Proceedings of the Design Society, Volume 5: ICED25
https://doi.org/10.1017/pds.2025.10229

Investigating Engineering Design Tasks – Iterative Testing in Large-Scale Engineering Courses

Mona Batora [1,✉], **Oliver Liewerenz** [2], **Patric Grauberger** [2], **Noreen Enseroth** [1], **Mathis Wolter** [1], **Sven Matthiesen** [2] and **Nikola Bursac** [1]

[1] ISEM - Institute for Smart Engineering and Machine Elements (ISEM, Hamburg University of Technology (TUHH), Hamburg, Germany, [2] Institute of Product Engineering (IPEK), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

✉ mona.batora@isem-tuhh.de

ABSTRACT: In this research a study environment is presented that enables iterative design in large engineering lectures and show possibilities for investigations at two example lectures from German universities. The initial results show that it is possible for large lecture-hall-based courses to engage in in-depth tasks of engineering design. Design researchers can use the generated data to measure infuences, e.g. the applied methods on specifc design tasks. Two key insights include the potential for large courses to serve as large-scale research environments for design research and the observed effects of infuences on students' decision-making processes. This approach offers a promising method to further explore the complexities of decision infuences and design optimization in educational settings.

KEYWORDS: Large-Scale, Design Education, Integrated Product Development, Prototyping, Collaborative Design

1. Introduction

In design education, tools, methods and processes of design are taught either in a theoretical way for large engineering courses including more than one hundred students or in a more practice-oriented way in smaller courses (Liewerenz et al., 2023). The education of future design engineers shows challenges similar to investigations of design research, where these tools, methods and processes are developed and evaluated through proband studies. Design tasks close to reality are hindered by the high effort needed for conducting them. This effort stems mainly from manufacturing and testing. However, these elements are necessary for experiencing the iterative character of design, where prototypes play an important role for success (Jobst et al., 2024).

Testing enables to evaluate decisions, as it gives feedback to designers regarding the functions they implemented through their design decisions (Liewerenz et al., 2023). Exams are used as an established measuring tool in large groups to test how decision-making has been communicated. Qualitative methods that can capture the procedure in great detail are also possible for a small number of people is investigated and statistical effects are not in focus, e.g. in case studies. The challenge lies in the absence of measurement tools for iterative design in large groups, making quantitative assessments of design tasks, such as the gain of design knowledge, highly elaborate. The problem is that no measurement tool is available for iterative design in large groups, so quantitative investigations of design tasks like the gain of design knowledge are still highly elaborate.

Our research question derived from this problem is "How can large courses be utilized as a large-scale research environment to conduct quantitative and realistic studies in engineering design?"

2. State of Research

2.1. SGE - System Generation Engineering

To investigate realistic tasks in engineering design, the consideration of evolutionary development of technical systems is necessary. The SGE - System Generation Engineering focuses on the systematic evolution of product families through the creation of successive product generations. Each generation builds on its reference product from previous iterations, incorporating improvements in performance, technology, and design (Albert et al., 2015). Consequently, product engineering work like engineering design is also conducted in engineering generations, comprising iterative phases of product engineering, including design, manufacturing, testing, and interpretation. The model of SGE describes the process of leveraging existing solutions from prior products, competitors, and external industry sources as references to drive innovation success, while reducing the risks associated with design changes (Albers et al., 2017). The model emphasizes these connections to facilitate the development of new and improved product generations. This approach is critical for industries that rely on product development, where products must adapt to market changes without losing the value of previous generations. Research in SGE emphasizes the importance of knowledge management across generations and the strategic reuse of design elements to ensure that new product generations meet market and technical requirements more effectively. (Albers and Rapp, 2022)

2.2. Testing for gain of knowledge

In product development processes, testing activities are essential activities to ensure gain of knowledge and validation, but they are also huge drivers of resources like costs and time. Testing in general is “the process of operating a system or component under specifed conditions, observing or recording the results, and making an evaluation of some aspect of the system or component” (610.12, 1990). Testing activities can be differentiated into planned and unplanned activities. Many of these activities can be defined in project planning and serve purposes like ensuring product functionality or fulflment of legal requirements (Engel, 2010). During product development, many different testing activities can occur (Tahera et al., 2018).

A case, where unplanned testing activities emerge, is caused by missing specifc design knowledge (Hubka and Eder, 1990), which cannot be gained through enquiry, simulation, or consulting of experienced engineering designers (Matthiesen, 2021). This is also a cause for iterations, in which problem and solution evolve together in cycles (Dorst and Cross, 2001; Meboldt et al., 2012).

Generating knowledge empirically is always necessary when valid models are lacking and the interactions within the product to be developed or with its surrounding systems are unknown or diffcult to understand. Development practice teaches us that it is practically impossible to “think of everything directly”. This is also the reason why a large number of verification tests are planned in product development processes. They are unavoidable and make product development processes expensive and time-consuming. In product development, empirical knowledge building is called testing. Testing involves operating, observing and evaluating an object. Developing knowledge empirically is often more time-consuming than researching or analyzing it. However, if models are missing or unknown, this empirical knowledge building is an indispensable part of the development project. (Matthiesen and Grauberger, 2024)

2.3. Research Environments in Engineering Design

Common research environments for studies are mainly feld studies, live labs and laboratory studies. Laboratory studies are conducted in controlled laboratory environments with often high numbers of participants, offering precise control over variables and high reproducibility and reliability of results. However, conducting experiments in laboratory settings isolates methods from practical application contexts, which may result in a lack of refection of the complexities of real-world environments. (Roe and Just, 2009)

Field studies are conducted in real-world settings, providing higher external validity for the results. However, the data interpretation is dependent of many uncontrolled variables, which can lead to variability in the results (Runeson and Höst, 2009). Cassidy (2015) investigates the disadvantages and challenges of considerable time and resources in preparation for the study ending often in quantitative smaller amounts of people participating.

When examining the negative aspects of the studies and the concept of lectures with hundreds of students, it opens up opportunities to combine both into an improved concept. With lectures providing a high amount of potential subjects, a wide range of data can be generated. To ensure that the data is collected in a research-appropriate manner, various methods are available. Surveys are an effective way to gather data from a large number of students and analyze it quickly. They represent a common quantitative method for capturing students' opinions and ideas (Malik et al., 2018). Recordings provide the opportunity to retrospectively analyze verbal interactions between students and the lecturer. Exams are the most established method for quantitatively assessing student performance, as they evalate students' knowledge and understanding while providing feedback to the students (Cortright et al., 2003).

2.4. A Snap-Fit Connection as an Object of Research and Assignment

A snap-fit connection is a widely used technical system that lacks standardization. A solution for this specific task is unknown in the literature due to the chosen material of high-density fbreboard (HDF). Therefore, conducting testing activities is essential to gain specifc design knowledge that can help optimize the main function of achieving a required holding force. The primary purpose of a snap-ft connection is to non-destructively release connected components at a specifc holding force threshold. The geometry of the snap-ft joint has an impact on the holding force. Adjusting the snap-ft joint geometry can result in a force range from 0 N (instant release) to a release that is only possible by destroying the snap-ft joint. The task in this study is to improve the holding force of the given snap-ft connection from approximately 0 N to over 200 N, matching the force attainable by a hypothetical competitor. A key requirement is the ability to use the snap ft connection twice in a self-releasing manner without breaking it. This ensures that the task cannot be achieved through a locking geometry that requires destruction, which would contradict the primary function of allowing non-destructive disconnection at a specifc holding force threshold. To accomplish the design task, participants use the Hyper Text Markup Language (HTML) front end of a web-based Computer Aided Design (CAD) environment (onshape®), where parameters can be adjusted in real time by moving sliders. This accelerates the process of modifying the design and preparing it for manufacturing compared to the original CAD tool. The adjustable parameters are shown in Figure 1 and could be modifed using sliders, also shown in the fgure. (Liewerenz et al., 2023)

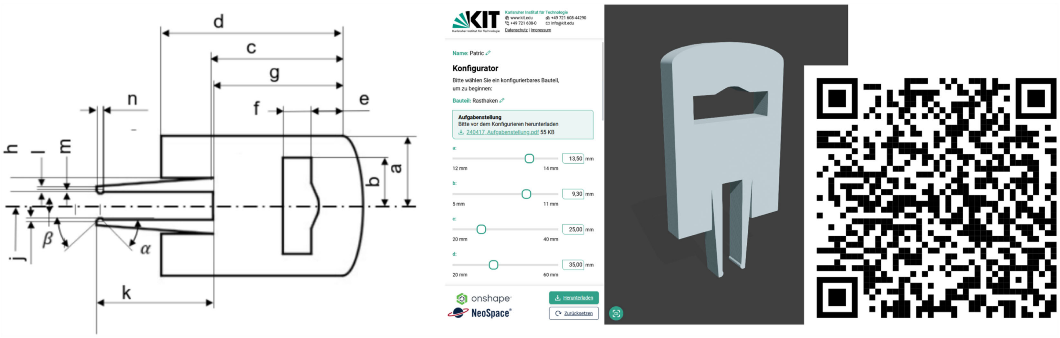

Figure 1. Technical Drawing of the snap-ft connector with 16 parameters (left) and the online confgurator featuring the QR-Code (right)

This virtual design environment enables close to reality design tasks to be conducted in parallel by large groups of engineering students and is therefore suitable for this investigation. Due to the environment also running on mobile devices, a wider range of students can participate.

3. Research Objective and Research Methodology

3.1. Research Objective

Laboratory studies enable reliable results through their controlled environment, however they do not accurately represent real-world environments. Field studies are are closer to reality, however controlling them is diffcult, which can lead to less validity. Studies with iterative design tasks are currently conducted mostly with smaller groups due to the high effort needed in manufacturing and testing. Since lectures with hundreds of students are conducted annually, this presents an opportunity to combine these settings effectively.

The aim of this study is to use large courses as a large-scale research environment to conduct quantitative and realistic studies in engineering design. This leads to the research question: How can large courses be utilized as a large-scale research environment to conduct quantitative and realistic studies in engineering design?

3.2. Methodology

Large courses with more than one hundred students in engineering design are selected to provide a homogeneous group of students with an engineering background. Students in these courses will likely have no practical experience with engineering design. The selected lectures set the basic requirements: according to the expected amount of participating students, capacities in the digital study environment like servers and wireless communication need to be provided. To follow the iterative steps (see Figure 2), manufacturing and testing capabilities need to be integrated in this environment.

When implementing and conducting these large-scale studies in two different universities, different pre-conditions apply. This includes, but is not limited to, different semesters of students, different sizes and topics covered in the courses. In addition, data points to be collected need to be determined beforehand to enable answering the research question.

3.3. Data Collection

The parameter set of each iteration's snap-ft connectors is captured automatically to investigate the participants' activities in the design data. It comprises 16 design parameters, 2 function fulfllment data and 4 metadata. The design parameters are used to capture the design activities. The function fulfllment data represent the two subsequently measured holding forces. The metadata consists of anonymized participant identifcation, iteration identifcation, and time course.

Out of 16 available parameters, 8 were relevant to improve the holding force. Via a wordcloud on the online survey platform Mentimeter, students were tasked with identifying irrelevant parameters. The platform was also used to rank multiple designs. Audio and video of lectures at Hamburg University of Technology (TUHH) were recorded.

To assess the impact of iterative design modifcations, mean values, standard deviations, and boxplots (Figure 4) were used to analyze holding force distributions. Correlation analyses examined relationships between design parameters and performance. Differences between moderated (TUHH) and unmoderated (KIT) iterations highlighted the infuence of structured guidance. Failure rates at KIT provided insights into the balance between stiffness and fexibility.

4. Large-Scale Research Environment Study

This chapter introduces the implementation and conduction of the Large-Scale Research Environment Study, which involved the snap-ft connector assignment given to students in lectures at two different universities, along with the data points collected. The second section focuses on various methods employed throughout the lecture. It explores the environmental, content-related, and personal infuences in an exploratory manner. Thus giving a congruent evaluation of the given data.

4.1. Implementation and Conduction of the Large-Scale Study in Two Lectures

Using the online confgurator from Liewerenz et al. (2023) different designs can be implemented, allowing for different topics to be covered in the scope of the lecture it is applied in. These options in the confgurator also each include assignments and boundaries to the design, one of the options being the used snap-ft connector and assignment. To create physically comparable results, material parameters like thickness are not to be changed. The sequence of iterative steps (design, manufacturing, testing, interpretation) are fxed. Using these materials for a large scale research environment, many parameters in the study design remain open. Among others, decisions on the setting in the lecture, the number of iterations, the degree of moderation, the product to be designed, and the optimization variables are up to the specifc implementation. For example, TUHH performed iterations in a moderated environment and a total of six iterations. Due to the selected assignment, both implementations optimized for a maximum holding force of the snap-ft connector. Due to supply problems, a different sheet material was used between the implementations, with Karlsruhe Institute of Technology (KIT) using 5 mm and TUHH 6mm thickness.

Figure 2 shows both approaches at KIT and TUHH. In both cases, students were moderated through the assignment, with TUHH performing a test on a reference snap-ft connector as well as two iterations during the lecture. To measure the holding force, the snap-ft-connector was attached to a stable mount. The students then used a luggage scale to pull on the connector to determine the maximum weight it could hold. Each iteration consists of a design, manufacturing, testing and interpretation phase, with exception of the reference connector. Manufacturing and testing at TUHH were live streamed into the auditorium from the workshop. To motivate students throughout the semester, awards were given to the best performing designs at each university. The lectures used as a large-scale research environment take place in Karlsruhe at the KIT and in parallel in Hamburg at the TUHH. In Karlsruhe, the bachelor course "Mechanical Design B" (second semester) with about 400 students enrolled and in Hamburg, the bachelor course "Fundamentals of Engineering Design" (second semester) with about 300 students enrolled are chosen, as these students have a rather homogeneous background and represent a large group of probands.

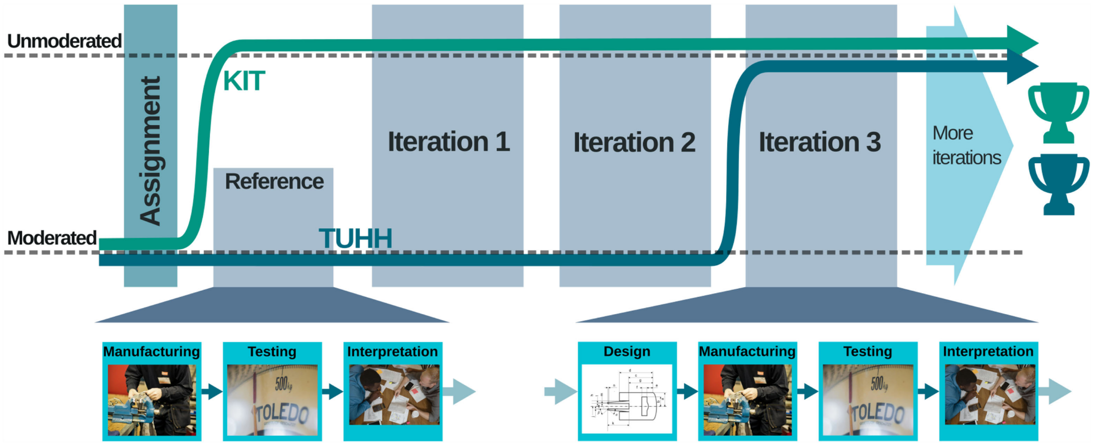

Figure 2. In both implementations the assignment was moderated. Students iterated their designs multiple times, with the best design being rewarded at the end of the semester

4.2. Explorative Part of the Large-Scale Study

In the following section, the infuences of the lecturer on the students at TUHH are investigated. The lecture recordings were analyzed to fnd factors, where the lecturer infuenced the group's decision process. Four main methodologies were found.

4.2.1. Method 1: Structured Learning through Facilitator-Led Guidance

In a large group such as basic lectures, multiple thought streams emerge, preventing a congruent thought process. Facilitator-led guidance is tasked with maintaining an overview and motivating and enabling individual students to make practical decisions based on information and ideas (Harvey et al., 2002). Over the course of the lectures held at TUHH and KIT, the lecturer assumes the role of a facilitator focusing on different aspects. At TUHH, he used guidance and nudges to steer the group's thought process. In both lectures, the lecturers clarifed the task by introducing the most important aspects and key factors of the snap-ft connector. As part of the moderated session, students were tasked with identifying the parameters of the snap-ft connector that most signifcantly infuenced the holding force. The following parameters have been highlighted as the most important ones (Figure 1), h, j, k, n, α and β. In further discussion in interpretation of the reference generation and design of the frst generation, the lecturer listens to the students' problem-solving assumptions and responds with probing questions. This approach provides indirect feedback, highlighting the plausibility or implausibility of suggestions in a way that is accessible to everyone. Furthermore, the lecturer can narrow the scope of possibilities by providing direct feedback on whether a suggestion is feasible. At KIT, the focus laid on experiencing iterations, so no steering of the lecturer is conducted. The processes of interpretation and new iterative design belongs fully to the student. As the design process progresses, the lecturer at TUHH transitioned to a more managerial and structuring role. He systematically addressed the parameters previously discussed for modifying the snap-ft connector, guiding the group step by step and thereby creating an organized process.

4.2.2. Method 2: Unifed Understanding of the Design Problem

Despite the inherent challenges of coordinating and facilitating large groups, they offer signifcant advantages. Groups are equipped with enhanced information-processing capabilities, allowing them to identify and correct errors made by individual members. This gives groups a distinct advantage in understanding the structure and rules of decision-making tasks compared to individuals (Kugler et al., 2012).

To make this possible, it is crucial that students achieve a comparable level of understanding of the design problem and expand upon their understanding throughout the lecture. Several strategies were thus employed at TUHH. During the discussion, the lecturer repeatedly engages with students' sugges-tions, encouraging them to elaborate further. This process ensures that all students in the lecture hall are brought to the same knowledge level, enabling them to follow and understand the reasoning behind decisions. Their understanding of the importance of mentioned parameters is also implied by Figure 3. This illustration shows a survey, conducted via Mentimeter, where TUHH students were asked to rank six drawings of snap-ft connectors. Based on the discussion about the parameters in the previous part of the lecture and their own intuitive understanding the students ranked the snap-ft-connectors. Each most selected rank corresponds to its true measured rank, especially clear distinctions were made for the top three connectors, signaling a good understanding of relevant design parameters. A correlation is indicated, where the holding force of the snap-ft connectors, designed by the students groups, improves signifcantly, as shown in Figure 4.

An overview on the lecture hall's blackboard shows the table, into which the holding force measurements are to be flled. This can be seen as the nudge "framing". Since decision-making in engineering design processes are also infuenced by cognitive biases, decisions made by leaders are not entirely rational. Batora et al. (2024) argues that understanding and addressing these biases in decision-making can signifcantly enhance outcomes in product development.

Once the foundational discussions have taken place and a suffcient level of understanding has been developed, the technical drawing model is introduced. This serves to minimize potential misunderstandings — both between the lecturer and the students, as well as among the students themselves. As a result, communication is signifcantly simplifed.

4.2.3. Method 3: Relevance of the Design Problem and Motivation

Tiedens and Linton (2001) found that individuals who experienced certainty appraisals — thereby becoming more confdent in the accuracy of their opinions — engaged in increased heuristic processing. These individuals exhibited greater confdence in the validity of their judgments and adopted bolder approaches. These fndings highlight the importance of certainty appraisals and their motivational impact. At TUHH, this strategy fnds application in the use of repeated direct feedback and evaluation by the lecturer. Phrases such as "well done" or "nice work" provide students with reassurance that their responses thus far are valid, fostering confdence. This encouragement, in turn, motivates students to propose further answers, including those that may involve greater risk or creativity.

In the further course of the lecture the lecturer established a recurring pattern where he lets the responsibility for the design process rest with the students. First students are initially primed to improve

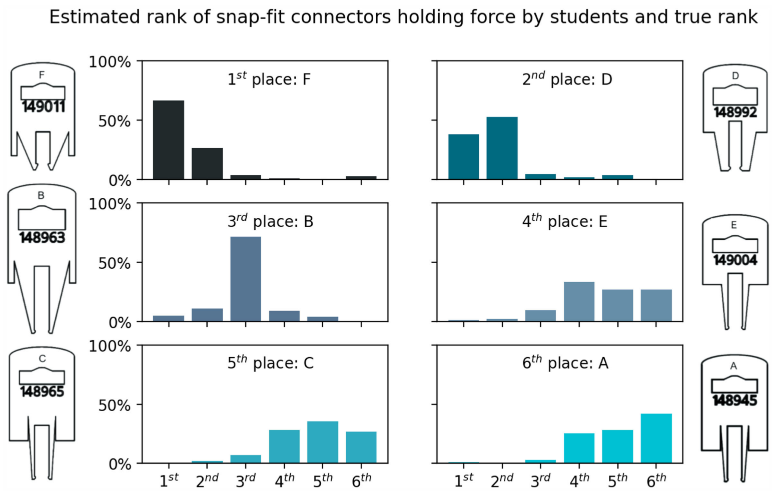

Figure 3. TUHH students ranked six drawings of snap-ft connectors

the default design. Then the lecturer keeps the pattern by consistently asking the students for their validation of the right settings for the snap-ft-connector parameters. By employing this methodology, students are provided with a sense of autonomy and responsibility, simulating a preliminary version of real-world work scenarios.

4.2.4. Method 4: Supportive Environment that Permits Mistakes

From the previous methodology, it becomes evident how crucial the emotional environment and an individual's emotional state are for their learning and decision-making processes.

In a study conducted by Holley and Steiner (2005), 63 % of high school students reported having at least one class where they felt unable to openly share their thoughts and ideas. Students felt insecure when instructors were overly critical, biased, or dismissive of their opinions.

A clear pattern can be observed at TUHH regarding the creation of an environment that encourages students to actively participate in the design process. An initial assessment is gathered through an online survey. This survey is designed to be anonymous, allowing students to provide uncertain answers without hesitation. The methodology of using online surveys is also applied repeatedly later in the lecture. The lecturer consistently emphasizes and primes students with the idea that they are engaging in design for the frst time and that there is no single correct solution in design. This approach conveys that high-performance outputs are not expected, and the initial demands on students are intentionally kept low. Subsequently, the Think/Pair/Share concept is implemented. In this method, students frst refect individually (Think), then exchange ideas with their neighbors (Pair), and fnally share their assumptions in a class discussion (Share). The Think and Pair phases provide a safe environment for students to make initial assessments, which can be verifed or refned through their peers. As a result, students feel more confdent in sharing their validated ideas during the larger discussion.

4.3. Performance of Snap-Fit Connectors

The holding forces of each generation for both implementations are shown in Figure 4. As each snap-ft connector was released two times, each holding force of both tests is shown for KIT. Connectors that irreversibly broke got scored with 0 N. At TUHH, only the holding force on the frst release is plotted. As the frst two iterations are made during the Live-Lab lecture, these are highlighted.

In both implementations the average holding force of the snap-ft connectors increased over the iterations. At TUHH, students were able to design their own generations at home in groups after the lecture, resulting in a wide spread of holding forces, as shown in Figure 4. Over the generations, students mostly widened the base of the arms and the height of the hooks (parameter "h" and "j" from Figure 1). Students reached a breaking point in iteration 4, where the connectors got too stiff to be reversibly separated from their base, resulting in a noted force of 0 N. The highest measured holding force was 163 N with no team reaching the set goal of 200 N. The stronger trend upwards for the holding force for iterations discussed and designed in the lecture suggests a high importance of nudging performed by the lecturer.

At KIT, the average holding force slowly climbs for the frst test, excluding the frst exercise. The mean holding force on the second release stagnated at 11 N. Except for test 1 of exercise 1, over half of all connectors broke on each test. Most designs failing suggests a good understanding of which parameters matter to improve stiffness of the connectors, yet not considering the needed fexibility in the compliant mechanism. This is in accordance with observations made at TUHH, where students did correctly match designs to their respective performance (Figure 3). These factors were also oftentimes mentioned in a survey performed during the lecture at TUHH, where students considered adding hooks the most important change to improve the reference products performance.

In general, students at KIT oftentimes started out with too stiff connectors, decreasing parameters like "h" and "j" (Figure 1) over generations. This results in less connectors breaking and getting scored 0 N, raising the average holding force. In contrast, TUHH students rather slowly increased these parameters, starting out with less stiff connectors which did not break so often.

5. Discussion

Based on the gained results, the research question of this study can be answered as follows: By combining the measurement of both embodiment and functional behavior in design tasks conducted within large courses, it

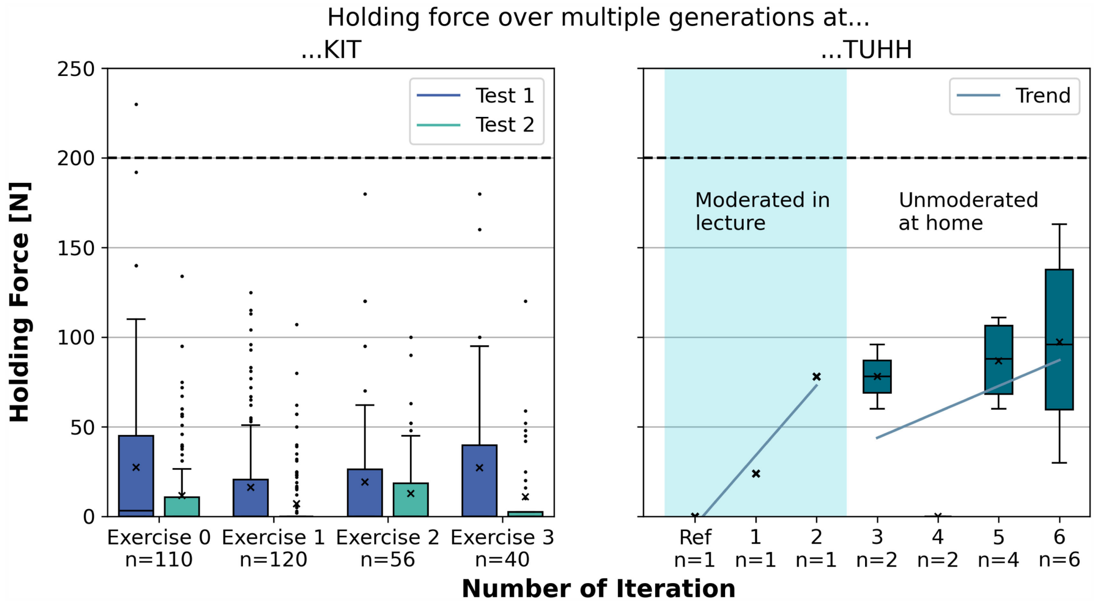

Figure 4. The resulting measured forces over multiple iterations at both universities. At TUHH, the frst two generations were done live in the lecture. Afterwards at both universities, students were able to hand in their own designs over the course of multiple lectures

becomes feasible to investigate the infuences on decision-making in a quantitative and realistic manner. Our study demonstrates that measuring design activities in real time within large groups is not only possible but also provides valuable insights into the design process. Furthermore, preliminary indications of the impact of nudging techniques on decision-making during the task were observed. However, due to the nature of this study, which is not an experimental design, only indices of effects can be reported, as the presence of numerous confounding variables limits the ability to draw defni-tive conclusions. Furthermore, it is essential to acknowledge that the present investigation is confned to the mechanical domain of design tasks. Consequently, additional research is required to elucidate the infuences on decision-making across other design domains.
The comparison between the two implementations in Table 1 demonstrates differences beyond the scope of the designed study. As students at KIT were designing their snap-ft connectors during the exercise, a higher number of designs was submitted in comparison to TUHH, where students were designing outside of the lecture after iteration 2 and were able to hand in designs in groups. Since lower performing students are less likely to participate in optional homework, this lowers the signifcance of TUHH results for measuring the success of teaching. Yet, TUHH designs had a higher rate of improvement over the iterations, especially during the moderated part (see sub section 4.3). Moderating the frst iterations forms a supportive structure of the procedure, which was done at TUHH. These weaknesses stem from the specifc study design, not the usage of a large-scale research environment.
The differing approaches at TUHH and KIT offer a valuable opportunity to compare teaching methods and their effects on iterative design learning. This contrast enables future research on the role of structured guidance and nudging, providing deeper insights into effective strategies for large-scale engineering education.

Table 1. Differences between moderated conduction of the large-scale study at TUHH and unmoderated conduction at KIT

	Topic	**Unmoderated** (KIT)	**Moderated** (TUHH)
1	Measuring success of teaching	●	◑
2	Supporting structure of procedure	○	●
3	Number of submitted designs	●	◑
4	Lecture recordings for analysis	○	●
5	Rate of improvement	◑	●

6. Conclusion and Outlook

With this study, we introduced an instrument that can be used to measure which lecture content has contributed to gain of design knowledge and how. In the background, the instrument used enables the low-threshold collection of data that is of interest for design research. This lays the basis for in-depth investigations of method steps in design research, where the high number of probands enable statistical evaluation, which was hindered before. In the context of education, a diverse set of methods can increase students motivation and skillset in engineering design beyond the lecture. Coupled with a competition for the strongest design, students are more likely to engage and gain a positive view on the lecture and topic. As collaboration between students and lecturer can be handled digitally, this method can be scaled to even larger lectures easily. As a measurement method for students' learning success and evaluation, examinations are an appropriate tool.

The ability to measure design activities in large groups in real-time has the potential to facilitate a more nuanced understanding of decision-making processes, which was previously not feasible. This advancement has signifcant implications for engineering design research, as it allows for more accurate, data-driven insights into how decisions are infuenced during the design process. In terms of infuencing decisions, this approach could be developed further to address objectives related to task optimization, understanding decision-making in design, nudgig, production effciency, and economic sustainability. Moving forward, expanding these methods to include more objectives, such as environmental impact and resource effciency, will further enhance the scope and relevance of this research.

Based on these fndings, a follow-up study is planned that utilizes the snap-ft-connector-task to investigate specifc research questions. The focus will be on decision-making in design and the impact of nudging. In addition to workshops and the implementation of nudges, a shared data repository will be established, enabling the integration of design data and test results from both institutions for comparative analysis. Testing machines will automate the measurement of the maximum holding force of the hooks and systematically store the results. This data can be linked with design parameters in the confgurator at both KIT and TUHH, making it centrally available for collaborative research. Thus, comparing the teaching approaches at TUHH and KIT can provide deeper insights into the effects of decision-making in design, nudging, moderation and independent iterative optimization.

References

610.12, I. S. (1990). Ieee standard glossary of software engineering terminology. *IEEE Std* 610.12-1990, pages 1–84.

Albers, A., Bursac, N., Rapp, S. (2017). PGE – produktgenerationsentwicklung am beispiel des zweimassen-schwungrads. *Forschung im Ingnieurwesen*, 81(1):13–31.

Albers, A. Rapp, S. (2022). Model of SGE: System Generation Engineering as Basis for Structured Planning and Management of Development, pages 27–46. Springer International Publishing, Cham.

Albert, A., Bursac, N., Wintergeist, E. (2015). Produktgenerationsentwicklung – bedeutung und heraus-forderungen aus einer entwicklungsmethodischen perspektive. In Stuttgarter Symposium für Produktentwick-lung 2015.

Batora, M., Ognevoj, R., Schöck, M., Ritzer, K., Bursac, N. (2024). Engineering better decisions: Exploring infuencing factors of behavioral economics on SGE - system generation engineering. *In DS 130: Proceedings of NordDesign 2024*, Reykjavik, Iceland, 12th - 14th August 2024, pages 888–897.

Cassidy, A. (2015). 'big science' in the feld: experimenting with badgers and bovine TB, 1995–2015. *History and Philosophy of the Life Sciences*, 37(3):305–325.

Cortright, R. N., Collins, H. L., Rodenbaugh, D. W., DiCarlo, S. E. (2003). Student retention of course content is improved by collaborative-group testing. *Advances in Physiology Education*, 27(3):102–108. Publisher: American Physiological Society.

Dorst, K. Cross, N. (2001). Creativity in the design process: co-evolution of problem–solution. Design Studies, 22(5):425–437.

Engel, A. (2010). Verifcation, validation, and testing of engineered systems. *Wiley series in systems engineering and management*. Wiley, Hoboken, N.J.

Harvey, G., Loftus-Hills, A.., Rycroft-Malone, J.., Titchen, A., Kitson, A., McCormack, B., Seers, K. (2002). Getting evidence into practice: the role and function of facilitation. *Journal of Advanced Nursing*, 37(6): 577–588. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-2648.2002.02126.x.

Holley, L. C. Steiner, S. (2005). Safe space: Student perspectives on classroom environment. *Journal of Social Work Education*, 41(1):49–64. Publisher: Routledge _eprint: https://doi.org/10.5175/JSWE.2005.200300343.

Hubka, V. Eder, W. E. (1990). Design knowledge: Theory in support of practice. *Journal of Engineering Design*, 1(1):97–108.

Jobst, B., Thoring, K., and Badke-Schaub, P. (2024). What can we learn from outstanding designers? the relationship between design expertise and prototyping. *Proceedings of the Design Society*, 4:603–612.

Kugler, T., Kausel, E. E., Kocher, M. G. (2012). Are groups more rational than individuals? a review of interactive decision making in groups. *WIREs Cognitive Science*, 3(4):471–482.

Liewerenz, O., Grauberger, P., Nelius, T., Matthiesen, S. (2023). IDENTIFYING SUCCESSFUL APPROACHES DURING TESTING ACTIVITIES IN ENGINEERING DESIGN. *Proceedings of the Design Society*, 3:2205–2214.

Malik, Z. A., Khan, S. S., Maqsood, M.-e. (2018). Exploring the relationship between student engagement and new pedagogical approaches. *Journal of Eduational Technology Systems*, 47(2):170–192. Publisher: SAGE Publications Inc.

Matthiesen, S. (2021). Gestaltung – prozess und methoden. In Bender, B. and Gericke, K., editors, Pahl/Beitz Konstruktionslehre: Methoden und Anwendung erfolgreicher Produktentwicklung, pages 397–465. Springer Berlin Heidelberg, 9 edition.

Matthiesen, S. Grauberger, P. (2024). KONSTRUKTIONSWISSEN FUR INGENIEURE: Innovative Produkte zielgerichtet gestalten. SPRINGER VIEWEG, Berlin, Berlin.

Meboldt, M., Matthiesen, S., Lohmeyer, Q. (2012). The dilemma of managing iterations in time-to-market development processes. *In 2nd International Workshop on Modelling and Management of Engineering Processes*, pages 127–139, Cambridge, UK.

Roe, B. E. Just, D. R. (2009). Internal and external validity in economics research: Tradeoffs between experiments, feld experiments, natural experiments, and feld data. *American Journal of Agricultural Economics*, 91(5):1266–1271. Publisher: [Agricultural & Applied Economics Association, Oxford University Press].

Runeson, P. and Höst, M. (2009). Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*, 14(2):131–164.

Tahera, K., Wynn, D. C., Earl, C., Eckert, C. (2018). Testing in the incremental design and development of complex products. *Research in Engineering Design*, 47(2):235.

Tiedens, L. Z. Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specifc emotions on information processing. *Journal of Personality and Social Psychology*, 81(6):973–988.