Proceedings of the Design Society, Volume 5: ICED25
https://doi.org/10.1017/pds.2025.10299

Scenario-based planning of design method validation studies and the associated effort

Olga Sankowski[1,✉], Lukas Paehler[2], Christoph Wittig[2], Selin Uereten[1], Matthias Eisenmann[2], Sven Matthiesen[2] and Dieter Krause[1]

[1] TUHH - Hamburg University of Technology, Germany, [2] KIT - Karlsruhe Institute of Technology, Germany

✉ o.sankowski@tuhh.de

ABSTRACT: Design method validation is fundamental to ensure that design methods achieve their objectives in the intended situations and are accepted in practice. Although various method validation approaches have been developed, there is still a lack of practical guidance for planning validation studies based on project characteristics. To address this, an intensity map of the validation effort is presented as the core of a scenario-based planning approach. It categorizes projects according to the novelty of the method and the state of research on the problem or the research area, enabling the required validation studies, their sequence and validation criteria to be identified. Thereby, researchers can plan validation studies and estimate the required effort situation-based, allowing for a better alignment with their individual project characteristics before starting studies.

KEYWORDS: research methodologies and methods, design methods, evaluation, validation, study planning

1. Introduction

1.1. Motivation

Design method research aims to provide engineering designers with valuable design methods for product development. Design methods are specifications of how certain results are to be achieved and can include the activities to be carried out, their sequence, or the information to be used for them (Gericke et al., 2017). The validation of design methods provided to engineering designers in practice is an important step, as it ensures that the developed methods fulfil their previously defined objectives and can be used in the intended real-world problems. Inadequate validation can lead to weaknesses in the design methods remaining unrecognised (Gericke et al., 2017) and consequently to the design methods not being accepted in practice (Wallisch et al., 2021). In addition, the ability to provide validated methods is necessary for method research as a scientific field to increase scientific, theoretical, and methodological rigour (Cash, 2018).

In the context of product development in general, the validation of a system is understood as the review of the system properties against a predefined purpose or benefit; in other words, it answers the question of whether the right system has been developed (Verein Deutscher Ingenieure, 2021). In contrast to this and as a preliminary stage to validation, a distinction is made with the term 'verification'. This refers to the review of the system in relation to the specification at different system levels and levels of concretisation and poses the question of whether the design process was correct and whether the requirements placed on the system were fulfilled (Verein Deutscher Ingenieure, 2021). Within the validation of design methods, the criteria for proving the achievement of objectives and their dependence on design methods themselves pose a challenge. For example, development projects in which design method validation is conducted are mostly unique and not reproducible. In addition, their results may be influenced,

consciously or unconsciously, by the design method researcher him/herself, resulting in a lack of objectivity.
Various approaches have therefore been developed to increase the objectivity and reproducibility of design method research and to assist researchers when assessing the fulfilment of a predefined purpose. For example:

- Gottfredson et al. (2015) outline three evaluation stages: efficacy (testing under optimal conditions), effectiveness (assessment in real-world settings), and dissemination (implementation on a broader scale). Each stage includes standards for planning, execution, and documentation. While relevant to prevention science, their application to design method research is limited, as real-world studies in companies often involve few participants.
- Eckert et al. (2003) describe in their eight fold model of design research a step-by-step process in which a theory is first derived based on empirical studies and then evaluated. Only afterwards, methodological tools and processes are developed, tested and finally transferred to industrial practice.
- Pedersen et al. (2000) emphasise step-by-step validation and distinguish between structural validation and performance validation. By structural validation, they understand the proof that a method 'provides design solutions correctly' and by performance validation, that a method 'provides correct design solutions' (Pedersen et al., 2000, p. 5).
- The Design Research Methodology (DRM) (Blessing & Chakrabarti, 2009) describes a comprehensive process that goes beyond the validation of the methods itself. It contains various aids relating to method research and a multi-stage validation procedure consisting of three validation steps: support evaluation, application evaluation and success evaluation. Within these steps, various validation criteria, such as functionality, usability and usefulness, are evaluated.
- Cash et al. (2023) describe four key properties of a 'good method', i.e. defined, predictable, useable, desirable. They further proposed a framework for assessment of design methods using a chain of evidence the method development and evaluation has to follow to get the label of a good method; the framework can also be used for development and reporting of a design method.

The examples show that many references and authors suggest a multi-step process, or different levels of quality for validation (see Blessing & Chakrabarti, 2009; Eckert et al., 2003; Gottfredson et al., 2015; Pedersen et al., 2000; Verein Deutscher Ingenieure, 2021). On further inspection, it is also evident that different authors admit a contextual and situational dependency of the validation process. Gericke et al. (2017) acknowledge that different elements of methods need different approaches to validation. Isaksson et al. (2020) refer to available resources and the intended scope of application of a method, such as the product to be developed, as influencing factors for the possibility of conducting multiple studies. The more specific the field of application, the less likely it is that several case studies can be carried out in this field. In the eight fold model of design research, a complete run-through of the process is according to the authors too extensive for a single doctoral thesis and requires a long-term project or several project groups (Eckert et al., 2003). Within the DRM approach up to 7 different types of method development and validation projects, i.e., doctoral projects, can be distinguished (Blessing & Chakrabarti, 2009). Which steps within the approach are treated extensively depends on the project characteristics 'state of research on the problem' and 'existence of prior work on the design method to be analysed' (Blessing & Chakrabarti, 2009). However, the authors do not explain how the researchers should identify or determine which type applies to their own research project. The assignment can therefore only be made retrospectively and cannot be determined in advance for planning purposes. In conclusion, existing approaches for method validation acknowledge the need to consider the characteristics of projects but do not guide researchers on how to take them into account when planning method validation studies.

1.2. Research focus

In summary, various authors have defined approaches for method validation and have also named some project characteristics that affect their exercise. However, the problem is that the existing approaches do not provide applicable scenario-based differentiation for supporting the planning of individual method validation studies, hindering the estimation of the related effort and necessary studies. Without an applicable scenario-based differentiation, which maps the project characteristics to the validation procedure, it is difficult to estimate which validation steps can be carried out within a time-limited methodological research project as well as which validation degree, i.e. validation step, can be achieved at the end of the project and how great the total validation effort in that case is. Without this distinction, it also stays unclear to what extent and by how much the validation effort is increased, if the new design method differs greatly from established ones.

The aim of this paper is therefore to take a first step towards closing this gap by proposing a scenario-based planning approach for individual method validation projects. Based on the existing literature, a procedural approach for the estimating of validation effort depending on project characteristics is proposed, that shall support the planning of method validation studies. In preparation for this, Section 2 provides the background on different validation steps and associated validation criteria named by different authors and concludes with a consolidation of the various views. The consolidation serves as a working basis to include different project characteristics in the proposed scenario-based planning approach in Section 3 and to link the characteristics to the effort and project scenarios. Finally, the approach for scenario-based design method validation is discussed (Section 4) and further research endeavours are described (Section 5).

2. Background on validation steps and validation criteria

As they build the basis for scenario-based planning, background on widespread design method validation approaches and their validation steps will be provided, divided by the type of research, i.e., theoretical or empirical work, and specified by the quality criteria to consider, e.g., internal validity, the type of the study, i.e., qualitative or quantitative, and the validation criteria. The terms from the various references in Section 2.1, which are taken up again in the consolidation in Section 2.2, are emphasised in italics.

2.1. Different views on validation steps and validation criteria

The most common approach is the DRM. It involves three validation steps: support evaluation, application evaluation and success evaluation. The first validation step, the support evaluation, is already achieved as part of method development. Among other things, the *internal consistency* of the method and the fulfilment of the requirements are checked (Blessing & Chakrabarti, 2009). *Internal consistency* is given if the method consists of a logical and structured flow of method steps that build on each other (Pedersen et al., 2000). The second validation step, the application evaluation, should be carried out as part of an empirical study to evaluate the design method. In addition to providing evidence of usability and *applicability* this validation step should also provide evidence that the key factors according to the impact model have been achieved (Blessing & Chakrabarti, 2009). The last validation step serves to prove the achievement of the usefulness of the method. This requires evidence of achievement of the (measurable) success factors.

In the Validation Square, the authors make a fundamental distinction between the main categories of *structural validation* and *performance validation* (Pedersen et al., 2000). In addition, there is a distinction between *qualitative* and *quantitative* studies as well as *empirical* and *theoretical* proof. Similarities can be found between the Validation Square and the DRM, for example in the first step, where *internal consistency* of the design method is analysed. By this the *theoretical structural validation* is established; it is thus closely related to the support evaluation from the DRM (Blessing & Chakrabarti, 2009). A difference can be found in the final step. It addresses the transferability of the experiment results (*theoretical performance validity*) to problems other than the example problem (Pedersen et al., 2000). Thereby, they refer to the *external validity* of the design method, which is thus to be proven on a theoretical level. *External validity* can be understood as a quality criterion of validation and is given if the results of a study can be generalised to the population or other situations (Hussy et al., 2013). According to Pedersen et al. (2000), proof of *external validity* at the theoretical level, i.e. based on reasoning and logical derivation, may be sufficient for the initial proof. However, they add that in the case of design methods, external validity can only be confirmed by various case studies and therefore cannot be provided in the same context as the initial proof of the method's performance. The stage of dissemination, introduced by (Gottfredson et al., 2015), goes even further than the examination of the generalisability of methodological results to larger populations and other cases of application (*external validity*) within the effectiveness trials or *performance validation*.

Eckert et al. (2003) explain that design methods and tools should first be tested for their usability and, in a further step, only in serious industry contexts with regard to the *effective achievement* of goals. The distinction between laboratory and industrial contexts and various validation criteria is emphasised here.

The same distinction is made by different authors (Daalhuizen & Cash, 2021; Gottfredson et al., 2015). Thus, the *efficacy* of the design method can be evaluated in a *controlled study environment*, and *effectiveness* of the design method under *real-life conditions*. According to Daalhuizen and Cash (2021), the sum of the two corresponds to the usefulness of a method. With regard to the proof of usefulness, Marxen and Albers (2012) add that this can only be provided by evaluating the method in *different real contexts*, i.e. different companies. This is an extension of the concept of *external validity*, the proof of which on a theoretical level is also part of the Validation Square (Pedersen et al., 2000).
Cash et al. (2023) do not clearly distinguish between different validation studies or steps, but they do identify properties of good methods. The properties ‘defined’ and ‘useable’ overlap with the validation criteria of *internal consistency* and *applicability*, which have already been mentioned elsewhere. The property of predictability has no equivalent in other works, but here there is also a need for operationalisation in order to be able to use it as a validation criterion. The property ‘desirable’ could also be defined as design method *acceptance*, which has been investigated in various publications. An elaborated model for *acceptance* is the Technology Acceptance Model (TAM) and its subsequent extensions TAM2 and TAM3 (Venkatesh & Bala, 2008). These take, for example, experience, voluntariness, job relevance, and perceived enjoyment as influencing factors into account. Wallisch et al. (2021) add to this and emphasise that *acceptance* is largely dependent on the organisational and task-related context. However, the cause-and-effect relationships of the influencing factors on the phenomena of *acceptance* formation have not yet been universally studied. As a result, its evidence is valid for the measured context and can only be predicted to a limited extent for further contexts. (Cash et al., 2023) further add that exogenous factors such as marketing and promotion have a major influence on the adoption and popularity of methods. Thus, these factors should be deliberately excluded in the validation process.

2.2. Consolidation of the different views

As a basis for scenario-based planning of design method validation studies, a consolidation of the different views is presented here. For this, we use the designations from Eisenmann and Matthiesen (2022); they have already consolidated three of the approaches mentioned above, i.e. (Blessing & Chakrabarti, 2009; Marxen & Albers, 2012; Pedersen et al., 2000), into three overarching validation steps. These are: structure validity, effect validity, and performance validity. These three steps are expanded to a total of six sub-steps, and further information about the research types, i.e., whether studies are empirical or theoretical, the study types, i.e., whether the research should be conducted qualitatively or quantitatively, as well as the quality criteria of the research are added. The overview is complemented by the consolidated validation criteria. The result is shown in Figure 1.
In Figure 1, an increasing and consecutive degree of validation can be found from left to right, consisting of three superordinate **validation steps** and six sub-steps of empirical and theoretical research. We refrained from an additional stage of dissemination, which would have to be categorised after performance validation or in parallel with empirical evaluation in further industry cases. The broad implementation and scaling of a method for long-term application by a large group of people depends, to a greater extent on exogenous factors such as marketing and promotion, see (Cash et al., 2023), as well as on top management support and the superordinate processes in the company.

Validation step	**Structure Validation**		**Effect Validation**	**Performance Validation**		
Research type	theoretical evaluation:	empirical evaluation in controlled test environment	empirical evaluation in controlled test environment	empirical evaluation in designated context	theoretical evaluation	empirical evaluation in further industry cases
Quality criteria to consider		(internal validity)	internal validity	(internal validity)		internal & external validity
Study type		(mostly) qualitative	quantitative & qualitative	quantitative & qualitative		(mostly) quantitative
Validation criteria	requirements-fulfilment, internal consistency, completeness	applicability, functionality	efficacy, applicability, acceptance	effectiveness, applicability, acceptance	theoretical estimation on external validity	effectiveness, efficiency, acceptance in further contexts/ with other users/ measuring further success factors

Figure 1. Consolidated validation process as working basis for scenario-based validation

Below the **research types**, the **quality criteria** of the validation studies to be observed are assigned. In addition to external validity, where the generalisability of the goal achievement to other populations or other situations is achieved and/or evidence is provided by means of other variables (Gottfredson et al., 2015; Hussy et al., 2013), internal validity is also mentioned. Internal validity ensures that observed changes in the dependent variable are actually caused by the independent variable, rather than by extraneous influences (Gottfredson et al., 2015; Hussy et al., 2013). It is a desirable quality criterion for empirical studies, but is not explicitly required in small studies with few participants, and is therefore placed in parentheses, as would be expected in empirical structure validation and empirical performance validation.

With regard to the **study types**, a distinction is made between qualitative and quantitative research methods. Pedersen et al. (2000) recommend first testing validation criteria qualitatively, then quantitatively. Individual case studies or studies with a small number of participants tend to be conducted qualitatively rather than quantitatively. A causal relationship (see internal validity) can also be approximated by repeating and varying studies (Hussy et al., 2013). Furthermore, there are also mixed methods approaches that are based on analysing relationships first quantitatively and then qualitatively with a reduced sample from the same participant group (Hussy et al., 2013). The decision on the research method and type of survey depends, among other things, on the success factors or key factors of the design method, that shall be examined, and is therefore situation-dependent.

In terms of **validation criteria**, the acceptance of design methods should be considered in all validation studies with realistic or real-world problems (effect and performance validation), following the findings of Wallisch et al. (2021). The other validation criteria are shown in the figure.

3. Scenario-based planning of the effort of design method validation

As outlined in the literature review in Section 2, design method validation is a multi-step, partly iterative process. A newly developed method should be validated step-wise, first with regard to its structure, then with regard to the effect in a laboratory environment and finally with regard to performance in a real context, while also considering applicability and acceptance in multiple evaluations. As studies in the industrial context are difficult to replicate and subject to various biases, such as the interviewer effect, it is considered critical to rely solely on a single application of the method to provide empirical performance validation without first effect validation. In comparison, a study under controlled conditions with a sample problem, i.e., effect validation, adds more weight to the performance validation statements. However, from the perspective of a time-limited individual research project, not all validation steps can always be carried out; instead, the steps required in the individual project scenario have to be carried out with the necessary rigour and focus.

Method research, like any other research, does not take place in a state of vacuum; often there are already established methods from other areas of application that can be modified, complemented or further developed for one's own problem so that the preliminary work on the established method can be taken as partial validation of the modified method. In another scenario, where research is based on an existing method, but the evaluation has not yet been successfully carried out, the first process stages up to the prescriptive studies might be completely literature-based. Furthermore, it is also possible that the research area in which the new design method is positioned is already well researched and many measurable success factors for proving effectiveness have been already derived here in the form of metrics or even impact models (regarding impact models, see (Blessing & Chakrabarti, 2009)). This preliminary work can be used directly for one's own validation studies. However, in the absence of measurable success factors in the literature, these must be derived first within comprehensive research and problem clarification studies before performance validation can be attempted. Therefore, in the following subsections, an approach is presented to include these aspects in the planning of design method validation studies. This approach consists of an intensity map of validation effort (Section 3.1) and recommendations on how to validate a design method based on scenarios distinguished within the intensity map (Section 3.2).

3.1. Intensity map of validation effort

An intensity map of the validation effort was derived from the literature, spanning the dimensions of 'novelty of the method' and 'state of research on the problem or the research area' (see Figure 2). Each of the dimensions can be divided into three areas. The validation effort increases towards the right upper

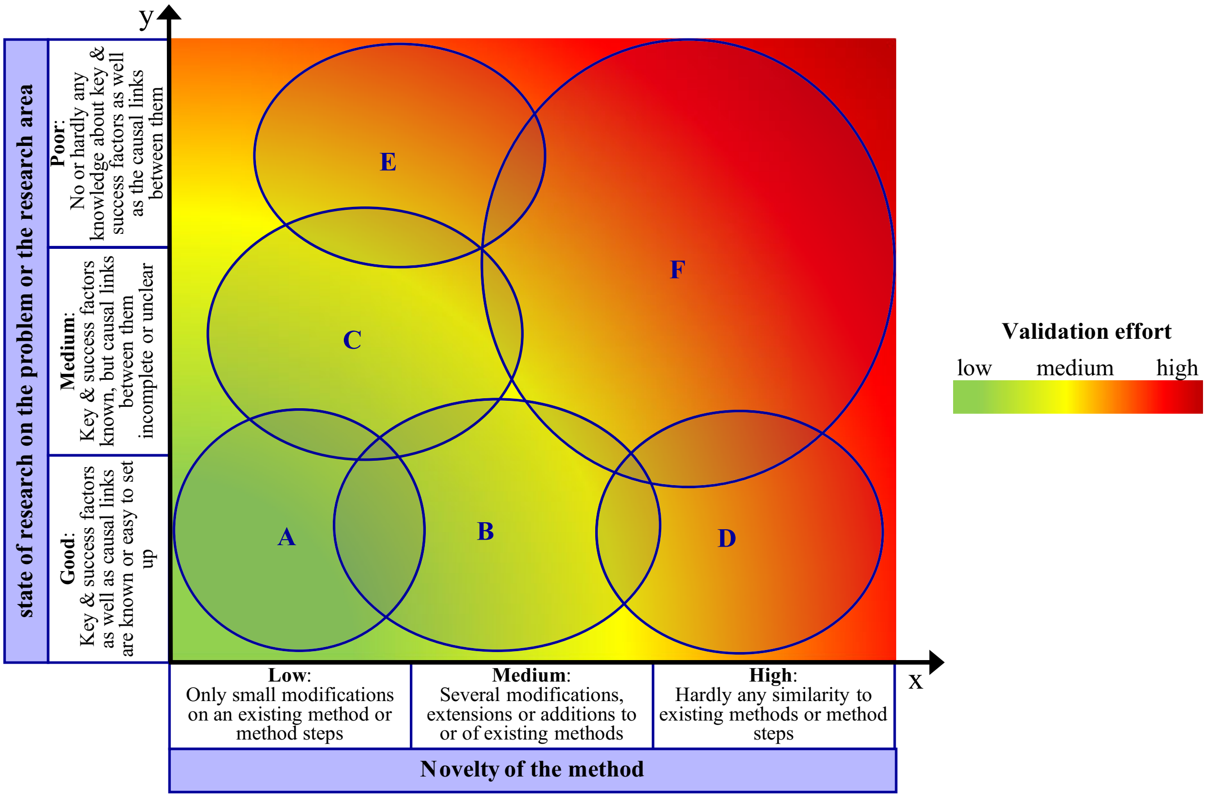

Figure 2. Intensity map of the validation effort

corner, as there is less preliminary work there, resulting in more intensive validation. This effort is illustrated using a colour spectrum from green then yellow to red. Within the intensity map, different scenarios can be distinguished; they are explained in detail in Section 3.2.

The x-axis of the map, 'novelty of the method', is divided into low, medium, and high as follows:

Low novelty means that the design method to be validated differs only slightly from existing, established, and already successfully validated methods. Since the original method is therefore a functional, consistent, and applicable method, this can also be considered to be the case for the method to be validated in the current project; structural validity is therefore already given. An example of such a case would be the transfer and specific adaptation of an existing method to a new industry. For example, the scenario management method (Gausemeier et al., 1998) could be adapted for aviation. In this case, it is sufficient to provide individual industry-specific information in the method step of the scenario field analysis or complementary evaluation criteria for the consequence analysis of the method.

Medium novelty of the method means that several additions or extensions are made to an existing method. This could also mean that these are extended by further method steps or even that two or more methods or tools from different areas are combined to create a new method for a new application. An example of this is a Persona-Scenario method (Madsen & Nielsen, 2010) whose components, the Personas method (Pruitt & Adlin, 2006) and the Storytelling method (Parrish, 2006), also exist as individual methods. The combination, in which personas are first created and then guided through various usage scenarios to identify usability problems, is still quite obvious. Nevertheless, the functionality and applicability of the new method should be validated at least once (empirical structure validation) before extensive experiments with several participants are planned.

High novelty refers to methods that bear little resemblance to existing validated methods and therefore require precise and intensive validation, beginning with the structure validity (theoretical and empirical structure validation). An example of this area could be a completely new creativity method based on the use of novel technologies, such as augmented reality tools.

The y-axis of the intensity map is divided into three areas which are distinguished by the 'state of research on the problem or the research area' to which the design method shall contribute. In other words, this is about how well the research and problem area addressed has already been investigated and how elaborate the theories and metrics are. The areas on the y-axis influence the extent to which proof of efficacy should be provided as part of the effect validation. However, this subdivision in the intensity map does not affect the extent to which applicability and acceptance should be measured quantitatively or qualitatively in the context of effect validation. This depends on the results of the previous structural validation and the probability with which applicability or acceptance problems are expected. In case of doubt, a more extensive data collection, i.e., a combination of qualitative and quantitative research methods, is recommended here.

At the bottom, visualised in the green, is the '**good**' area, which refers to a research field in which there are already many measurable key factors that can be combined into a well-documented impact model with little effort. There may even be complete impact models in reference work that can be used. One example is the area of creativity and the theories that establish the connection between quantity, novelty, variety, quality, usefulness, innovation and product success (Genco et al., 2012; Shah et al., 2003). In the case of a complete impact model, as is the case with a 'good state of research', a single quantitative measurement setup could be sufficient to prove the efficacy and direction of the effect of the method. In the example presented, it would be sufficient to measure the number of ideas based on the new creativity method as part of the controlled laboratory study.

The '**medium'** area refers to a problem for which there are already several measurable key and success factors, but the connection between them, i.e., the links in the impact model, cannot be easily established as there may be a lack of theoretical foundations. This is often the case when two different research areas are brought together, i.e., when, for example, flexibility and capacity utilisation in assembly is to be increased through clever product structuring. In this example, there are already metrics for evaluating a product structure, e.g., by comparing the dependency of the components on each other by comparing the degree of theoretical combinability with the actual combinability (Rapp, 2010). On the other hand, there are also ways of measuring the capacity utilisation and the downtime of an assembly line or cell (Corrado & Mattey, 1997) and estimating its flexibility. How these factors are interrelated and influence each other, i.e., what the causal link is from one to the other, still needs to be investigated further. As a consequence of this gap in the state of research, a purely quantitative study within the framework of empirical effect validation may be difficult to conduct without a good theoretical foundation, see also (Wacker, 2008). Instead, it is recommended to collect complementary qualitative measures to support the formation of causal theories. In the example shown, a mixed methods approach could be pursued. Depending on the size of the gap in the impact model, a further empirical study in a laboratory environment is then required, in which quantitative proof of efficacy is provided.

A **poor** research foundation exists when there is hardly any basis for both key factors and measurable success factors and there is also a lack of theories on causal relationships. An example of such a case could be the improvement of sustainability through suitable product design, which shall go beyond the improvement of material recycling shares and the reduction of CO2 values. This would require further operationalisation of the sustainability goals (United Nations, 2024) and the derivation of the causal relationship between product design and these sustainability aspects. Depending on the scope of the impact model to be developed several iterations of qualitative and quantitative measurement of the efficacy of the key factors and the success factor(s) may be required in the case of a poor research basis. The iteration is complete once proof of causal links up to the success factor(s) has been provided by empirical studies or references.

It is important to mention that the areas within the intensity map are not always sharp and can also be scalable per level of consideration. In case of doubt, a more comprehensive validation should be favoured and any similarities to other methods should be disregarded; i.e., the positioning of one's own method on the intensity map shifts further into the yellow or red area, i.e. to the upper right. Further, the performance validation of a method is in turn necessary in every scenario and must be carried out regardless of its degree of novelty or the state of research. However, its scope, duration, number of repetitions and the degree of proof of external validity depend on the method itself as well as the time, personnel and financial resources of the research project. As a minimum, the first two sub-steps of performance validation should be aimed for, to ensure initial performance validity of the design method.

3.2. Scenario-based planning of design method validation studies

The two dimensions described above create an intensity map in which various scenarios for method validation projects can be distinguished. This results in a scenario-based validation effort, as well as suitable recommendations and focal points for the step-by-step validation of the design method. It can be

summarized that the more novel a method is, the more structural validation and thus proof of functionality and applicability it requires. There is a similar correlation between the state of research on the problem and the scope and type of effect validation required. However, the type and scope of the measurement of applicability and acceptance in the context of effect validation remain largely independent of this project characteristic. The same applies to performance validation. This should always be planned as part of the validation, regardless of the degree of novelty and the state of research. The following six scenarios with the recommended validation process result from the intensity map:
For **scenario A**, in which both the degree of novelty of the method is very low and the state of research is very good, structural validation can be omitted completely. A single quantitative study with proof of efficacy in a laboratory setting for effect validation is sufficient before starting performance validation.
In **scenario B**, validation starts with the empirical structural validation due to the higher degree of novelty. The process then follows the flow of scenario A.
In **scenario C**, the degree of novelty is not necessarily higher than in A, but the theoretical foundation from the state of research is not sufficient. The effect validation should be carried out more thoroughly, e.g. in the form of a mixed methods approach or through iterations, until the gaps in the causal chain of effects in the impact model are closed. The performance validation can then be carried out.
In **scenario D**, the degree of novelty of the method is such that it requires comprehensive and complete structural validation in form of both, theoretical and empirical studies. Subsequently, effect and performance validation take place in a similar way to the scenarios A or B.
Scenario E shows the other extreme compared to scenario D, this time on the y-axis. However, it is unlikely that a fully functional and effective method already exists for a little-researched problem area. As there are only a few operationalised metrics or theoretical foundations for deriving causal links in the impact model, a comprehensive and iterative sequence of empirical studies is recommended here, in which both qualitative and quantitative measurement methods are used to investigate causal links between the method and the problem and to demonstrate the effect of the method on the various key factors. This should be done iteratively until the impact model has been completely built up to the success factors that the method promises to achieve. After that performance validation can be evaluated in the designated context.
In **scenario F**, both extremes from scenarios D and E come together. The validation procedure here is extensive and the number of studies to prove the various validation criteria is correspondingly large. The complete execution of such a validation process in a limited period of time of a single methodological research project becomes increasingly unlikely as the validation effort increases.
There are various validation activities known for carrying out the individual validation steps necessary for the different scenarios. A validation activity is an umbrella term for various methods, models, tools and any instructions that represent a premise or support for the proof of a validation criterion as part of a validation step.
The best-known validation activity is the proof of performance in the form of a case study in a company, while e.g., questionnaires and statistical evaluation procedures are additional research methods that can be used in it. Yee et al. (2013) differentiate in this context between different types validation studies, e.g. retrospective validation or contemporaneous validation, in respect to strength of evidence they provide. Laboratory studies, i.e., those conducted in a controlled study environment using a sample problem, could in turn be assigned to the effect validation step. Pre-tests are often mentioned as preparation and risk reduction before conducting any empirical study. They can be used for the verification of empirical structure validation. According to Pedersen et al. (2000), flowchart-like representations of the method can be created to prove internal consistency and completeness of the method in the theoretical structure validation; more formalised approaches for this are, for example, the genome approach (Birkhofer et al., 2002) or the method and process visualisation according to (Beckmann & Krause, 2013) though both approaches were not originally created for this purpose. No specific activities have been named yet for the theoretical performance validation.

4. Discussion

The proposed approach combines a scenario-based differentiation of the validation procedure with existing validation steps and enables planning of the steps to be carried out and the expected effort. The approach thereby supplements existing validation approaches with an applicable differentiation based on

the individual characteristics of the project. Using the proposed approach, individual validation steps can be logically excluded or reduced depending on the initial situation of the method.
Still, there are two central unknowns with regard to the scenario-based approach: the applicability of the intensity map and the scenario-based recommendations itself, and the theoretical basis for the same. Regarding applicability, it is still to be evaluated whether the descriptions and classifications of the project characteristics are sufficient and complete to be a suitable scenario-based planning support for method researchers in actual research projects. With regard to the theoretical basis for the intensity map, it is unclear whether the six scenarios need to be further differentiated or more clearly distinguished from one another. It is also unclear how often the scenarios in the extremes of the intensity map exist in research. Finally, it cannot be ruled out that other decisive project characteristics open up further scenarios for validation projects. Further, the consolidation of different validation approaches used as a working basis for deriving different scenarios in method validation does not claim to be complete or final. There is not yet a recognised standard for this in the context of design method research, see (Cash et al., 2023). Further development in this regard could also contribute to a deeper understanding of the scenarios.

5. Conclusion

This paper proposed a scenario-based approach for planning design method validation studies and their effort consisting of an intensity map of validation effort and the recommended validation process for six scenarios. The approach was derived from the state of research on design method validation and established validation models. The six different scenarios of method research projects were distinguished based on the novelty of the design method and the state of research on the problem or the research area. For each scenario, recommendations were provided regarding the scope, sequence and iterations of the validation steps and criteria. Through the positioning of projects within the intensity map, researchers can plan validation studies and estimate the required effort, allowing for a better alignment with their individual project characteristics before starting studies. Thereby, there are three core aspects of the proposed approach: (1) validation is a multi-stage process, (2) validation depends on research project characteristics, and (3) the validation scenario is defined by the state of research on the problem and the novelty of the method.
In the future, a goal should be to assess the applicability and adequacy of the scenario-based validation approach to real or more elaborated design method research projects. This could help to evaluate and advance the proposed scenario-based approach. Furthermore, validation activities, i.e., pre-test, laboratory test and case studies, are usually not very formalised and therefore often lack objectivity and reproducibility. Even though efforts are being made to remedy this situation with regard to laboratory studies and effect validation, see (e.g., Eisenmann & Matthiesen, 2022; Üreten et al., 2020), there still appear to be major gaps and a need for support for method research; particularly with regard to performance validation. Future research efforts on design method validation should therefore encompass formalised validation activities for objective and reproducible findings.

Acknowledgments

The research documented in this contribution has been funded by the German Research Foundation (DFG) within the project “MProvE - Validation of Methods in Product Development through Experiments – Development of a Procedural Model” (447357425). The support by the German Research Foundation (DFG) is gratefully acknowledged.

References

Beckmann, G., & Krause, D. (2013). Process visualisation of product family development methods. In *ICED13*, Seoul, Korea.
Birkhofer, H., Kloberdanz, H., Berger, B., & Sauer, T. (2002). Cleaning up design methods-describing methods completely and standardised. In *Design 2002*, Dubrovnik, Croatia.
Blessing, L., & Chakrabarti, A. (2009). *DRM, a Design Research Methodology*. Springer. https://doi.org/10.1007/978-1-84882-587-1
Cash, P. (2018). Developing theory-driven design research. *Design Studies*, 56, 84–119. https://doi.org/10.1016/j.destud.2018.03.002

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

Corrado, C., & Mattey, J. (1997). Capacity Utilization. *Journal of Economic Perspectives*, 11 (1), 151–167. https://doi.org/10.1257/jep.11.1.151

Daalhuizen, J., & Cash, P. (2021). Method content theory: Towards a new understanding of methods in design. *Design Studies*, 75, 101018. https://doi.org/10.1016/j.destud.2021.101018

Eckert, C. M., Clarkson, P. J., & Stacey, M. K. (2003). The Spiral of Applied Research: A Methodological View on Integrated Design Research. In *ICED'03*, Stockholm, Sweden.

Eisenmann, M., & Matthiesen, S. (2022). *Guidance for design method validation: Operationalisation of efficacy* (KIT Scientific Working Papers No. 205).

Gausemeier, J., Fink, A., & Schlake, O. (1998). Scenario Management. *Technological Forecasting and Social Change*, 59 (2), 111–130. https://doi.org/10.1016/S0040-1625(97)00166-2

Genco, N., Hölttä Otto, K., & Seepersad, C. C. (2012). An Experimental Investigation of the Innovation Capabilities of Undergraduate Engineering Students. *Journal of Engineering Education*, 101 (1), 60–81. https://doi.org/10.1002/j.2168-9830.2012.tb00041.x

Gericke, K., Eckert, C. M., & Stacey, M. (2017). What do we need to say about a design method? In *ICED17*, Vancouver, Canada.

Gottfredson, D. C., Cook, T. D., Gardner, F. E. M., Gorman-Smith, D., Howe, G. W., Sandler, I. N., & Zafft, K. M. (2015). Standards of Evidence for Efficacy, Effectiveness, and Scale-up Research in Prevention Science: Next Generation. *Prevention Science : The Official Journal of the Society for Prevention Research*, 16 (7), 893–926. https://doi.org/10.1007/s11121-015-0555-x

Hussy, W., Schreier, M., & Echterhoff, G. (2013). Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor (2., rev.). Springer-Lehrbuch. Springer. https://doi.org/10.1007/978-3-642-34362-9

Isaksson, O., Eckert, C., Panarotto, M., & Malmqvist, J. (2020). You need focus to validate. *Proceedings of the Design Society: DESIGN Conference*, 1, 31–40. https://doi.org/10.1017/dsd.2020.116

Madsen, S., & Nielsen, L. (2010). Exploring Persona-Scenarios - Using Storytelling to Create Design Ideas. In D. Katre, R. Orngreen, & *et al.* (Eds.), *Human work interaction design: Usability in social, cultural and organizational contexts* (pp. 57–66). Springer. https://doi.org/10.1007/978-3-642-11762-6_5

Marxen, L., & Albers, A. (2012). Supporting Validation in the Development of Design Methods. In *DESIGN 2012*, Dubrovnik, Croatia.

Parrish, P. (2006). Design as Storytelling. *TechTrends*, 50 (4), 72–82. https://doi.org/10.1007/s11528-006-0072-7

Pedersen, K., Emblemsvag, J., Allen, J., & Mistree, F. (2000). The "Validation Square" - Validating Design Methods. In *ASME Design Theory and Methodology Conference*, Baltimore, Maryland.

Pruitt, J. S., & Adlin, T. (2006). *The persona lifecycle: Keeping people in mind throughout product design*. Morgan Kaufmann/Elsevier.

Rapp, T. (2010). Produktstrukturierung: Komplexitätsmanagement durch modulare Produktstrukturen und -plattformen (2. Aufl.). *ID-Consult Wissen für die Praxis*. Books on Demand.

Shah, J. J., Smith, S. M., & Vargas-Hernandez, N. (2003). Metrics for measuring ideation effectiveness. *Design Studies*, 24 (2), 111–134. https://doi.org/10.1016/S0142-694X(02)00034-0

United Nations. (2024). *Sustainable Development Goals*. https://sdgs.un.org/goals

Üreten, S., Spallek, J., Üreten, E., & Krause, D. (2020). Validation of the Design for Mass Adaptation Method – A Case for Higher Medical Treatment Quality. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, 9 (1), 88–100. https://doi.org/10.1177/2327857920091059

Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39 (2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Verein Deutscher Ingenieure (November 2021). *Development of mechatronic and cyber-physical systems: Entwicklung mechatronischer und cyber-physischer Systeme (VDI/VDE 2206).*

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. https://doi.org/10.1111/j.1745-493X.2008.00062.x

Wallisch, A., Nicklas, S. J., & Paetzold, K. (2021). Method Accepted - Fields of Action for Increasing Methods Application in Product Development. In *ICED21*, Gothenburg, Sweden.

Yee, P., Fischer, M., & Kam, C. (2013). Prospective validation of virtual design and construction methods. *Journal of Information Technology in Construction (ITcon)*, 18, 214–239.