	Strömberg et al. (2018)
Application	IV	Wodehouse and Ion (2010)
Application	V	Baylis, Zhang, and McAdams (2018)
Application	V	Ćatić and Malmqvist (2013)
Application	V	Hofstetter and Crawley (2013)
Application	V	Kimita, Sakao, and Shimomura (2018)
Application	V	Koh et al. (2015) [Study 1]
Success (lab)	II	Cardin et al. (2013)
Success (lab)	IV	Kroll and Weisbrod (2020)
Success (lab)	IV	Nam and Kim (2011)
Success (lab)	IV	Tromp and Hekkert (2016)
Success (lab)	V	Hamraz et al. (2015)
Success (lab)	V	Jung and Sato (2010)
Success (lab)	V	Karana et al. (2015)
Success (field)	III	Snelders, Morel, and Havermans (2011)
Success (field)	IV	Baek et al. (2018)
Success (field)	IV	Stark et al. (2017)
Success (field)	V	Eisenbart and Kleinsmann (2017)
Success (field)	V	Pakkanen, Juuti, and Lehtonen (2016)