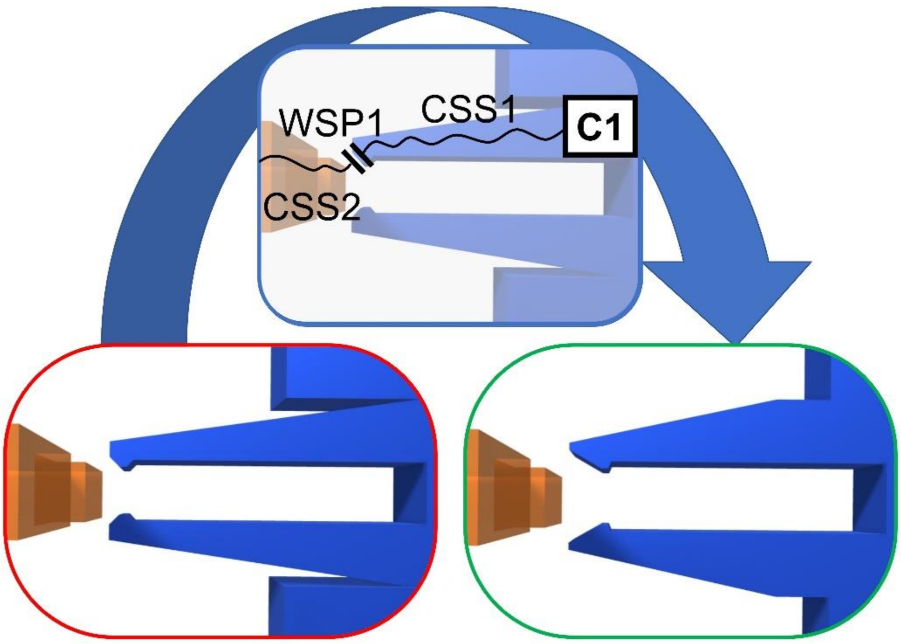
Masterarbeit

Konzeption, Durchführung und Analyse eines Probandenexperiments zur Wirksamkeit der qualitativen Modellbildung mit dem Contact & Channel Ansatz

Design, Implementation and Analysis of a Human Subjects Experiment on the Efficacy of Qualitative Modelling with the Contact & Channel Approach

B.Sc. Lukas Pähler



Betreuer: Univ.-Prof. Dr.-Ing. Sven Matthiesen
Co-Betreuer: Matthias Eisenmann

Nr.: 4397

Karlsruhe, März 2022

Kurzfassung

Diese Arbeit untersuchte die Wirksamkeit der qualitativen Modellbildung mit dem Contact & Channel Ansatz bezüglich der Verbesserung des Systemverständnisses von Probanden mit geringem Vorwissen. Zu diesem Zweck wurde ein Detailprozess zur Entwicklung von quantitativen Studien mit dem Fokus Wirksamkeit angewendet. Die daraus resultierende Forschungsmethode untersuchte die Wirkung qualitativer Modellbildung auf das Systemverständnis von Probanden in einer kontrollierten Testumgebung auf zwei Verständnisebenen: der System- und Detailebene. Die Messung auf der Systemebene erfolgte durch sechs Aufgaben, in denen jeweils das Verhalten von technischen Systemvarianten bestimmt werden musste. Das Verständnis auf der Detailebene wurde durch eine webbasierte Konstruktionsaufgabe abgefragt, in der eine bestimmte Funktionsbeschreibung erfüllt werden musste. Die Forschungsmethode wurde zur Untersuchung der qualitativen Modellbildung mit dem Contact & Channel Ansatz in die ILIAS-Plattform des Karlsruher Instituts für Technologie implementiert. Die Lehre der Modellbildungsmethode erfolgte dabei durch ein Trainingsvideo und eine geführte Modellbildung mit Rückmeldung. Die Studie wurde in Lehrveranstaltungen mit Studierenden im ersten und dritten Bachelorsemester durchgeführt. Dabei wurde auf keiner der Verständnisebenen eine signifikante Verbesserung des Systemverständnisses identifiziert. Auf der Detailebene wurde in dem Durchlauf mit den Drittsemestern eine signifikante Verschlechterung des Systemverständnisses identifiziert. Die Untersuchung der Eignung der Forschungsmethode sprach für die Plausibilität der Messungen, auch wenn weitere Forschung zu den Einflüssen innerhalb der Forschungsmethode und Operationalisierung von Systemverständnis notwendig ist. Die plausiblen Ergebnisse deuteten demnach darauf hin, dass die Verschlechterung auf die Modellbildung zurückzuführen war. Demzufolge muss der Grund für die Verschlechterung untersucht werden, wobei das individuelle Verhalten der Probanden bei der Anwendung und deren Umgang mit der Modellbildung betrachtet werden sollte.

Karlsruher Institut für Technologie

IPEK Institut für Produktentwicklung

Abstract

This thesis investigated the efficacy of qualitative modelling with the Contact & Channel Approach regarding the improvement of system understanding of subjects with no to minimal prior knowledge. For this purpose, a detail process for developing quantitative studies with focus on efficacy was applied. The resulting research method investigated the impact of qualitative modelling on system understanding of subjects in a controlled test environment on two levels of understanding: the system level and the detail level. The measurement on the system level was performed by six tasks, in which the behaviour of technical system variants must be determined for each of them. Understanding on the detail level was assessed by a web-based design task in which a functional description has to be fulfilled. The research method was implemented in the ILIAS-platform of the Karlsruhe Institute of Technology to investigate qualitative modelling with the Contact & Channel Approach. Teaching of the modelling method was done through a training video and guided modelling with feedback. The study was carried out in lectures with students in their first and third bachelor's semester. No significant improvement in system understanding was identified on any level of understanding. On the detail level in the run with the third semesters, a significant decrease in understanding was identified. The investigation on the suitability of the research method showed that the results were plausible, however additional research is required as to the influences within the research method and the operationalisation of system understanding. The plausible results suggested that the decrease was due to the modelling. Hence, the cause of the decrease is to be investigated, which may need to consider the individual behaviour of the test persons and their handling of the qualitative modelling approach.

Contents

1. Introduction

Design methods provide their users with support in design activities for the purpose of creativity, selection of a design variant, understanding a design problem, and many other objectives. The validation of a design method represents a key factor in the development process of the method (Blessing & Chakrabarti, 2009), as well as a contributing factor to the success of the introduction into practice (Jagtap et al., 2014). However, a closer look at the validation practice reveals a wide range of approaches, research methods, metrics, and objectives (Eisenmann, Grauberger, Üreten, et al., 2021). This disagreement increases the development effort for validation activities and reduces the comparability of their results. The disagreement consequently has a negative impact on the development and introduction to practice of methods.

This disagreement also affects the validation of qualitative modelling with the Contact and Channel Approach (C&C²-A). Modelling with the Contact & Channel Approach has been investigated in previous studies with regard to applicability (Eisenmann, Grauberger, & Matthiesen, 2021) and efficacy in terms of system understanding (Grauberger et al., in review). The study on efficacy could only partially validate the positive impact of the modelling method. Due to the novelty of the used research method compared to the established modelling approach and its successful applications (Grauberger et al., 2020), enhancement of the research method is pursued. Hence, this thesis is focused on the design, implementation, and analysis of a human subject experiment on the efficacy of qualitative modelling with the Contact & Channel Approach.

The human subject experiment investigates the impact of the method on system understanding and is developed by using a reference process that builds on the findings of two previous studies. For this purpose, the current state of research in design method validation, validation of modelling with C&C²-A and the reference process is presented. Then, the developed experiment and its implementation are explained. Finally, the results are discussed in respect to their implications for the research method and validation of the design method.

2. Design method validation

As a starting point, a general introduction to terms, approaches and the current practice in design method validation is given. Thereafter, the scope of observation is narrowed down to the validation of the qualitative modelling method under investigation. Lastly, the reference process used in this thesis is described.

A central term that needs to be defined is design method. In this thesis, the definition by Gericke et al. (2017, p. 105) is used, as it separates design methods from other forms of design supports by defining them as:

> "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."

In Table 1, the explanations of the terms design process, guideline, and tool are provided for comparison. As Eisenmann, Grauberger, Üreten, et al. (2021) conclude, tools, processes, and guidelines restrict the user's scope of action with respect to the execution of the task and the interpretation of the results more than design methods.

Table 1. Explanations for terms related to the term design method (based on Gericke et al. (2017, p. 105)).

Term	Explanation
Design process	"In design, (1) A formally specified sequence of activities to be carried out in developing a particular design, or a class of designs, which will often be an application or customization of a methodology to a particular problem. (2) The actual sequence of activities carried out in the development of a design, which may correspond more or less well to any formally specified process."
Guideline	"In design, a statement of what to do when, or what should be the case under particular circumstances. A should only be violated for good reason, with a careful consideration of the consequences."
Tool	"An object, artefact or software that is used to perform some action (for example to produce new design information). Tools might be based on particular methods, guidelines, processes or approaches or can be generic environments that can be used in conjunction with many methods."

The second key term that needs to be defined is method validation. For a general definition of validation, which can be applied to design methods, an IEEE standard can be consulted. The Institute of Electrical and Electronics Engineers (1998, p. 71) defined validation as:

> "Confirmation by examination and provisions of objective evidence that the particular requirements for a specific intended use are fulfilled."

Transferring this definition to the validation of design methods, validation must confirm that the design method fulfils the requirements for a specific intended use by examination and providing objective evidence. For a more precise definition of design method validation, requirements to be met and examination procedures are necessary in addition to this general definition. Comparisons of various approaches to design method validation by Eisenmann, Grauberger, Üreten, et al. (2021) and Frey and Dym (2006) revealed a broad range of requirements and necessary examinations (see Figure 1). The following paragraph presents several such approaches to illustrate the differences, compatibility, and points of overlap. Since this is not a comprehensive listing, refer to the publications by Eisenmann, Grauberger, Üreten, et al. and Frey and Dym for a more detailed discussion.

Figure 1 provides an overview of the stages of the validation approaches, which are shown in the respective upper arrows, as well as the requirements that are evaluated in each stage. Blessing and Chakrabarti (2009) define three stages which build on each other. The first stage, called support evaluation, is aimed at ensuring the internal consistency of the method during the development to enable further evaluation. The second stage, called application evaluation, is aimed at the evaluation of the applicability and usability regarding the directly influenced key factors. The final stage, called success evaluation, is aimed at assessing the usefulness and is to be done in the intended situation of method use. Pedersen et al. (2000) defined a four step process consisting of two qualitative structural validation steps and two quantitative performance validation steps. During these steps, six criteria are examined, which are: (Pedersen et al., 2000, pp. 382–384):

(1) "Accepting the individual constructs constituting the method;"
(2) "Accepting the internal consistency of the way the constructs are put together in the method;"
(3) "Accepting the appropriateness of the example problems that will be used to verify the performance of the method."
(4) "Accepting that the outcome of the method is useful with respect to the initial purpose for some chosen example problem(s);"
(5) "Accepting that the achieved usefulness is linked to applying the method;"
(6) "Accepting that the usefulness of the method is beyond the case studies."

The structural validation steps aim to evaluate the effectiveness of the method, while the performance validation steps aim at evaluating the efficiency. To further clarify the terms, Pedersen et al. (2000, p. 382) describe effectiveness as "whether the method provides design solutions 'correctly'" and efficiency as "whether it provides 'correct' design solutions", while usefulness is the combination of both. In contrast, Olewnik and Lewis (2005) propose a working definition for decision support tools which can nonetheless be applied to design methods. According to this definition, for a method to be valid it must fulfil the following three elements: Be logical, use meaningful, reliable information, and not bias the designer. 'Be logical' is not to be confused with internal consistency but aims at the necessity of the results being consistent with intuition.

Comparing the procedures, Blessing and Chakrabarti (2009) and Pedersen et al. (2000) define different stages which have to be completed evaluating multiple criteria to reach validity in the end. These procedures start with the consistency of the method and proceed towards the usefulness. Olewnik and Lewis (2005) do not define different stages but a group of equivalent elements to fulfil for validity even though they mention the need for a combination of qualitative and quantitative evaluations. The criteria examined are referred to differently, but due to close similarities in content, such as usability and efficiency, they can be regarded as overlapping. Hence, despite some differences of the approaches, they may be considered compatible in some respects, as the combination of criteria or stages is a possible option.

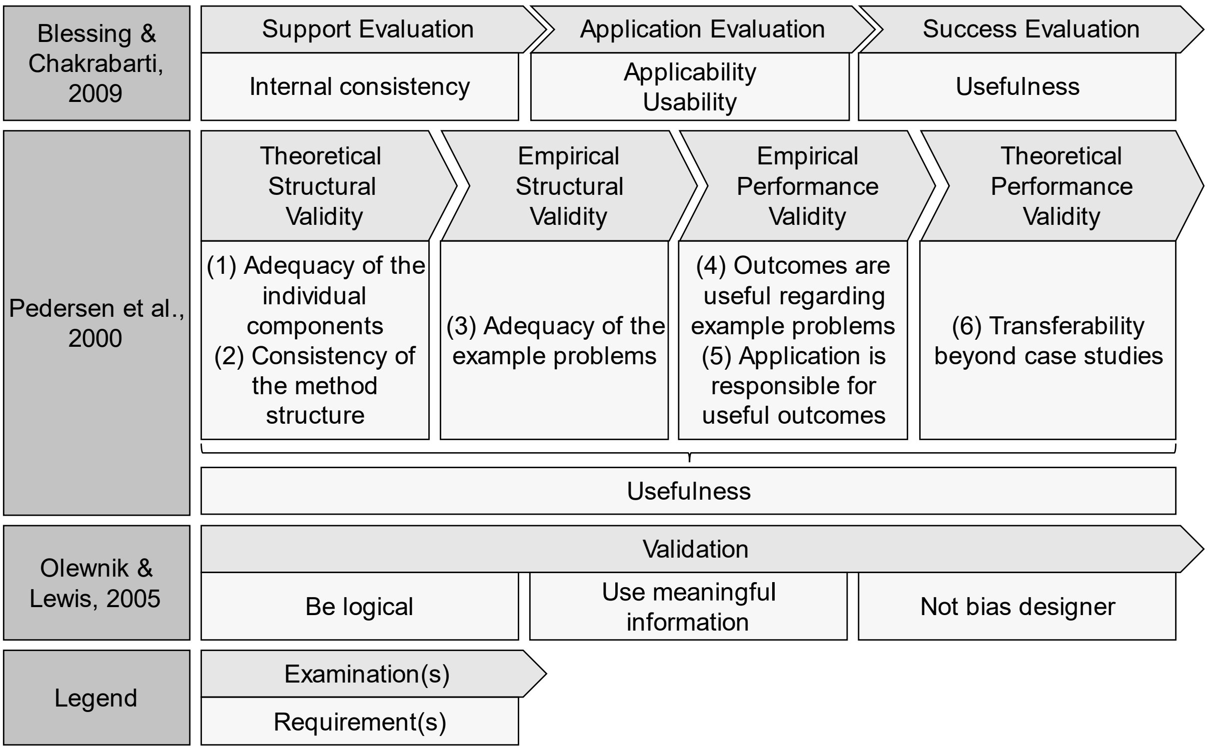


Figure 1. Comparison of the procedures of approaches towards method validation.

To perform design method validation activities, high-level criteria such as applicability or effectiveness must be operationalised, i.e., broken down, to allow observation and measurement (Eisenmann, Grauberger, Üreten, et al., 2021). This can best be explained by an example, like the operationalisation of ideation by Shah et al. (2003). Shah et al. aim at the measurement of ideation effectiveness and therefore propose four metrics: *novelty*, *variety*, *quality*, and *quantity*. Each of this metrics is further defined by Shah et al. and its measurement procedure described to enable the measurement in the study context.

The practice of design method validation also reveals an incoherent situation overall. This situation can be illustrated with the example of studies on selection methods. Reich (2010) visualized the tangled situation of studies on methods for the selection among alternatives in design which is shown in Figure 2. The visualization shows a variety of studies that rely on each other, criticize each other, suggest improvements, or merely reference each other without moving in the direction of a more consensual position regarding the capabilities of the methods.

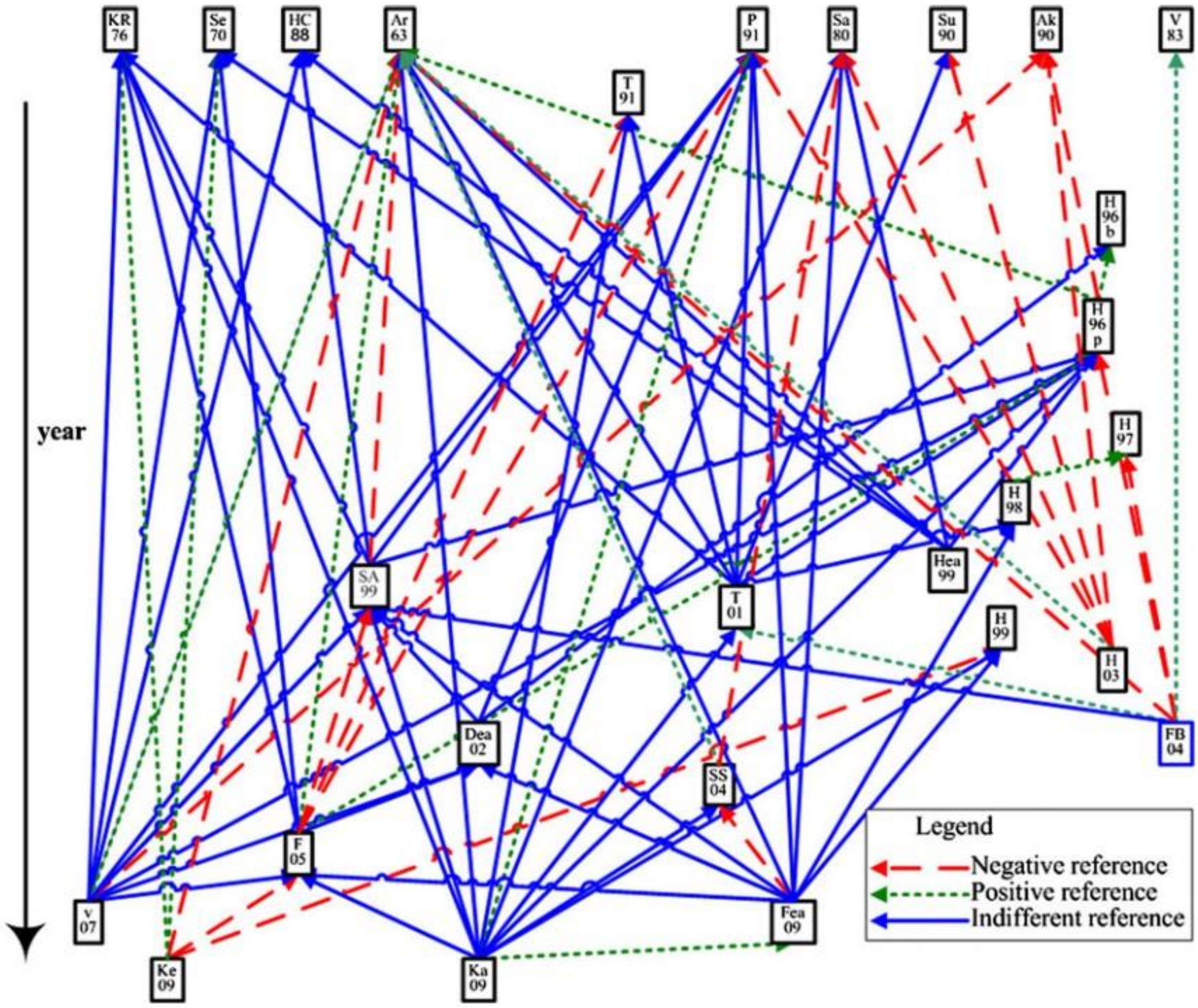


Figure 2. Part of the web of relationships between studies that address team-based selection among alternatives (Reich, 2010, p. 138).

Eisenmann, Grauberger, Üreten, et al. (2021) conducted a systematic literature review to examine how design methods are currently validated and to identify best practice approaches. The identified studies were categorized by evaluation type and levels of evidence. The evaluation type corresponds to the evaluation stages of Blessing and Chakrabarti (2009), support evaluation, application evaluation, and success evaluation. The evidence levels represent the possible internal validity, range from expert opinions to meta-analyses, and were adopted from evidence levels for therapeutic studies from medicine. The identified studies ranged over all categories, except the highest level of evidence, meta-analysis. To reach this level, multiple studies with the same metrics and investigated design method must be conducted. A key finding of the examination of the studies was:

> "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." (Eisenmann, Grauberger, Üreten, et al., 2021, p. 13)

To approach this lack of standardisation, Eisenmann, Grauberger, Üreten, et al. propose two approaches: Defining common goals and metrics and using common tools and models to identify similarities. The studies on ideation methods show the positive effect of common goals and metrics, with eight of 15 identified studies building on the four metrics proposed by Shah et al. (2003). This results in comparability of results which further fosters the development of the design method and metrics. The identification of similarities between different methods aims at a discussion on the direct effects of those elements. Elements can for example be tools or models which help to document knowledge. Connecting this documented knowledge with criteria for successful design enables the targeted use of the elements for methods with the same success criteria.

In summary, the general state of research of design method validation shows a disagreement regarding the necessary examinations for validation activities, requirements to be tested, metrics, and objectives. This disagreement increases the effort of design method validation, as existing research methods can only be used in individual cases. Regarding the metrics, the area of ideation methods shows a possible solution, as similar metrics provide a reusable framework for validation activities.

2.1. Preceding validation of the modelling method based on the C&C²-A.

Two studies are available that validated the qualitative modelling method based on the Contact & Channel Approach. The first study by Eisenmann, Grauberger, and Matthiesen (2021) investigated the applicability and the second study by Grauberger et al. (in review) investigated the efficacy in terms of the impact on system understanding. As a basis for further validation activities on this method, both studies are examined.

Eisenmann, Grauberger, and Matthiesen (2021) investigated applicability using three metrics: *ease of understanding*, *ease of use*, and *extent of being followed correctly*. These metrics were based on the metrics defined by Kroll and Weisbrod (2020) who built their metrics on the framework of Motte and Eriksson (2016). The study was designed as an experiment with double use of the participants (for an example, see Figure 3), with the participants analysing a technical system without the modelling method first, then receiving the introduction to the method and finally analysing another system with the support of the method. The aim of the analysis was to identify the embodiment parameters which influence the function of the system substantially. While the participants analysed the systems, the whole room was recorded on video and a researcher observed each participant. Additionally, all means of documentation were collected and the participants were interviewed afterwards as a group. The metrics were assessed qualitatively based on all collected data. The investigation regarding *ease of understanding* was mainly based on the observation and interviews and resulted in a modification of the introduction. *Ease of use* was assessed based on the documentation and interviews and revealed a critical method step. The *ease of being followed correctly* was investigated based on the documentation and is influenced by the two previous metrics. Three iterations were conducted for optimization purposes, the first with undergraduate students and the other iterations with students with a bachelor's degree in mechanical engineering. Based on the findings and after the optimization, the method met the preconditions for further validation activities.

Grauberger et al. (in review) investigated the impact on system understanding compared to an intuitive procedure. To measure the impact, system understanding was divided based on the levels of function description by Eckert et al. (2011) into two levels: system level and detail level. The system level was measured by the *correct assignment of system variations to corresponding behaviour*. On the detail level, the *correct selection of system states for certain behaviour* and *correct selection of details of the embodiment for a certain behaviour* were used as measurements. As study design, the design of the study on applicability was adopted and the contents adapted to the new objective. The resulting study design is shown in Figure 3. The study design was implemented in an online learning system in which participants solved a group of tasks per metric. The *assignment of system variations* and *selection of system* *states* were designed as multiple-choice questionnaire, while the *selection of details of the embodiment* was performed using image maps. The questionnaires and image maps resulted in test scores based on the number of correct answers which were analysed using non-parametric statistics. The study was conducted in two runs with a total of 36 master's students and research assistants. The results revealed a significant gain of system understanding on the system level, while no statistically significant effect on the detail level was detected.

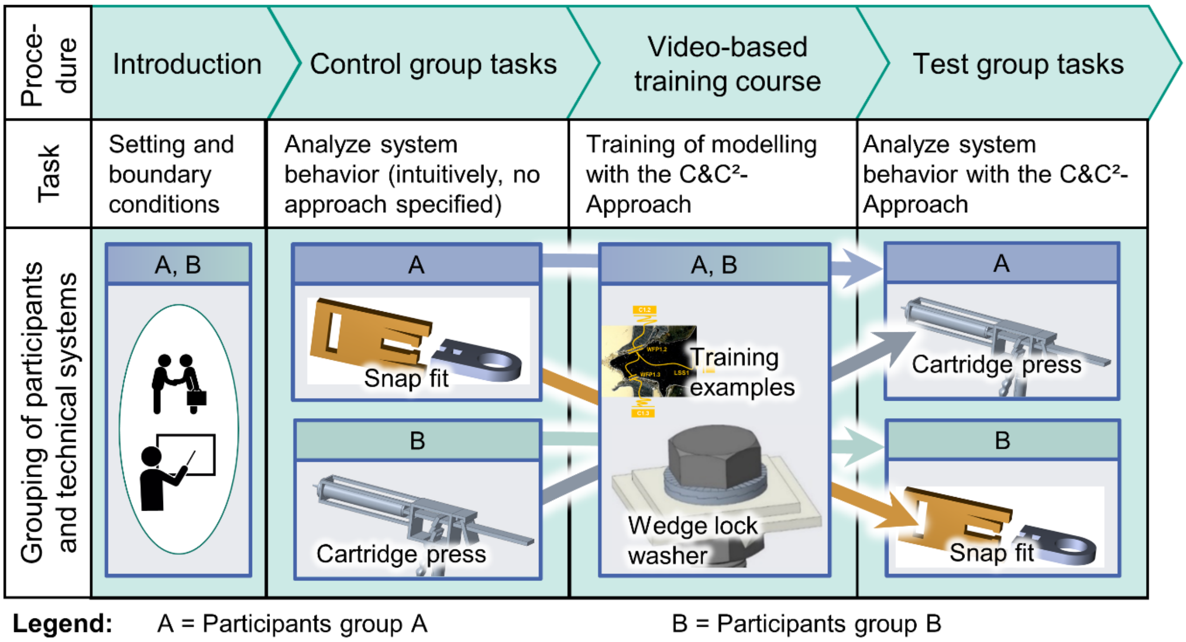


Figure 3. Structure of the experimental study on efficacy (Grauberger et al., in review)

In summary, these two studies show two complementary validation studies with different objectives. The validation of applicability has been successfully completed with subjects with a bachelor's degree. The efficacy study builds directly on the findings of the applicability study, which is supported by similar subject groups. The validation of the efficacy could not be completed based on the results so far.

2.2. Reference process for the validation of design method efficacy

Eisenmann (in progress) developed a reference process for the validation of design method efficacy in human subject experiments. This reference process serves as a support for the development of studies, as existing approaches do not provide sufficiently detailed guidance. Since this thesis is integrated into the reference process, the basic structure is explained in this subchapter. This subchapter is therefore based on Eisenmann, unless another source is quoted.

The validation procedure is composed of three overarching components: theoretical structure validation, empirical validation of efficacy, and empirical validation of effectiveness. The empirical validation of efficacy, which is the focus of the reference process, is further composed of a qualitative study on applicability and a quantitative study on efficacy. The qualitative applicability study aims to identify factors influencing the applicability of the design method in order to optimize the applicability ahead of the quantitative follow-up studies. By developing a quantitative data collection method to measure efficacy, the quantitative study on efficacy builds the foundation for follow-up studies in a more real-world context.

In the course of the validation procedure, the criteria of the validation square by Pedersen et al. (2000) are used as test criteria that must be met. The criteria are listed in Section 2 and located in the validation steps of the validation square in Figure 1 (page 4). The reference process is aimed at the fulfilment of criteria (3), (4), and (5). Criterion (3) checks the suitability of the tasks and is investigated in the qualitative study. Not fulfilling the criterion results in a revision of the research method and must be ensured during both studies. This is particularly important to note when there is a significant change in tasks between qualitative and quantitative study. Criterion (5) checks the applicability of the design method and is investigated in the qualitative study. Like criterion (3), fulfilment must be ensured across both studies. In contrast to criterion (3), failure to fulfil criterion (5) results in a change in the design method. Criterion (4) tests the efficacy and is investigated in the quantitative study. Nonfulfillment results in the modification of the design method. The criteria therefore serve as milestones for the successful completion of the validation stages. As a result of not fulfilling a criterion and new insights gained during the studies, iterative revision of research and design methods may occur. This process is called Co-evolution of design and research method. Figure 4 shows the described validation process and illustrates the Co-evolution.

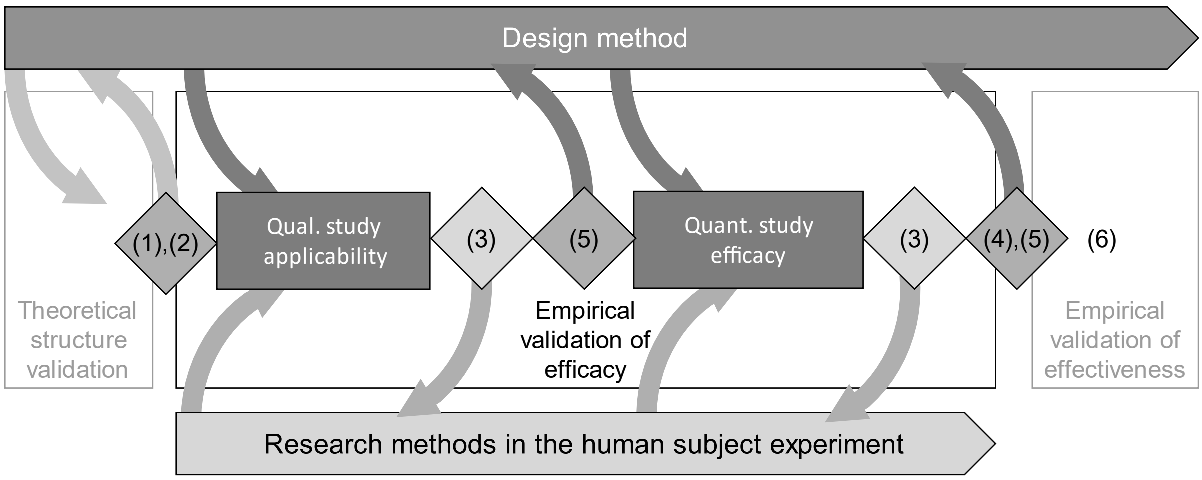


Figure 4. "Reference process of co-evolution of design and research method for method validation in human subject experiments." (Eisenmann, in progress)(translated by the author)

Within the reference process, there are also separate detail processes for the development of the qualitative and quantitative study. These detail processes provide assistance for operationalising the success criteria, conducting the study, and interpreting the study results. The study by Eisenmann, Grauberger, and Matthiesen (2021) was developed using the detail process for qualitative studies with a focus on applicability, and the study by Grauberger et al.

(in review) can be assigned to the detail process for quantitative studies with a focus on efficacy. Since the focus of this thesis is on efficacy, only the corresponding detail process for quantitative studies will be described in further detail.

The detail process uses the findings and study design of the qualitative study for applicability as a starting point and adapts them to the new objective of efficacy and quantitative data collection. This process begins with the operationalisation of the objective and then progresses through the implementation and execution of the study to the review of the criteria for data collection and efficacy. If the criteria are not met, iterations occur in which the design or research method is modified. This process is shown in Figure 5. In the following, the steps and final reviews of the criteria are described in detail.

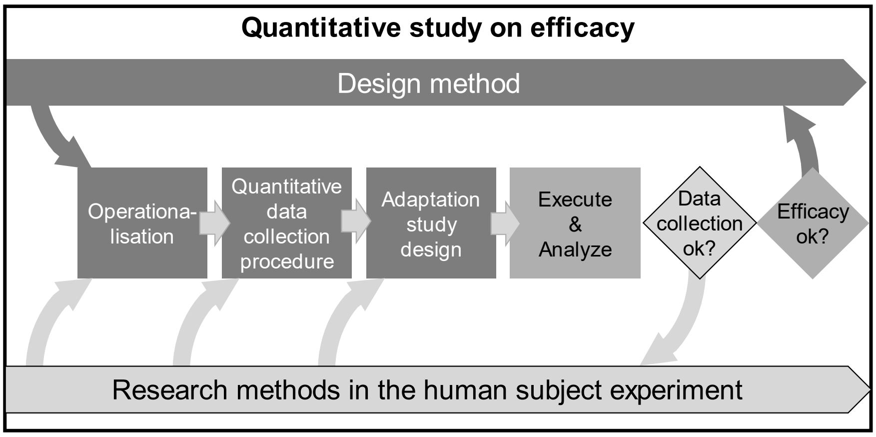


Figure 5. "Detail process for a quantitative study on efficacy" (Eisenmann, in progress) (translated by the author)

Operationalisation. In this step, the definition of metrics to be measured in order to assess the efficacy of the design method is carried out. As a basis for defining metrics, the desired impact of the design method on the designer must be described as precisely as possible. In the subsequent definition of metrics, existing metrics should be used whenever possible. However, since there are no existing metrics in most research areas, it is often necessary to define metrics that are specifically designed to meet the individual objectives of the design method. In this case, metrics should be as general as possible in order to enable follow-up use.

Quantitative data collection procedure. The purpose of this step is to obtain a quantitative data collection procedure for the metrics of the previous step related to the tasks. To achieve this goal, the study design of the qualitative study must be extended to include a quantitative data collection procedure. It is advantageous to reuse means of documentation from the qualitative study and to generate traceable, separate outputs against which the metrics can be measured. As far as possible, the collection of metrics should be automated to increase objectivity and reproducibility.

Adaption of the study design. The study design of the qualitative study should be maintained as far as possible, except for changes due to the requirements of the quantitative study. Keeping the study design allows the use and transfer of experiences from the qualitative study.

Study participants should be inexperienced subjects, since the impact of the design method tends to be stronger due to the lower level of previous experience and is consequently easier to capture. Here, care must be taken to ensure that the selection of participants allows for a number that is sufficiently large for a statistically significant statement regarding the efficacy.

Decision on the suitability of the data collection procedure for efficacy. The decision regarding the suitability of the data collection procedure is focused on its ability to actually measure the defined metrics. In this context, the criteria of objectivity, reliability and validity must be examined individually for each case. To meet these criteria, contributing measures are specified, which are listed in Table 2.

Table 2. Contributing measures to the fulfilment of the criteria objectivity, reliability and validity (based on Eisenmann (in progress)).

Criterion	Measure
Objectivity	• Automation of data collection and analysis • Reduction of the influence of the study supervisor by automating the introduction of the methods • Execution with different groups of participants and comparison of the results
Reliability	• Stability by conducting the study multiple times and comparing the results of the runs
Validity	• Argumentative by including alternative explanations for how the results occurred • Use of multiple different task designs to reduce its influence • Participants are in the control and test group in order to reduce the influence of personality characteristics • Participants are only in the control or test group to reduce learning effects

Decision on the efficacy of the design method. If the suitability of the data collection procedure as well as the suitability of the metrics and application of the design method are considered to be satisfied, the efficacy can be evaluated. The evaluation of efficacy with regard to criterion (4) of the validation square is based on the results of the metrics and provides the basis for a further validation procedure.

2.3. Conclusion on the state of research of design method validation

In conclusion, the state of research on design method validation shows a lack of agreement on theoretical approaches and practical research methods, including the goals and metrics. This results in a high effort to validate methods thoroughly, complicates comparison of results and therefore does not achieve the maximum positive contribution to the development of methods possible. A positive exception is the area of creativity methods, where over several years there has been a tendency towards common goals and metrics that are being used continuously in research practice. Consequently, the clear specification of links to existing research, the definition of method-overarching objectives and metrics, and reuse of research methods should be aimed for to avoid further divergence.

In the case of the validation of the qualitative modelling method based on the C&C²-A, an ongoing process can be observed. After a successful validation of the applicability with mechanical engineering students with bachelor's degree, the validation of efficacy was reached. An initial human subject experiment on efficacy has delivered only partially positive results. Efficacy was assessed in terms of system understanding, which was divided into two sub-levels: the system level and the detail level. A significant effect was identified on the system level, but not on the detail level. Consequently, a revised efficacy validation is required according to the reference process. Due to the novelty of the research method used in the study by Grauberger et al. (in review) compared to the established modelling and successful applications of the modelling approach (Grauberger et al., 2020), the following iteration of the detail process should target the research method.

3. Experimental study on data-driven measurement of system understanding

This chapter presents the progress of the research method and its application to continue the validation of modelling with the C&C²-A. For this purpose, it begins with the relevance of the objective, the improvement of system understanding, and its existing measurement approaches. From the existing measurement approaches, the need for a general research method for methods with the objective system understanding is derived and formulated as a research question. This research question forms the basis for the further investigation of modelling with the C&C²-A based on two research hypotheses.

> This chapter is based on the publication "Data-driven measurement of system understanding using web-based design tasks" (Grauberger et al., in progress). The numbering of the cross references and sections was adapted to this thesis.

3.1. Introduction

Decisions in the early design of a product are often based on uncertain knowledge, and reducing this uncertainty early contributes to the efficiency and effectiveness of a product development project (Verworn et al., 2008). Uncertainty is "the difference between the amount of information required to perform a particular task, and the amount of information already possessed by the organization" (Galbraith, 1973, p. 5). For this reason, the early acquisition and application of knowledge, applicable information, is crucial to reduce inaccurate decisions and their potential negative consequences. This can be achieved through a gain in understanding, as understanding contributes to the expansion of the existing knowledge and is fundamental to the application of knowledge (Royer, 1986). Acquiring an understanding of the product, and by consequence of the system, at an early stage is therefore a key component of an efficient product development project.

Johnson and Satchwell (1992) merged the definitions by Royer (1986) and De Kleer (1984) to define technical system understanding as "[...] the ability to use system knowledge in a meaningful way and to qualitatively reason about three aspects of the system: (1) the structure of the system, (2) the function of the components within the system, and (3) the behavior of those components as they interact with other components in the system." Johnson and Satchwell measured technical system understanding by using standardised tests and the ability to reproduce understanding in the form of diagrams. The standardised tests were subdivided according to three aspects of systems: structure, function, and behaviour. In comparison, Grauberger et al. (in review) measured system understanding on two levels:

system level and detail level. The system level was evaluated by the ability to identify system behaviour, while the detail level was evaluated by the ability to identify function-relevant system states and function-critical details. These studies thereby demonstrate two possibilities for more differentiated measurements of system understanding, compared to the measurements by, for example, Hmelo-Silver and Pfeffer (2004) or White and Frederiksen (1986), which view system understanding as an overall concept. The approaches for the operationalisation of system understanding are examined in more detail in Section 3.1.1.

The standardised tests used by Johnson and Satchwell (1992) revealed a significant impact of the investigated pedagogical method on the overall understanding and the understanding of the behaviour but not on the understanding of the structure of the system and function of the components. In contrast, a replication of the study conducted by Satchwell (1997) found a significant impact on the understanding of the function of components, but not for the other aspects or overall. Grauberger et al. (in review) were able to demonstrate significant impact of the investigated qualitative modelling method only on the system level. In their study procedure, the tasks for measuring the system level were completed by the subjects before the detail level tasks. Due to the time the subjects spent with the technical system for the system level tasks, an influence on the tasks and consequently measurements of the detail level can not be excluded. Consequently, the studies indicate that it is not possible to conclude from an impact on the overall system understanding that all sub-aspects are affected. Additionally, the contrasting results of Johnson and Satchwell's studies and the study procedure used by Grauberger et al. raise questions about the maturity of the measurement methods for independently measuring the sub-aspects of system understanding.

As Forbus (2011, p. 385) states, "understanding mechanical system has been one of the major successes of qualitative modeling, [...]". Qualitative modelling ranges from the casual creation of sketches or symbolic representations (Andreasen et al., 2015) to the methodical generation of models with more precise syntax (Matthiesen et al., 2018). The study by Grauberger et al. (in review) investigated such a methodical approach, more specifically a seven-step qualitative modelling method based on the Contact and Channel Approach. The Contact and Channel Approach is a modelling language that has been successfully applied in a variety of projects (Grauberger et al., 2020) and has shown its efficiency in the context of failure analysis (Gladysz & Albers, 2018). In these applications, the users were mostly experienced in using the method or had acquired general prior knowledge as master students. However, a differentiated validation of the impact on system understanding, especially for subjects with no to minimal prior knowledge, is still missing.

The validation of design methods is fundamental to the improvement of design theory and professional design practice, as the validation guides the development of methods and enables a targeted, situational application (Frey & Dym, 2006). Therefore, a differentiated measurement method is necessary for the development of methods with the objective of improving system understanding. This measurement method would enable a validation of the impact and an optimization of the impact of the method regarding the aspects of system understanding. However, there is no such differentiated research method in the literature. To address this problem, the following research question should be answered:

How can system understanding be operationalised in a differentiated way and how can these metrics be implemented in a study design for a controlled environment?

The resulting study design can then be applied as a sample for the validation of the modelling method based on the C&C²-Approach. So far, validation studies have not been able to prove the impact of the method on the detail level of system understanding and do not include any studies regarding the impact with subjects who have no to minimal prior knowledge. Since these subjects have fewer experiences to fall back on, a stronger impact is expected. To fill these gaps, the following hypotheses are investigated in a study based on the research method with subjects with no to minimal prior knowledge:

H1: *Qualitative modelling with the C&C²-Approach assists in gaining a deeper understanding of technical systems on the system level.*

H2: *Qualitative modelling with the C&C²-Approach assists in gaining a deeper understanding of technical systems on the detail level.*

To investigate these hypotheses, an experimental study was designed to validate the efficacy of qualitative modelling in terms of system understanding. As a basis for the development of the study, Section 3.1.1 provides the state of research on method validation and operationalisation of system understanding. The study design is then applied to investigate qualitative modelling with the C&C²-Approach. Section 3.1.2 therefore introduces the modelling method.

3.1.1. Method validation in design research

As a basis for the development of the study design, this section outlines the state of research on the design method validation process and the intended position of the study within that process. Furthermore, approaches to operationalising of system understanding are reviewed.

The examination and comparison of different theoretical validation approaches revealed a divergence regarding the necessary investigations, their focus and their contents (e.g. Frey and Dym (2006); Eisenmann, Grauberger, Üreten, et al. (2021)). This can be illustrated by the following examples: Blessing and Chakrabarti (2009) described three types of studies with different validation goals, referred to as support, application, and success evaluation, Pedersen et al. (2000) described the combination of a qualitative structural and a quantitative performance validation to assess six criteria, and Olewnik and Lewis (2005) suggested three criteria specifically for design decision methods. This diversity was also reflected in a literature review of validation studies conducted by Eisenmann, Grauberger, Üreten, et al. (2021), which revealed a range of target objectives, metrics, and study designs.

Since there is no established validation procedure whose application was established, this work is built on the compatible concepts by Blessing and Chakrabarti (2009) and Pedersen et al. (2000). The aim of the study design is to investigate the impact of a design method on the system understanding of subjects in a controlled environment. This study design can therefore be classified as an application evaluation according to Blessing and Chakrabarti, since it investigates whether the directly influenced factor, system understanding, is addressed in the intended way. In the context of the validation square by Pedersen et al., this study design is a quantitative performance validation to investigate criterion 4, which evaluates whether "the outcome of the method is useful with respect to the initial purpose for some chosen example problem(s)" (Pedersen et al., 2000, p. 384).

For an investigation of this type, the applicability of the design method must be ensured. This condition was fulfilled for the modelling method based on the C&C²-Approach by a study by Eisenmann, Grauberger, and Matthiesen (2021).

The approaches by Blessing and Chakrabarti (2009) and Pedersen et al. (2000) do not include detailed instructions for the setup of a corresponding study or the derivation of metrics. However, the lack of standardisation in metrics is known as a crucial issue in the development and comparison of validation studies (Eisenmann, Grauberger, Üreten, et al., 2021). Therefore, the approaches for the operationalisation of system understanding in the current literature are listed below:

I. White and Frederiksen (1986) assessed system understanding regarding electrical circuits by requiring subjects to explain the behaviour of circuits while the states of components within were manipulated. Manipulating states included, for example, flipping switches or increasing the voltage in the circuit. The effect of the manipulation on a light bulb was considered as a behaviour.
II. Johnson and Satchwell (1992) and Satchwell (1997) divided system understanding into the elements of structure, function, and behaviour. These elements were addressed by specific questions in exams and the answers were evaluated. Additionally, they assessed the ability to reproduce understanding in the form of diagrams. The electrical systems and subsystems of a small aircraft were used as the object of investigation.

III. Hmelo-Silver and Pfeffer (2004) assessed system understanding by examining thinking out loud protocols of interviews. For this purpose, the protocols were analysed with a structure-behaviour-function coding scheme. The interviews started with drawing an aquarium while thinking out loud and moved into questions, and problem-solving tasks.

IV. Grauberger et al. (in review) distinguished a system and detail level of system understanding and measured them separately. The system level was measured by identifying system behaviour, and the detail level by identifying function-relevant system states and design details.

A comparison of the approaches reveals overlapping as all of them consider system behaviour. On top of system behaviour, most approaches look at other aspects, such as function and structure, to achieve a broader range of measurements. For measurement methods, mostly exam-like approaches such as questions or tests were chosen, except for the coding of thinking out loud protocols of approach III.

3.1.2. The Contact and Channel Approach

For easier understanding of this contribution, a detailed description of the C&C²-Approach is given. This section is based on Matthiesen et al. (2019). Parts of the following text are taken from that paper without changes.

The C&C²-Approach is a thinking tool for embodiment design. It aims to support design engineers in recognizing function-related parameters of the embodiment. As a meta-model it contains elements and rules to build up explicit C&C²-Models. It consists of three key elements and three basic hypotheses that define the usage of its key elements. An overview of the three key elements Working Surface Pair (WSP), Channel and Support Structure (CSS) and Connector (C) is depicted in Figure 6 (left side). A WSP describes the interface where parts of the system connect while it fulfils its function. The CSS goes through system parts and connects the WSP. A CSS can include parts of components or whole subsystems depending on the modelling purpose. The Cs represent a model of the surrounding systems and transmit influences from outside the system boundaries into the system (Gladysz & Albers, 2018). The basic hypotheses describe the possibilities and boundaries of the modelling with the C&C²-Approach. They are depicted in Figure 6 (right side).

A C&C²-Model (Figure 6, centre) is derived by using the key elements and basic hypotheses. For modelling state-dependent embodiment-function-relations, the C&C²-Sequence model is used, where the created C&C²-Models are structured according to their temporal sequence and also different levels of detail can be considered. (Matthiesen et al., 2019)

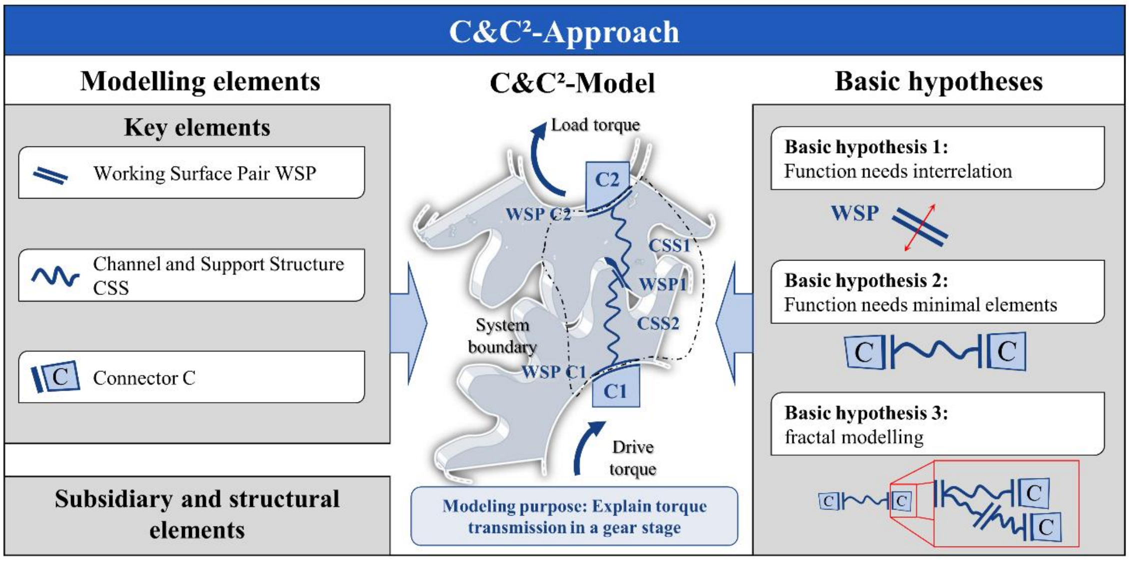


Figure 6. The C&C²-Approach according to Matthiesen, Grauberger, Sturm, and Steck (2018)

A modelling method emerged almost two decades after the model was introduced. This indicates that the need for modelling methods might often not be considered by the model's developers, especially if the model itself seems lean and easy to use. This method differentiates the modelling according to the activities in embodiment design that are addressed. In the activity of analysis, a seven-step method is described. This method is used in the majority of modelling tasks with the C&C²-Approach. It is therefore in focus of this investigation. The four-step modelling method for synthesis remains for further investigations. Figure 7 shows an overview of the modelling method for analysis.

First, the purpose of the model is noted to comprehend the valid scope of this model. Since each model represents only a section of reality, the C&C²-Model is defined in its dimensions of space and time. Then an appropriate depiction of the system is identified, in which the interactions of its components in the function fulfilment are recognizable. All function-relevant energy, material, and information flows in the analysed system pass through the embodiment function elements. By tracking the flow of system variables that is done in step d) (Figure 7, centre), unknown embodiment function elements can be identified.

The identified embodiment function elements are integrated into the created representation of the system under consideration of the basic hypotheses. In the next step, functionally relevant embodiment parameters (characteristics and properties) are identified in the embodiment function elements and their relevance for function fulfilment is formulated.

At the end of the modelling, verification of the model is necessary to check whether the model correctly depicts the relations. The model building in the analysis is completed when the embodiment-function-relation is understood sufficiently, i.e., when characteristics in the embodiment are identified and the model is verified. Then a synthesis can be started in which these understood embodiment-function-relations are used to develop an embodiment capable of fulfilling the function.

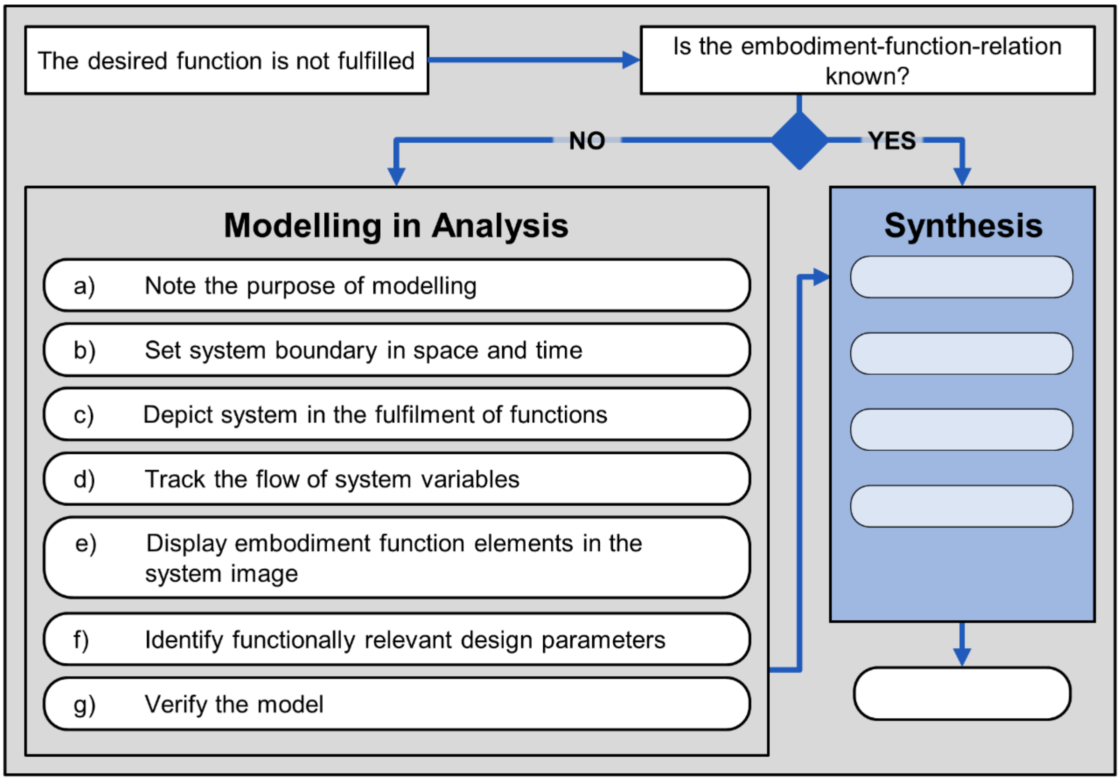


Figure 7. The modelling method of the Contact and Channel Approach translated from Matthiesen et al. (2018)

3.2. Materials and Methods

This section presents the developed experimental study to investigate the impact of qualitative modelling on subjects' system understanding in a controlled environment. The study is first described for a general investigation of qualitative modelling. This is followed by a description of the training course and participants with which the study design is applied to investigate the hypotheses regarding qualitative modelling with the C&C²-Approach.

3.2.1. Operationalisation and measurement of variables

Figure 8 provides an overview of the hypotheses, their segmentation, and the corresponding variables. The general research hypothesis under investigation is:

Qualitative modelling assists in gaining a deeper understanding of a system and its details.

This general research hypothesis must be segmented because of the need for a differentiated measurement, as identified in the introduction (see Section 3.1). As a result of the positive findings by Grauberger et al. (in review) in the context of qualitative modelling on the system level and the possibility to compare results, the segmentation into two general sub-hypotheses for two levels of system understanding, the system level and detail level, is adopted. The system level covers how the system behaves without looking at the reason for this behaviour. The detail level contains the influence of design parameters on the fulfilment of the function. The independent variable across both levels of understanding, and therefore across the general sub-hypotheses, is the training. The training consists of a video-based training and guided modelling. The dependent variable is system understanding on the different levels. For the operationalisation of the dependent variable, measurable variables are derived from the hypotheses based on the state of research.

Research hypothesis		**Qualitative modelling assists in gaining a deeper understanding of a system and its details**	
Measurable sub-hypotheses	System level	How does the system behave?	
		GH1: Qualitative modelling assists in gaining a deeper understanding of technical systems on the system level.	
		IV	Video-based training and guided modelling
		DV	Correct assignment of system variations to corresponding behaviour
		Test	Wilcoxon test, Mann Whitney U test
	Detail level	How do the design parameters influence the function?	
		GH2: Qualitative modelling assists in gaining a deeper understanding of technical systems on the detail level.	
		IV	Video-based training and guided modelling
		DV	Number of corrected function -restricting faults
		Test	Wilcoxon test, Mann Whitney U test

Legend:
GH1-GH2 = General sub-hypotheses DV = Dependent variable IV = Independent variable

Figure 8. Overview of the hypotheses, their variables, and statistical tests. Refinement of (Grauberger et al., in review).

The operationalisation on the system level is adopted from Grauberger et al. (in review). They operationalise the system level by assessing the ability to assign the correct behaviour to a given system variant. This exclusively measures the extent to which the behaviour of the system in a given operation can be predicted, not the reason for the behaviour. Subjects are given six tasks, each with one system variant and four possible system behaviours. One point is awarded for each correctly assigned behaviour, resulting in a maximum of six points. No feedback is given. The resulting variable is metrically scaled, allowing statistical evaluation with the Mann-Whitney U test or Wilcoxon test.

On the detail level, the number of corrected function-restricting faults by modifying design parameters is used as operationalisation. The subjects are given an initial configuration of a technical system with four faults restricting its main functions in a web-based CAD configurator. The web-based CAD configurator is a website on which the participants can manipulate the design parameters of a 3D model and view their adjustment live. In addition to the initial configuration, they receive the description of the requested function in the form of instructions with a drawing and sliders to change the design parameters. Each slider controls the value of a design parameter. To illustrate this more clearly, the web-based configurator is explained in more detail in the following section. It is evaluated how many of the faults restricting the functions are corrected by the subjects by modifying the sliders. The calculation formulas are listed in the appendix. The faults of the initial configurations of the technical systems are shown in Figure 9. One point is awarded per corrected fault, resulting in a maximum of four points. No feedback is given during editing. The resulting variable is metrically scaled, allowing statistical evaluation with the Mann-Whitney U test or Wilcoxon test. The analysis, including the selection of statistical tests, is described in more detail in Section 3.2.3.

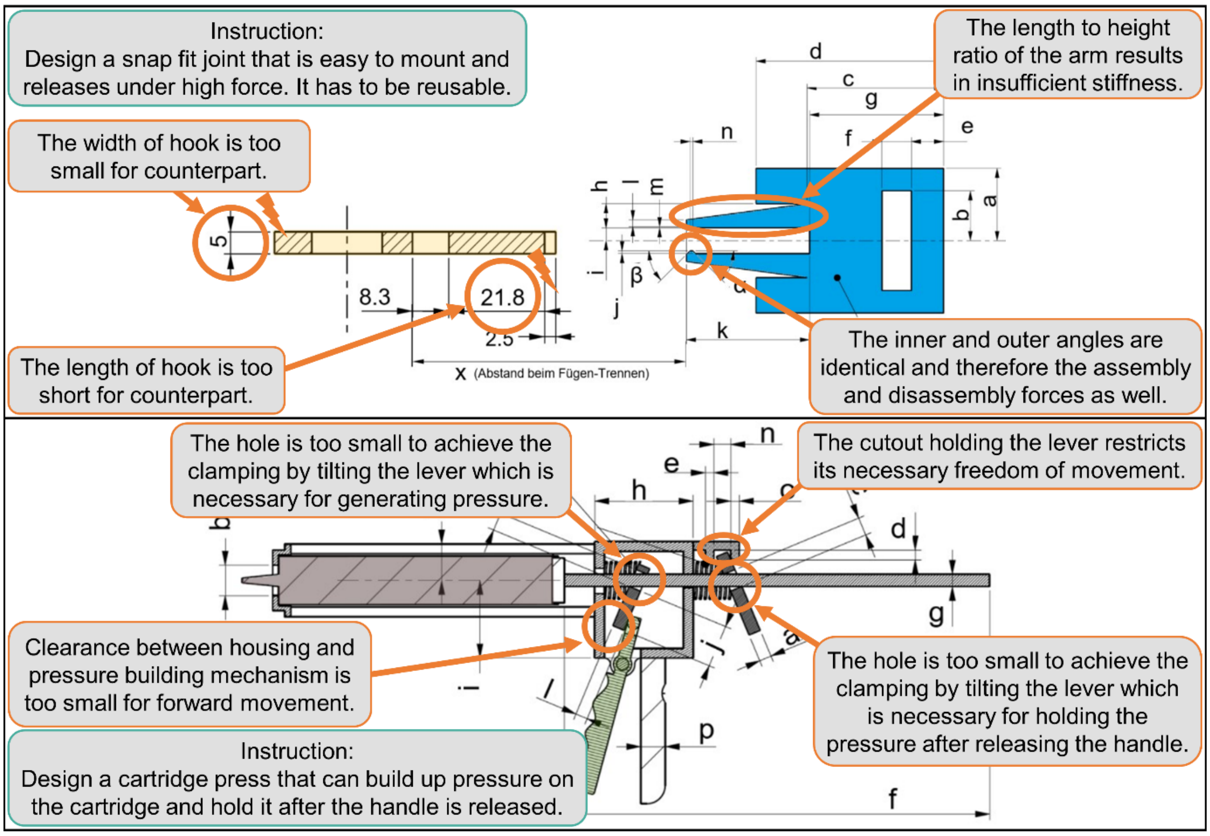


Figure 9. Faults restricting the main functions of the technical systems in its initial configurations with the functional requirements.

3.2.2. Investigation setup and procedure

An overview of the investigation setup including the assignment of the technical systems and subjects is shown in Figure 10. Due to the adoption of the segmentation of system understanding, the investigation setup by Grauberger et al. (in review) is adopted as well. This allows to build on the experience and creates comparability. In this investigation setup, every participant is put once in the control group and once in the test group after the training on qualitative modelling. This investigation setup allows for a doubling of the data collected per subject. The control group preceded the test group as permanent effects of the training on qualitative modelling are expected. To minimize learning effects between the measurements, two different technical systems are used, a snap fit joint and cartridge press. In the previous studies by Eisenmann, Grauberger, and Matthiesen (2021) and Grauberger et al. (in review), these systems showed a manageable complexity and they contain challenging embodiment function relations.

The procedure starts with the web-based design for the detail level measurement and then continues with the system level tasks. The sequence thus exchanges the order of the understanding level tasks compared to the study by Grauberger et al. in order to exclude any influence of the system level tasks on the detail level tasks. Group A starts the control group tasks with the snap fit joint, while group B starts with the cartridge press. This is followed by the training course on qualitative modelling. This course includes a video-based introduction and guided modelling with a wedge lock washer for both groups. Afterwards, both groups redo the tasks with the opposite technical systems while using qualitative modelling. The control and test group tasks are each designed to take 30 minutes to complete.

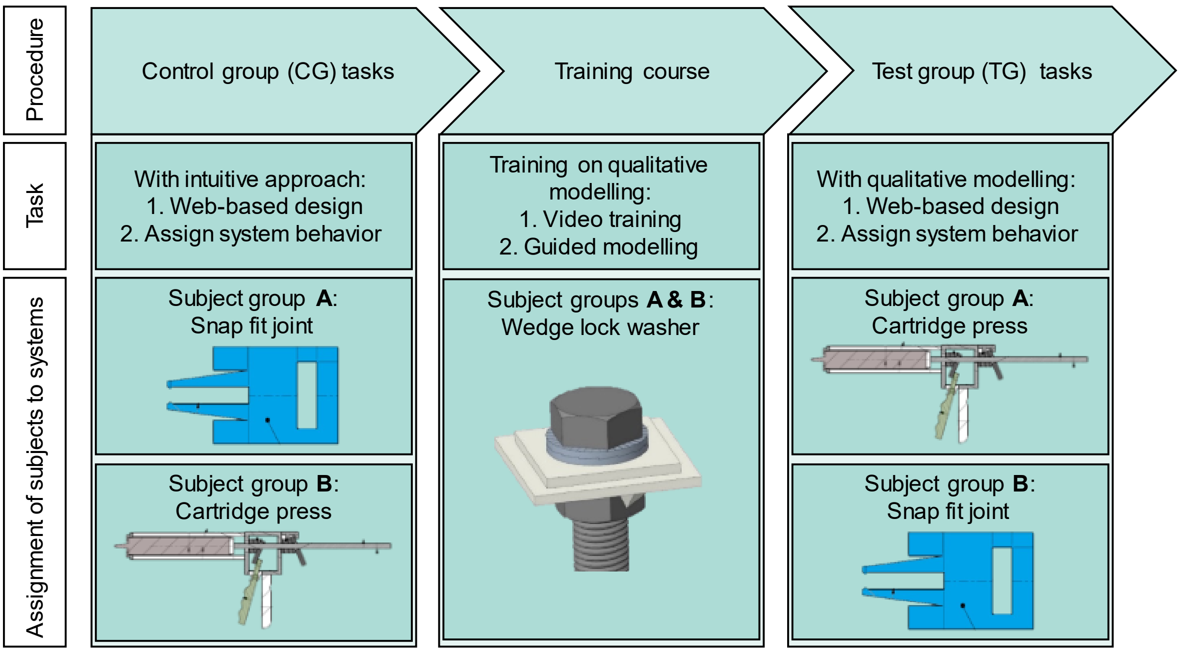


Figure 10. Investigation setup (based on Grauberger et al. (in review)).

Detail level task

The participants begin the task packages with a web-based CAD design task. In this task, the subjects are given a web-based configurator with an initial configuration of the snap fit joint or cartridge press. In addition, a set of sliders for changing the configuration, a functional description and a sketch with the design parameters are given. The task for the subjects is to fulfil the functional requirement by changing the sliders.

Figure 11 shows, as an example, the configurator of the snap fit joint with additional labels for better comprehensibility. On the left side, the design parameters are listed with the currently selected values and sliders. On the right side, from top to bottom, is a preview of the 3D model, the functional requirement and dimensioned sketch. The task of the cartridge press is built similar from the same elements.

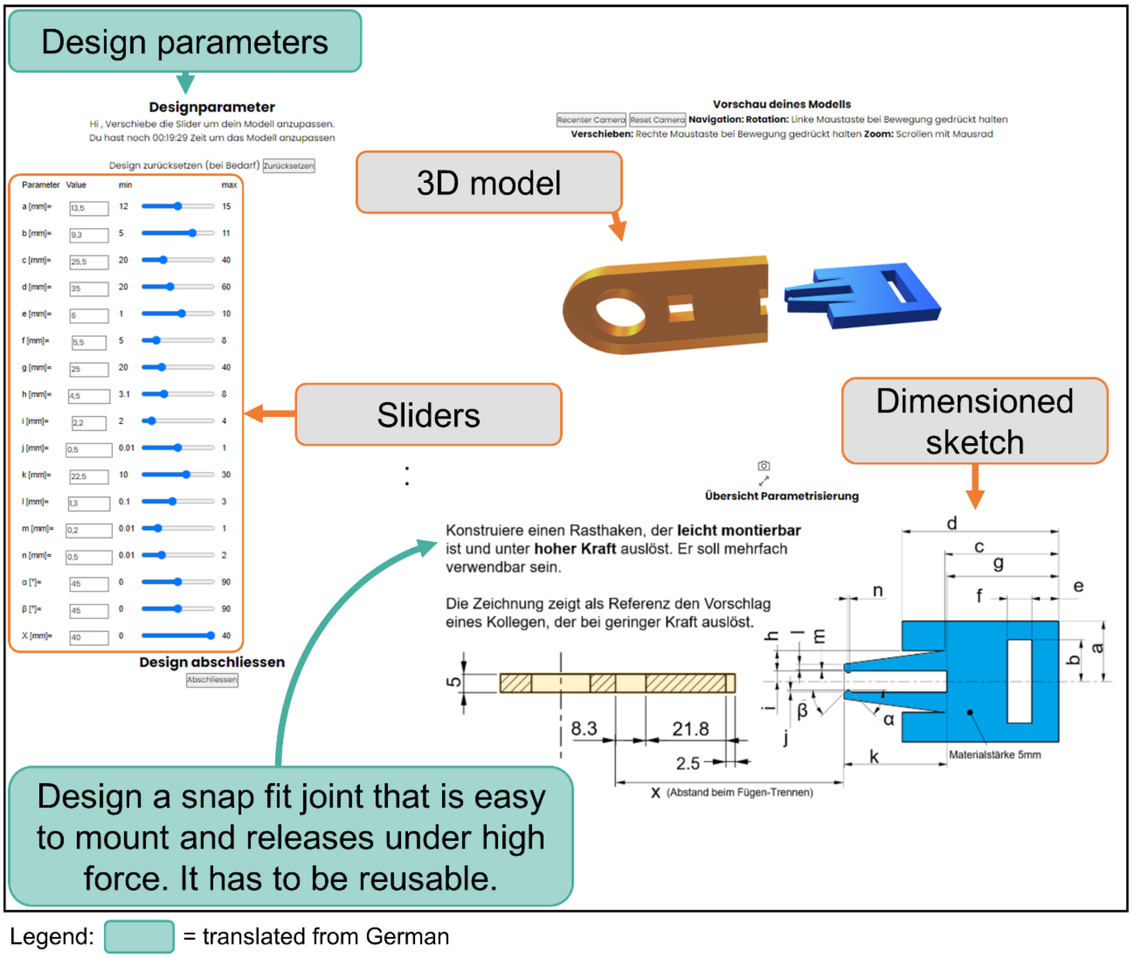


Figure 11. Web-based design task for the snap fit joint.

System level task

As already mentioned, the task on the system level is adopted from Grauberger et al. (in review). Only the representation is adapted based on the feedback in a preliminary study. The task consists of the evaluation of the system behaviour of six different system variations of the snap fit joint or cartridge press. The same four system behaviours per technical system are available for selection for all respective system variants.

The task is illustrated by an example system of the snap fit joint in Figure 12. The task description is written on the top. Below is the picture of the system, the selectable system behaviours, and the selection boxes, one after the other. The selectable system behaviours are characterized by a description, force arrows and force specifications. The task of the cartridge press is built similar from the same elements.

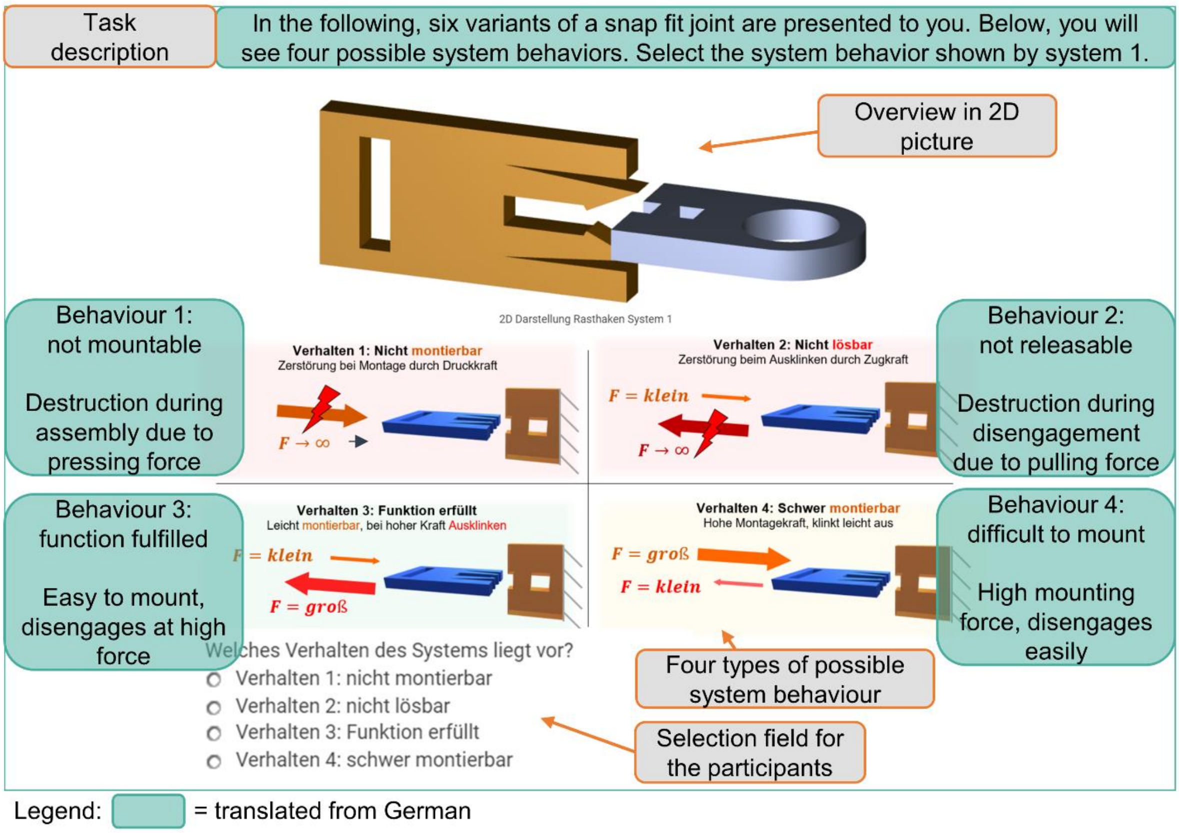


Figure 12. Example of a system behaviour assignment task (based on Grauberger et al. (in review)).

3.2.3. Data analysis

As a basis for the data analysis, the raw data must be pre-processed. In pre-processing, possible reasons for the exclusion of data sets are checked. The reasons for exclusion and checking procedures are listed in Table 3. In addition, the submitted configurations of the web-based design tasks must be assigned to the individual subjects, as the configurators of the technical systems use different files.

Table 3. Reasons for the exclusion of data sets from the data analysis.

Reasons for exclusion	Checking procedure
Non-final data sets	If more than one data set was created by a single subject, all data sets except the last one is excluded.
Unfinished data sets	Data sets in which at least one task was not completed are excluded.
Method not applied	If not all processing steps of the method were completed, the data set is excluded.
Name not linkable	If no identifiable name was entered in the web-based design task, the data set is excluded.
Invalid input	Inputs outside the value range of the model lead to exclusion.
Note by subject	Remarks about personal or technical problems lead to exclusion.

The results of the system level tasks are evaluated directly by the number of points reached. In contrast, for the web-based design tasks, the number of corrected faults restricting the main function of the system is determined using the design parameter data sets. For this purpose, the requirements are expressed as a function of the design parameters. This results in 12 out of 16 design parameters being used to evaluate the fulfilment of the function for the snap fit joint and 10 out of 16 for the cartridge press. The remaining design parameters are intended to make the identification of function-relevant parameters more difficult. After the assessment is performed, the scores achieved in the four tasks are available for each subject.

Non-parametric tests are selected for further analysis of the data due to the assumed non-existent normal distribution. Due to the identical scale levels, system level and detail level can be evaluated identically. The Mann-Whitney U test is used for independent samples, which is the case for the comparison of groups, runs, and difficulties of the technical systems. The Wilcoxon test is used for dependent samples, which is the case for the differences between the control and test groups, since these are taken from the same subjects. The chosen significance level is $p = 0.05$.

3.2.4. Application of the study design to investigate qualitative modelling with the C&C²-Approach

To investigate qualitative modelling with the C&C²-Approach, the study design is implemented in the ILIAS-platform[1] of the Karlsruhe Institute of Technology to enable reproducible experiments with large numbers of subjects. The web-based configurator for the detail level tasks is accessed via a link from the ILIAS-platform and the results are collected in separate tables. The system level tasks are performed directly on the ILIAS-platform. The study is conducted as part of teaching courses and is designed to last 90 minutes. Participants are randomly divided into the groups A and B and assigned to separate virtual Zoom meeting rooms. After the group assignment, subjects are able to enter the guided task procedure without further introduction. All tasks are provided with the necessary instructions on the ILIAS-platform, and help is provided via Zoom meetings if there are any problems with the instructions or the technology. The training course and subjects are presented below.

3.2.4.1. Video-based training and guided modelling

The training teaches modelling with the C&C²-Approach through a video introduction and guided modelling using a wedge lock washer. The training is implemented in the ILIAS platform, independent of the instructor, allowing participants to work at their own personal pace. In addition, the subjects receive a handout with the key elements of the model, the steps of the method, and an example. The training is adopted from the study by Grauberger et al. (in review).

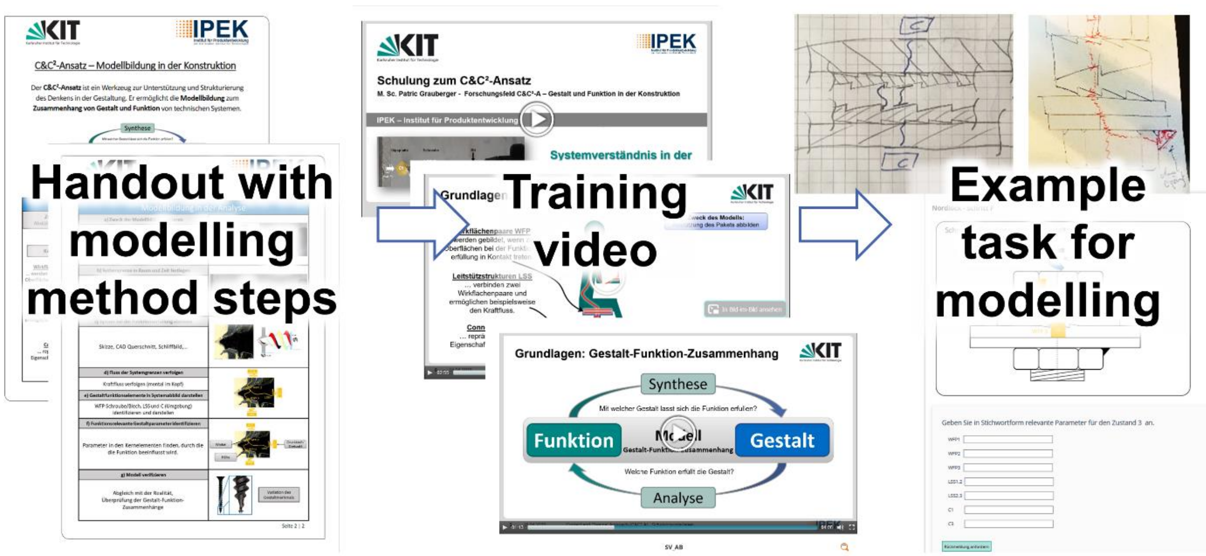


Figure 13. Overview of the training in modelling with the C&C²-Approach (Grauberger et al., in review).

[1] For more information on the ILIAS-platform, visit https://www.ilias.de/en/

The C&C²-Approach is introduced illustratively using a person carrying a package. The example is used to illustrate the key elements of the modelling approach. Subsequently, the sequence of the modelling steps is briefly introduced. In total, the video introduction lasts about 5 minutes. Afterwards, the participants receive a handout. Then the guided modelling begins, in which each modelling method step is performed with automated feedback on the entered solution by providing an example solution. An overview of the training is shown in Figure 13.

3.2.4.2. Participants

The study was conducted in two runs with a total of 306 participants without expert knowledge in the C&C²-Approach at the Karlsruhe Institute of Technology (KIT) in the context of the course mechanical design. In the first run, 159 students in their first semester of chemical engineering and process engineering participated. The second run was conducted with 147 students in their third semester of a mechanical engineering bachelor. In both runs, the execution took place approximately in the middle of the semester and the subjects have already heard about the C&C²-Approach in the courses beforehand. However, the subjects have not received a detailed introduction to the application of the approach.

3.3. Results

This chapter presents the results of the experimental study. As a foundation for the analysis, the pre-processing of the raw data is presented first. After that, the results of the applied study design, the results on the system level (H1) and detail level (H2) are presented.

3.3.1. Pre-processing of the raw data

As a starting point for further statistical analysis, the raw data must be pre-processed. Due to the separate data sets, the results of the system level in the ILIAS-platform and the results of the detail level in separate tables, there are two initial data pools. The pre-processing which is based on the reasons for exclusion (see Table 3) is shown in Figure 14.

The data set of the system level data is sorted by participants and contains 159 participants of run 1 and 147 participants of run 2 at the beginning. After the reasons for exclusion are applied, 111 participants from run 1 and 88 participants from run 2 remain for the analysis on the system level. This corresponds to a data loss of 35%.

223 configurations of the web-based design task were submitted in run 1 and 214 in run 2. The application of the exclusion reasons, except for "method not applied", results in 201 configurations for run 1 and 186 configurations for run 2. Matching these configurations to the individual subjects results in 72 subjects in run 1 and 64 subjects in run 2. After checking the application of method, 71 subjects from the first run and 58 subjects from the second run remain for the investigation on the detail level. This corresponds to a data loss of 41%.

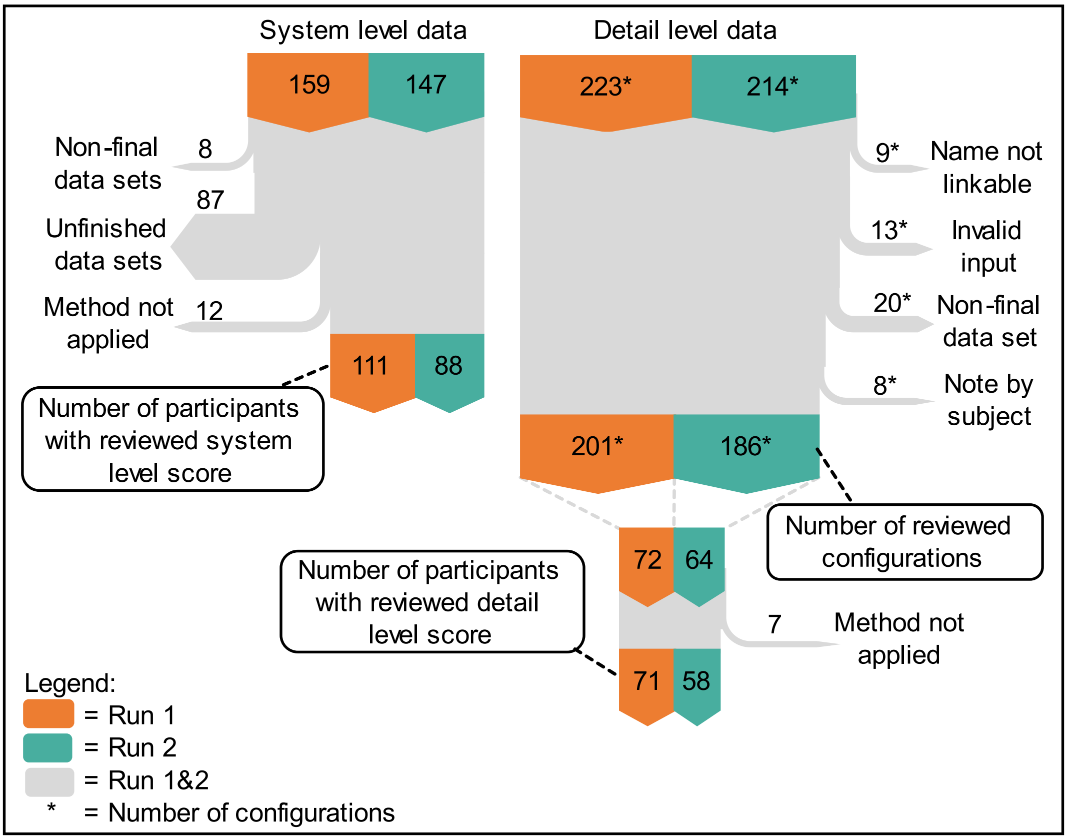
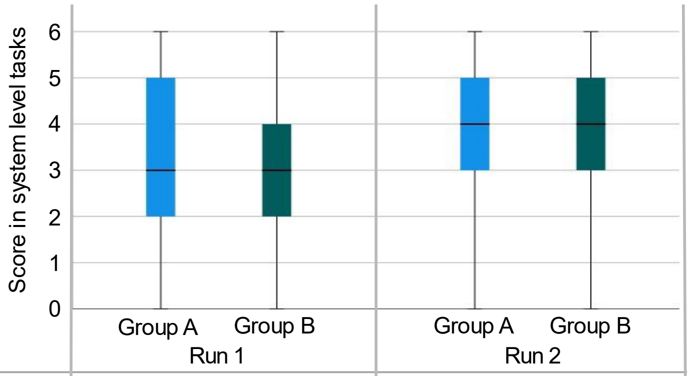


Figure 14. Pre-processing of the data sets.

3.3.2. Investigation of the impact on the system level (H1)

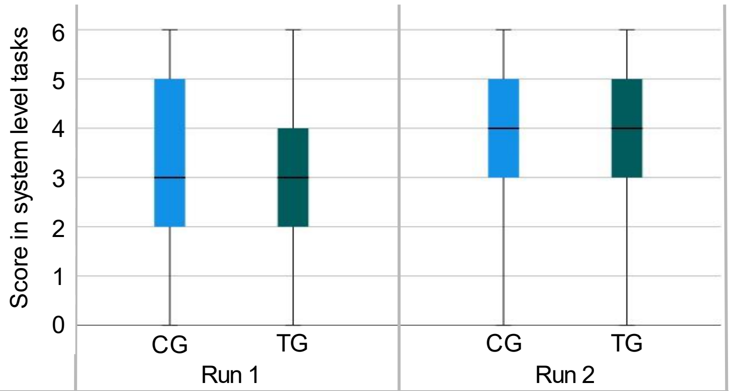
Figure 15 presents the box plots and results of the Mann-Whitney U tests comparing the system level tasks scores of the groups within runs and comparing the runs to each other. In both runs, there is no significant differences between group A and group B. This is the expected result due to the random assignment of the participants to the groups. The results of the Mann-Whitney U test for the comparison of the first run (first semester students) to the second run (third semester students) shows a significant effect (U=23769, p=<0.001). According to Cohen (1988), this is a weak effect (r=0.189).



Group comparison	Run 1	Run 2
Mann-Whitney U	6179.5	3859.0
Z	0.127	-0.015
p	0.899	0.988
r (Cohen 1988)	-	-
Run comparison		
Mann-Whitney U	23769.0	
Z	3.772	
p	<0.001	
r (Cohen 1988)	0.189	

Figure 15. Results of Mann-Whitney U tests comparing the groups within runs and comparing the runs regarding the system level tasks.

The assessment of the impact of the modelling method on system understanding on the system level is based on the number of correct assigned system behaviours to system variants. Figure 16 shows the comparison of the control and test groups and the results of the Wilcoxon tests. No significant effect is found in either run. Consequently, hypothesis H1 was rejected.



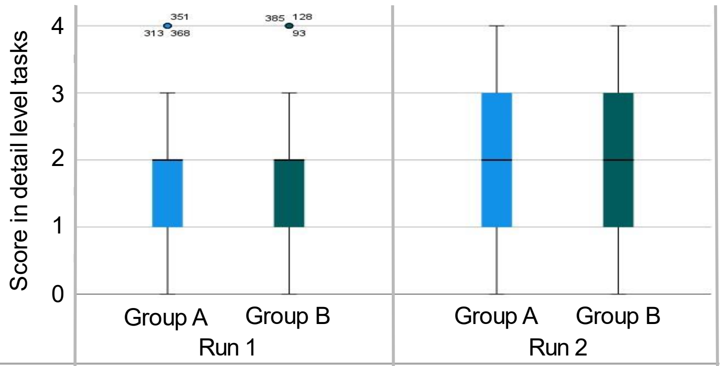
Impact investigation	Run 1	Run 2
N	111	88
Z	-1.144	-1.708
p	0.253	0.088
r (Cohen 1988)	-	-

Legend: CG = Control group; TG = Test group

Figure 16. Results of Wilcoxon tests comparing the control and test groups within runs regarding the system level tasks.

3.3.3. Investigation of the impact on the detail level (H2)

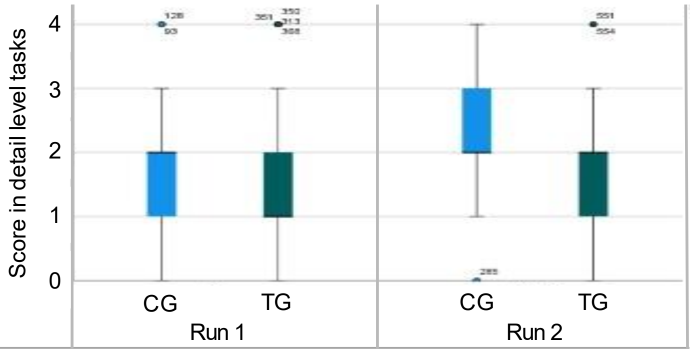
Figure 17 presents the box plots and results of the Mann-Whitney U tests comparing the detail level tasks scores of the groups within runs and comparing the runs to each other. In both runs, as expected, there are no significant differences between the groups. In contrast, the Mann-Whitney U test for the comparison of the first run (first semester students) to the second run (third semester students) shows a significant effect (U=9727.5, p=0.009). According to Cohen (1988), this is a weak effect (r=0.16).



	Run 1	Run 2
Group comparison		
Mann-Whitney U	2744.5	1601.0
Z	1.001	-0.286
p	0.317	0.775
r (Cohen 1988)	-	-
Run comparison		
Mann-Whitney U	9727.5	
Z	2.595	
p	0.009	
r (Cohen 1988)	0.16	

Figure 17. Results of Mann-Whitney U tests comparing the groups within runs and comparing the runs regarding the detail level tasks.

The assessment of the impact of the modelling method on system understanding on the detail level was based on the fulfilment of the functional requirement by correcting design faults. Figure 18 shows the comparison of the control and test groups and the results of the Wilcoxon tests. The Wilcoxon test shows no significant effect in the first run. The results of the Wilcoxon test of the second run shows a significant effect (Z=-3.65, p=<0.001, N=58). This corresponds to a medium effect (r=0.479) according to Cohen (1988). The boxplot shows that the effect is not an improvement, but a significant decrease in system understanding on the detail level in the run with the third semester students. Consequently, hypothesis H2 is rejected.

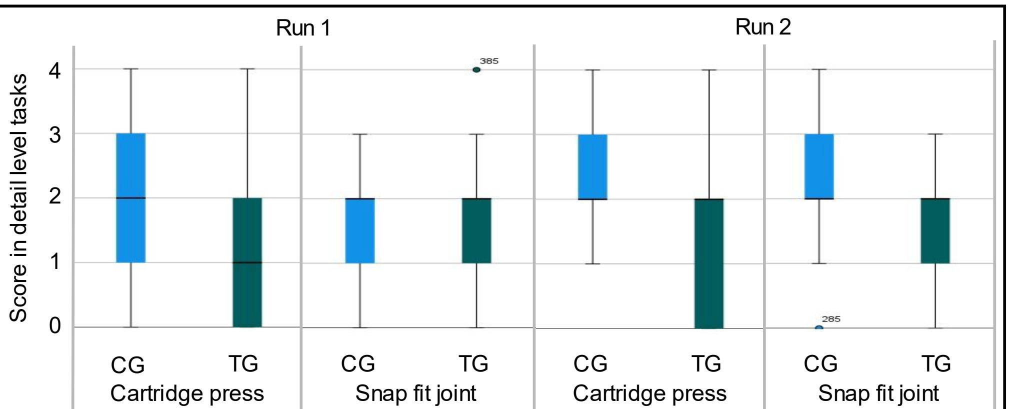


Impact investigation	Run 1	Run 2
N	71	58
Z	-1.516	-3.650
p	0.129	<0.001
r (Cohen 1988)	-	0.479

Legend: CG = Control group; TG = Test group

Figure 18. Results of Wilcoxon tests comparing the control and test groups within runs regarding the detail level tasks.

In order to examine the impact in a more differentiated way, Figure 19 shows the comparison of the control and test groups additionally divided according to the technical systems. Since each subject works on each system only once and the samples are therefore independent, the Mann-Whitney U test is used in this case. The results of the first run show no significant effect for any of the technical systems. In the second run, the Wilcoxon test shows a significant effect for the cartridge press (U=270, p=0.021) and snap fit joint (U=229, p=0.002). These results correspond to a medium effect for both technical systems according to Cohen (1988). The effect for the cartridge press is stronger (r=0.403) than the effect for the snap fit joint (r=0.304).



Impact investigation	Run 1 Cartridge press	Run 1 Snap fit joint	Run 2 Cartridge press	Run 2 Snap fit joint
Mann-Whitney U	479.5	591.5	270.0	229.0
Z	-1.752	-0.428	-2.315	-3.069
p	0.080	0.669	0.021	0.002
r (Cohen 1988)	-	-	0.304	0.403

Legend: CG = Control group; TG = Test group

Figure 19. Results of Mann-Whitney U tests comparing the control and test groups within runs differentiated according to the technical systems regarding the detail level tasks.

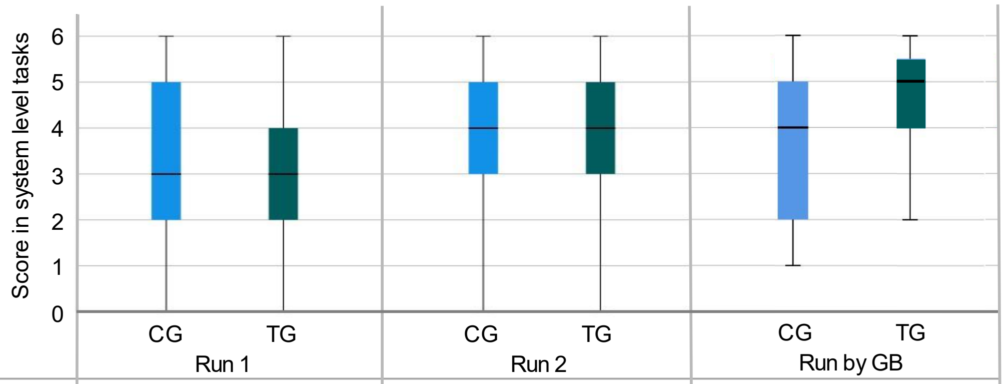
3.4. Discussion

By building on previous operationalisations and study designs, a study design for the validation of the efficacy of qualitative modelling on system understanding could be developed. Using this study design, it was possible to conduct a digital, experimental study with a high number of subjects within a teaching course. However, due to the realization in a teaching course without compulsory attendance and the technical implementation, there was a high loss of data in the pre-processing of the raw data. On the detail level, technical weaknesses, such as the possibility to enter an unclear name, were noticed, which can be eliminated by optimizing the technical implementation. On both levels of understanding, not completing the tasks was the biggest reason for data exclusions. This problem may be attributed in part to the lack of compulsory attendance, but possibly also to the size of the study. One way to reduce the time required is to adapt the study design so that each subject only completes the control *or* test group tasks.

The selected technical systems should be challenging for the subjects in order for qualitative modelling to be necessary as a design support. Nevertheless, the complexity should not lead to overwhelming tasks which would negatively influence the performance of the subjects. Figure 16 and Figure 18 show that the mean values of the control groups are in the mid-range of the scales on both levels of understanding. Accordingly, the complexity of the technical systems fitted the subjects' abilities. This ensures the possibility of perceiving both an improvement and a decrease in system understanding of the test groups.

The comparison between the runs shows a significant higher score in run 2 on both levels of understanding (see Figure 15 and Figure 17). This finding is in line with the expected difference between the groups due to the subject selection. The subjects of run 2 had already been studying longer and were mechanical engineering students in contrast to students of chemical engineering and process engineering in run 1. The field of study is relevant due to the selected strictly mechanical systems. This is a positive indication regarding the suitability of the measurements used on the two levels of system understanding.

As discussed, the difference in achieved scores between the groups of subjects is in accordance with the prior knowledge of the groups. The differences between groups with different prior knowledge becomes increasingly clear when the scores achieved on the system level are compared with the scores of the study by Grauberger et al. (in review), which are referred to as "run by GB" (See Figure 20). In the control group, the second run of this study achieved similar values to the run by GB. The influence of prior knowledge on the impact of the modelling method is reflected when the difference in prior knowledge between the subject groups increases. The comparison thus indicates that a transfer of findings between groups of subjects is only possible if the prior knowledge of the groups is almost identical.



Impact investigation			
N	111	88	35
Z	-1.144	-1.708	-2.223
p	0.253	0.088	0.026
r (Cohen 1988)	-	-	0.375

Legend: CG = Control group; TG = Test group; GB = Grauberger et al. (in review)

Figure 20. Comparing the results of the Wilcoxon tests on the system level tasks to the previous run by Grauberger et al. (in review).

However, the reliability of the comparison of system-level results between studies is limited because, in addition to the subjects' prior knowledge, the sequence of tasks was changed. Changing multiple influencing factors prevents the ability to definitively associate results with the responsible influencing factors. For this purpose, the development steps between research methods that build on each other must be reduced to the systematic modification of individual influencing factors.

As in the study design used by Grauberger et al. (in review), order effects can not be excluded in the study design due to the double use of subjects and sequence of tasks (for the investigation setup, see Figure 10). However, the lack of a significant improvement between the control and test groups is an indicator against an order effect due to the double use of the participants. A comparison with the previous study by Grauberger et al. to investigate the effect of the sequence of tasks is not possible due to the different subject groups and statistical results. Consequently, no absolute statement can be made about the impact of the order effect.

The conducted study provided new insights regarding the impact of the modelling method based on the C&C²-Approach on system understanding. Thereby, the rejection of the research hypotheses H1 and H2 serves as important guidance for the further development of the method.

The most notable result is the decrease in system understanding on the detail level in the second run. The reason for the decrease can not be understood from the statistical analysis of the final results. However, on the basis of the data, possible causes can be discussed, which can be examined in future investigations.

Looking at previous applications of the C&C²-Approach, most applications were performed by developers of the approach with extensive knowledge regarding its use (Grauberger et al., 2020). Therefore, the correct application of the approach was ensured by the competence of the developers. In this study, the competence of the users must be assumed lower and only the completion of the modelling steps was ensured. The suitability of the training course for subjects with no to minimal prior knowledge to reach a sufficient competence to apply the approach was also not ensured beforehand. The Applicability study by Eisenmann, Grauberger, and Matthiesen (2021) reached sufficient applicability only with subjects with more advanced prior knowledge. Therefore, an investigation regarding the correctness of application should be performed based on the modelling steps in the ILIAS-platform and created C&C² models to allow conclusions of the results about the modelling method. Based on the studies conducted on the C&C²-Approach, it can be assumed that the applicability increases with more advanced prior knowledge. The decrease in understanding on the detail level (see Figure 18) is in contrast to this assumption. In run 2, a better applicability and consequently efficacy would be assumed due to the more advanced prior knowledge. The inconsistency in run 2 therefore opposes the applicability and consequently training course as a cause.

Based on the indications, the application of the modelling method based on the C&C²-Approach itself must be considered as the cause for the decreased scores on the detail level. Using this modelling method represents an unaccustomed approach for the participants. However, the current measurement does not cover the subjects' behaviour during the tasks, motivation, concentration, or mental workload. To investigate this cause, further studies are therefore necessary that focus on the subjects during the tasks. These studies may include think-out protocols, eye tracking studies or questionnaires on the perceived workload, like the NASA Task Load Index. These studies would allow a more detailed comparison of the intuitive approach and combined approach plus insights into the application of qualitative modelling among subjects with minimal prior knowledge.

The results of the study can also be considered as indicators for the application of qualitative modelling in general. It was shown that compulsory application of qualitative modelling can lead to lower performance than an intuitive approach with subjects with no to minimal prior knowledge. Accordingly, the application must always be decided depending on the situation and user. Investigating the causes for worsened performance, studying other example problems, and using other modelling methods could provide further insight into when an application is effective.

3.5. Conclusion

This contribution described the development of a study design to investigate the impact of qualitative modelling on the system understanding of subjects in a controlled test environment. The study design relies on a combination of behaviour identification tasks and technical system design tasks for the measurements, marking a step in the direction of more mature, practical research methods. However, as the study design provides plausible results, it still offers room for improvement, particularly with respect to the amount of data losses and investigation of order effects. For this purpose, further studies will need to be conducted in which individual influencing factors, e.g., task order, are intentionally varied. In order to gain further insight, also with regard to the requirements of other design methods regarding the research method, comparative applications with other qualitative modelling methods will have to be carried out in the future.

The study design was applied to proceed with the validation of the modelling method based on the C&C²-Approach. For this purpose, a study was conducted in two runs in teaching courses with a total of 306 students, who made two homogeneous groups with different levels of prior knowledge. No significant improvement in system understanding was found. Consequently, both hypotheses were rejected. Therefore, for further validation of the method, more insights need to be obtained on how the method is applied by subjects with no or minimal prior knowledge. The focus should be on identifying the cause for the decrease in system understanding on the detail level. For this purpose, studies with talking out loud, eye tracking or measurements of mental stress are suitable. These would also provide insights into when qualitative modelling does not contribute to improving system understanding.

4. Discussion of the study within the detail process

In this chapter, the results and discussion are considered from the perspective of the detail process for quantitative studies with a focus on efficacy. In the detail process, the execution of the study and its analysis is followed by the *decision on the suitability of the data collection procedure for efficacy*. This decision is to be made based on the criteria objectivity, reliability, and validity, which are examined here.

The digital implementation of the training course without an instructor and automation of data collection contributes to the objectivity of the study. Data evaluation was also largely automated but had a weakness in the pre-processing of the raw data. The weakness was the interpretation of problem descriptions of subjects regarding technical or personal problems. However, relative to the number of subjects, these steps excluded only a few subjects from the analysis. Consequently, the best way to further assess the objectivity of the study is to replicate it with additional different groups of subjects.

To ensure reliability, it is recommended to conduct several runs and compare their results. This goal was pursued by conducting the study with a total of 306 subjects in two runs. As the participants between the runs were significantly different, only one run with participants with homogenous prior knowledge is currently available. Therefore, further runs with identical prior knowledge to one of the existing runs are necessary. The requirement regarding the prior knowledge of the subjects in further runs is thus oppositional to the requirements for objectivity.

To establish validity, it is recommended to vary the study design in terms of the assignment of subjects to control and test groups, to vary the task formulations, and to discuss different explanations for the results. The variations were not covered within this study and therefore require further study runs. The argumentative consideration of alternative explanations for the occurrence of the results represents the most far-reaching possibility. The discussion at this point complements the discussion of possible causes in Section 3.4. Looking at the previous studies, the study by Grauberger et al. (in review) is an indicator for the validity of the training course, study design, and system level measurement. As the study by Grauberger et al. was conducted with subjects who had more prior knowledge, the adoption of results must be done with caution. Therefore, central elements of the research method are discussed below with the study by Grauberger et al. as a comparison:

- *Measurement approach.* The segmentation of system understanding was adopted, and the resulting scores are in line with the expectation based on the prior knowledge of the subjects (see Section 3.4). However, the sufficiency of dividing system understanding into two sub-level is questionable. The investigation of existing operationalisations in Section 3.1.1 revealed up to three measured sub-aspects.

Vermaas (2010) defined five key-concepts describing different meanings of the term function. These key-concepts were goal, action, function, behaviour, and structure. It is therefore unclear to what extent the two measurements cover the system understanding of participants. Further studies using more measurements and investigating the correlations between those could extent the knowledge of representative measurements for system understanding. The goal would be a set of widely used metrics and measurement procedures, like the metrics for ideation by Shah et al. (2003)(see Section 2).

- *Study design*. As described in Section 3.4, effects due to the double use of subjects and sequence of tasks can not be excluded. However, the missing improvement between control group and test group is an indicator against a learning effect. Though, the relevance of learning effects due to the sequence of tasks will gain importance when the number of measured aspects of system understanding increases. In the case of using the five-key concepts of Vermaas (2010), an increase in system understanding due to preceding tasks must be anticipated. The decrease in the detail level scores might indicate a decrease in the motivation or concentration of the subjects due to working on the identical task types in the test group. To investigate this effect, the use of the Solomon four-group design[2] or an additional questionnaire could be used in future studies.

In summary, this consideration demonstrates a problem of the application of the detail process. A single study is not sufficient to fulfil the criteria of the research method. Instead, several studies must be carried out with varying objectives. In addition to studies that intentionally vary individual influencing variables, the segmentation of system understanding should be further investigated. For this purpose, other modelling methods should be used in order to capture the different impacts of the methods on the levels of understanding. This would also counteract an unintentional adaptation of the research method to the design method. Only after the successful measuring of system understanding has been ensured, a *decision on the efficacy of the design method* is possible.

[2] For more information on the Solomon four-group design, see Solomon (1949).

5. Summary and Outlook

This thesis investigated the impact of qualitative modelling with the Contact & Channel Approach on the system understanding of subjects. For this purpose, the reference process according to Eisenmann (in progress) was applied. Based on this process, the experiences of two previous studies were used to develop an evolved study. This study used the combination of two types of tasks, each addressing one level of understanding. For the measurement on the system level, tasks were used in which the subjects had to assign system behaviours to system variants. The detail level was addressed via a design task in which subjects had to fulfil a functional description by adjusting design parameters in a web-based CAD configurator. At the beginning of this task, the technical systems contained four function-critical faults. This set of tasks was completed twice by each subject, before and after a training course on the modelling method based on the C&C²-Approach. This study was conducted in two runs with subjects with different levels of prior knowledge. The purpose of these runs was to examine the impact of the modelling method on the detail level and to compare the impact between the runs.

The results revealed no significant positive effect of the modelling method on any level of understanding in any of the runs. Consequently, both research hypotheses were rejected. Nevertheless, insights into the research method were achieved by the two runs. The execution of the implemented research method showed high data losses and a lack of control of influential factors. The data losses were mostly caused by incomplete data sets and technical weaknesses in the implementation. The influential factors task order, order effects due to double use of subjects, and prior knowledge of the subjects should be mentioned. For example, findings from a previous applicability study could not be transferred because the prior knowledge of the subjects was different. Because of the change of more than one of these factors compared to previous studies, it was not possible to conclude from the results to the cause. Therefore, the suitability of the research method could not be confirmed. Due to the consistency of the results of the measurements with the expectations based on the prior knowledge of the subjects, implications for qualitative modelling with the C&C²-Approach can nevertheless be derived.

The most significant result is the decrease in system understanding on the detail level. The discussion of alternative causes indicated that the modelling itself is responsible for this decrease. It is therefore necessary to conduct further investigations on how the qualitative modelling impacted the subjects' actions in the web-based design tasks. These investigations may use think-out protocols, eye tracking studies or questionnaires on the perceived workload, like the NASA Task Load Index to identify the cause.

For the further development of the research method, further studies with intentional variation of influencing factors and additional replications for reliability are necessary. In this process, the development steps must be reduced in order to increase the possibilities for determining the influences. These development steps can also be implemented through a study in which more than two subject groups are created. For example, by forming three groups of subjects, a control group with study design A and task order A, a test group with study design A and task order B, and a test group with study design B and task order A. In doing so, the first group would replicate the existing study for reliability. The other two groups would specifically investigate influencing factors. By developing such targeted study designs within a study, the number of studies can be reduced. On the other hand, the number of studies required can be reduced by establishing a relationship with the research methods used in other design method validations. Eisenmann, Grauberger, Üreten, et al. (2021) identified 54 design method validation studies published between 2010 and 2020. 7 of these studies were experimental application evaluations. Examination of these studies and others from before 2010 would potentially allow reuse of previously tested study design elements. This would be a step toward looking at research methods across design methods rather than specifically developing them for each design method, leading towards a more sustainable development.

List of figures

List of tables

References

Andreasen, M. M., Hansen, C. T., & Cash, P. (2015). *Conceptual design: Interpretations, mindset and models*. Springer. https://doi.org/10.1007/978-3-319-19839-2

Blessing, L. T., & Chakrabarti, A. (Eds.). (2009). *DRM, a Design Research Methodology*. Springer London. https://doi.org/10.1007/978-1-84882-587-1

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.

De Kleer, J. (1984). How circuits work. *Artificial Intelligence*, *24*(1-3), 205–280. https://doi.org/10.1016/0004-3702(84)90040-7

Eckert, C., Alink, T., Ruckpaul, A., & Albers, A. (2011). Different notions of function: results from an experiment on the analysis of an existing product. *Journal of Engineering Design*, *22*(11-12), 811–837. https://doi.org/10.1080/09544828.2011.603297

Eisenmann, M. (in progress). *Validierungs-Navigator – ein Referenzprozess für Validierungsstudien zur Wirkung von Konstruktionsmethoden.*

Eisenmann, M., Grauberger, P., & Matthiesen, S. (2021). SUPPORTING EARLY STAGES OF DESIGN METHOD VALIDATION - AN APPROACH TO ASSESS APPLICABILITY. *Proceedings of the Design Society*, *1*, 2821–2830. https://doi.org/10.1017/pds.2021.543

Eisenmann, M., Grauberger, P., Üreten, S., Krause, D., & Matthiesen, S. (2021). Design method validation – an investigation of the current practice in design research. *Journal of Engineering Design*, 1–25. https://doi.org/10.1080/09544828.2021.1950655

Forbus, K. D. (2011). Qualitative modeling. *Wiley Interdisciplinary Reviews. Cognitive Science*, *2*(4), 374–391. https://doi.org/10.1002/wcs.115

Frey, D. D., & Dym, C. L. (2006). Validation of design methods: lessons from medicine. *Research in Engineering Design*, *17*(1), 45–57. https://doi.org/10.1007/s00163-006-0016-4

Galbraith, J. (1973). *Designing Complex Organizations*. Addison-Wesley.

Gericke, K., Eckert, C., & Stacey, M. (2017). What do we need to say about a design method? In *DS 87-7 Proceedings of the 21st International Conference on Engineering Design (ICED 17) Vol 7: Design Theory and Research Methodology, Vancouver, Canada, 21-25.08. 2017*.

Gladysz, B., & Albers, A. (2018). HOW DO C&C²-MODELS IMPROVE EFFICIENCY, COMPREHENSIBILITY AND SCOPE IN FAILURE ANALYSIS-AN EMPIRICAL STUDY BASED ON TWO LIVE-LABS. In *15th International Design Conference DESIGN 2018,* Dubrovnik, Croatia.

Grauberger, P., Eisenmann, M., Windisch, E., & Matthiesen Sven. (in review). *Investigating the impact of qualitative modelling in embodiment design – an experimental study on the Contact and Channel Approach*.

Grauberger, P., Pähler, L., Eisenmann, M., Beutler, P., Meboldt, M., & Matthiesen, S. (in progress). *Data-driven measurement of system understanding using web-based design tasks* [Version: March 2022].

Grauberger, P., Wessels, H., Gladysz, B., Bursac, N., Matthiesen, S., & Albers, A. (2020). The contact and channel approach – 20 years of application experience in product engineering. *Journal of Engineering Design*, *31*(5), 241–265. https://doi.org/10.1080/09544828.2019.1699035

Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, *28*(1), 127–138.

Institute of Electrical and Electronics Engineers (1998). *IEEE Standard for Software Verification and Validation*.

Jagtap, S., Warell, A., Hiort, V., Motte, D., & Larsson, A. (2014). Design methods and factors influencing their uptake in product development companies: A review. In *International Design Conference-Design 2014* (pp. 231–240). University of Zagreb and Design Society.

Johnson, S. D., & Satchwell, R. E. (1992). The Effect of Functional Flow Diagrams on the Technical System Understanding of Apprentice Aircraft Maintenance Mechanics.

Kroll, E., & Weisbrod, G. (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. https://doi.org/10.1007/s00163-019-00324-6

Matthiesen, S., Grauberger, P., Hölz, K., Nelius, T., Bremer, F., Wettstein, A., Gessinger, A., Pflegler, B., Nowoseltschenko, K., & Voß, K. (2018). *Modellbildung mit dem C&C²-Ansatz in der Gestaltung - Techniken zur Analyse und Synthese.* https://doi.org/10.5445/IR/1000080744

Matthiesen, S., Grauberger, P., & Schrempp, L. (2019). Extended Sequence Modelling in Design Engineering – Gaining and Documenting Knowledge about Embodiment Function Relations with the C&C 2 -Approach. *Proceedings of the Design Society: International Conference on Engineering Design*, *1*(1), 1483–1492. https://doi.org/10.1017/dsi.2019.154

Motte, D., & Eriksson, M. (2016). Assessment framework for a methodology under development–Application to the PDA methodology. In I. Horváth, J.-P. Pernot, & Z. Rusák (Eds.), *Proceedings of the 11th International Symposium on Tools and Methods of Competitive Engineering - TMCE'2016* (pp. 373–388). Delft University of Technology.

Olewnik, A. T., & Lewis, K. (2005). On Validating Engineering Design Decision Support Tools. *Concurrent Engineering*, *13*(2), 111–122. https://doi.org/10.1177/1063293X05053796

Pedersen, K., Emblemsvåg, J., Bailey, R., Allen, J. K., & Mistree, F. (2000). Validating Design Methods and Research: The Validation Square. In *Volume 4: 12th International Conference on Design Theory and Methodology* (pp. 379–390). American Society of Mechanical Engineers. https://doi.org/10.1115/DETC2000/DTM-14579

Reich, Y. (2010). *My method is better!* Springer.

Royer, J. M. (1986). Designing instruction to produce understanding: An approach based on cognitive theory. *Cognitive Classroom Learning: Understanding, Thinking, and Problem Solving*, 83–113.

Satchwell, R. E. (1997). Using Functional Flow Diagrams to Enhance Technical Systems Understanding. *Journal of Industrial Teacher Education*, *34*(2), 50–81.

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

Solomon, R. L. (1949). An extension of control group design. *Psychological Bulletin*, *46*(2), 137–150. https://doi.org/10.1037/h0062958

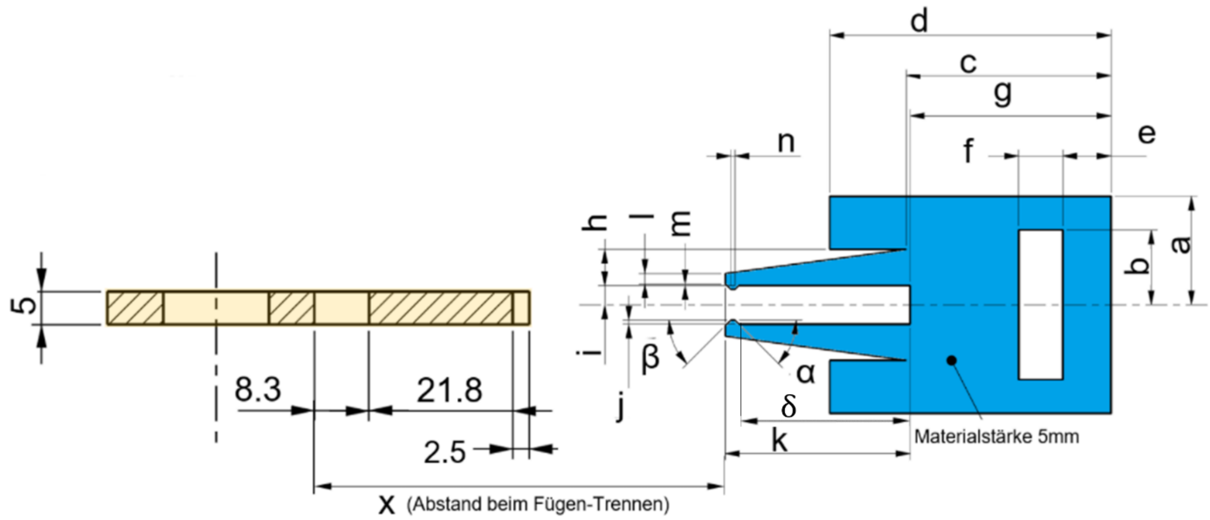
Vermaas, P. E. (2010). Technical functions: towards accepting different engineering meanings with one overall account. In *Eighth International Symposium on Tools and Methods of Competitive Engineering (TMCE 2010), Ancona, Italy* (pp. 183–194). Delft University of Technology, Faculty of Design Engineering.

Verworn, B., Herstatt, C., & Nagahira, A. (2008). The fuzzy front end of Japanese new product development projects: impact on success and differences between incremental and radical projects. *R&D Management*, *38*(1), 1–19. https://doi.org/10.1111/j.1467-9310.2007.00492.x

White, B. Y., & Frederiksen, J. R. (1986). Progressions of Qualitative Models as a Foundation for Intelligent Learning Environments. Report No. 6277.

Appendix

Calculation formulas for the snap fit joint



1. The assembly force must be smaller than the disassemblyforce.

$$90° > \alpha > \beta > 0$$

2. The width of the hook must be sufficient for the counterpart.

$$2{,}5 \leq i < 2{,}5+j$$

3. The length of the hook must be sufficient for the counterpart.

$$21{,}8 \leq \delta$$

$$\delta = \begin{cases} k - n - \frac{j+m}{\tan(\beta)} - \frac{j}{\tan(\alpha)} & , if\ i \leq 2{,}5 \\ k - n - \frac{j+m}{\tan(\beta)} - \frac{j}{\tan(\alpha)} + \frac{i-2{,}5}{\tan(\alpha)} & , if\ 2{,}5 < i < j + 2{,}5 \\ k - n - \frac{j+m}{\tan(\beta)} & , if\ j + 2{,}5 \leq i \end{cases}$$

4. The stiffness must be increased.

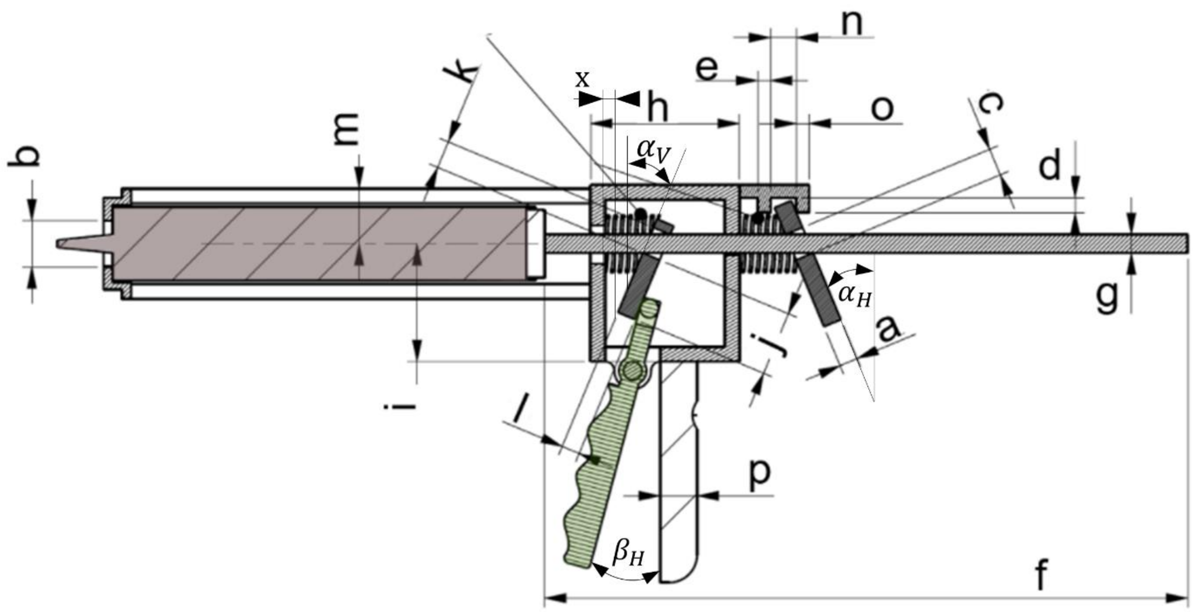
$$\vartheta < 1$$

$$\vartheta = \left(\frac{L_{eff}\, H_{eff\,IC}}{L_{eff\,IC}\, H_{eff}}\right)^3 \qquad (IC\ =\ Initial\, Configuration)$$

$$H_{eff} = \frac{h + l}{2}$$

$$L_{eff} = \begin{cases} k & , if\ g \geq c \\ k - (c - g) & , if\ g < c\ and\ d \geq c \\ k & , if\ g < c\ and\ d < c\ and\ d < g \\ k - (d - g) & , if\ g < c\ and\ d < c\ and\ d \geq g \end{cases}$$

Calculation formulas for the cartridge press



1. Tilting of the front lever must be possible.

$$0° < \alpha_V$$

$$\alpha_V = \begin{cases} 90° - sin^{-1}\frac{g}{\sqrt{l^2+k^2}} - tan^{-1}\frac{l}{k} & , if\, k > g \\ 0° & , if\, k \leq g \end{cases}$$

2. Tilting of the back lever must be possible.

$$0° < \alpha_H$$

$$\alpha_H = \begin{cases} 90° - sin^{-1}\frac{g}{\sqrt{a^2+c^2}} - tan^{-1}\frac{a}{c} & , if\, c > g \\ 0° & , if\, c \leq g \end{cases}$$

3. The cutout holding the back lever must allow tilting.

$$n > n_{Limit}$$

$$n_{Limit} = \mathrm{a} * \cos(\alpha_H) + (d - 0.491809[mm])\tan(\alpha_H)$$

4. Clearance between housing and pressure building mechanism is increased.

$$x_{Max} > x_{IC} \qquad (IC = Initial\, Configuration)$$

$$x_{IC} = 13.301mm$$

$$x_{Max} = \frac{(-j * \sin(\alpha_V) + 35 * \sin(\beta_H) + h - 63) * \cos(\alpha_V)}{\cos(\alpha_V)} + \frac{(-35 * \cos(\beta_H) - 0.5 * g + i + 5) * \sin(\alpha_V) - l - 5}{\cos(\alpha_V)}$$

$$\beta_H = 31.526°$$