

# Experimental Studies in Design Science and Engineering Design Science – A Repository for Experiment Setups

Carlo Kriesi, Stephanie Balters, Martin Steinert

*Department of Engineering Design and Materials, NTNU Trondheim  
carlo.kriesi@ntnu.no, stephanie.balters@ntnu.no, martin.steinert@ntnu.no*

## Abstract

This paper answers the design science's call for the systematical implementation of research methodology in order to empirically study the principles, practices and procedures of design. As a research methodology, experiments depend on robust study designs that quantify selected dependent variables, independent of time, location, or conducting researcher. However, it has been shown, that experimental science struggles with the repeatability of experiments, as context and participants introduce many unexpected independent and dependent variables to what is otherwise a robust experimental protocol. Crucially, the lack of detailed information in the description of the experimental setup may prevent the fellow researcher to comprehend data outcome along with corresponding cause-effect interpretation and, moreover, may lead to contradicting results when conducting confirmatory studies and/or meta-analyzes.

Grounded in the positive sciences, we strive to minimize this hurdle and present the concept of the *experimental item-mining repository*, targeting to generate a standardized document with detailed information about the experiment. This repository can potentially become mandatory for certain types of journals in the field, so that any published study is properly documented and does not jeopardize the foundations of science. The level of detail required for a successful submission in the repository is very high: from environmental influences, such as noise and temperature, to detailed plans of the rooms, and lists of hardware and software - incl. version number - involved. Along with a call for open source science and the necessity of detailed empirical information, we provide and present the concept of this online repository by providing a design observation example from our own studies. A later version might be incorporated into a new CERN IdeaSquare Journal of Experimental Innovation (CIJ).

**Keywords:** *Design Science Methodology, Open Source Science, Experiment Information, Repository, Repeatability*

# 1 Design Science's Call for Design Methodology and Confirmatory Studies

Design Science (Fuller et al., 1963; Gregory, 1966) and Science of Design (Cross, 2011; Gasparski and Strzalecki, 1990) call for the systematical implementation of research methodology in order to empirically study the principles, practices and procedures of design (Cross, 2011; Bender et al., 2002; Blessing and Chakrabarti, 2009). This 'includes the study of how designers work and think, the establishment of appropriate structures for the design process, the development and application of new design methods, techniques and procedures, and reflection on the nature and extent of design knowledge and its application to design problems' (Cross, 1984). Research methodologies to empirically analyze design activity are proposed to be literature review, observation, interview, case study, participatory design, and experiments, for example (Denzin, 1978; Descomb, 1998; Blender et al., 2002; Blessing Chakrabarti, 2002). With the focus on experiments as empirical method, we aim to generate quantitative (objective) outcome in order to test and derive the impact and effect of the independent variable on the dependent variable. The research question is approached by formulating a hypothesis which is to be falsified (Popper, 2005) in order to generate paradigms (Kuhn, 2012) which eventually form into theories (Lakatos, 1970).

Correspondingly, Lucienne Blessing and colleagues state: 'Unfortunately, many publications do not provide details of their research, such as data collection context and data analysis methods, and validation of the results is rather uncommon' (Blender et al. 2002). Due to this lack of information, fellow researchers are in consequently unable to entirely comprehend and validate the data outcome and proposed interpretation. The basis for empirical scientific discussion is, thus, severely endangered.

In addition to this issue, studies attempting to validate and meta-analyze data, indicate that slight variation in the contextual setting of (1) human-human interaction, (2) machine-machine interaction, (3) human-machine interaction, (4) human-environment, and (5) machine-environment in the experiment, may cause substantial variation in data outcome. Illustrative examples from our field of research of interaction design experiments and affective engineering (Balters and Steinert, 2014, 2015) are among others: simple pictures prime people into voluntarily paying more for their tea (Kahneman, 2011); one vs. two physiological sensors connected to an Arduino microcontroller leading to a difference in sampling rate (Kittilsen et al., 2016) or a different OS does not allow for the same data transfer protocols (Reime et al., 2015); participants performing the same experimental task, yet using different-sized screens (Gilbert et al., 2016); the impact of variation in sunlight onto the data outcome of skin conductance stress measures (Jung et al. 2015); conducting an experiment in the US versus Europe, leading to changes in electromagnetic fields interferences (50 Hz in Europe vs. 60 Hz in the US) (Balters and Steinert, 2015). In order to conduct comparable confirmatory studies and/or meta-analysis, it is, thus, of crucial importance to be, firstly, aware of the existence of confounding variables and, secondly, to give most detailed information of these human-machine-environment contextual settings, in order to allow the fellow researcher to test the dependent variable(s) under the same experimental conditions.

Addressing the issues of detailed information capturing stated above, this paper proposes the concept of an *experimental item-mining repository (EIMR)*. Final output of this repository is a standardized document including detailed information about an experiment. We propose to submit this standardized EIMR-document *supplementary* to the scientific paper, respectively, to make this document accessible for fellow researchers. This will provide standardized, yet

holistic information data in order to (1) allow the fellow researcher to validate empirical results and to, moreover, (2) enable confirmatory studies and meta-analysis.

The main challenge in generating a standardized EIMR-document is the potential complexity of an experiment itself along with the potential great variation between experiments. We therefore aim for a highly adaptive repository, respectively a subsequently generated standardized document, comprising a very broad coverage while still scanning for, and in detail. Additionally and application-wise, we aim for an intuitive usability when guiding the user through the complex matter. Envisioned is a repository with browser-based interface and complementary app support that guides sequentially along the experimental time line, asking the experimenter to short and precisely characterize the details of and the interactions within the experiment. The semi-active, meaning adaptive, property of the application, aims for use-case-friendly and, thus, time efficient benefits. In addition, the EIMR may serve as a sort of guideline and/or checklist, when setting up experiments.

Along with a call for open source research in engineering and design science in order to enable and promote quantitative-data-based discussion in the community, we will present the concept of the EIMR in this paper. In section 2, we will present the logical structure behind the repository, followed by an illustration of the latter via an exemplary application (section 3). A mock-up of the future graphical interface will be shown in section 4. Conclusively, we will call for more transparent science and propose to support the scientific discussion in design science and engineering design science by means of the *experimental item-mining repository* (section 5).

## 2 The Building Blocks of an Experiment

A successful repository needs to provide both: Usability - a user-friendly interface and an easily comprehensible structure, and completeness – the possibility to capture all the details necessary. We therefore use our own experimental setups as case studies from which we deduct insights on the meta-level of setting up experiments and highlight essential components thereof, which in return will be the building blocks of EIMR. Before we go through an exemplary use-case scenario of EIMR (see section 3), we introduce the logic behind the repository in this section.

### 2.1 Component Classes

There are four classes of components that need to be defined in detail:

- Physical Environment: e.g. Building, Rooms, Climate
- Hardware: e.g. Paper, Pens, Sensors, Computers, Screens
- Software: e.g. OS, Programs and their Versions, Resolution Settings
- Individuals: e.g. Participant(s), Supervisor(s), Assistant(s)

#### 2.1.1 Physical Environment

The physical environment is reaching further than just the room we are in. Background noises from cars, construction sites or airplanes need to be described as well. Additionally, all the rooms that are involved have to be characterized in detail. The repository allows for uploading sketches and pictures of the blueprint of a room. Furthermore, details like the height of a room, light sources, direct sunlight, temperature and humidity are asked for. As rooms can be divided for an experiment, one can also define areas of a room, e.g. the material repository, or where the supervisor is located.

### *2.1.2 Hardware*

Any hardware involved in the experiment needs to be described. This can be the writing material that is provided for the participant, or written instructions, but also computers and sensors that are used within the experiment. It is especially important that any electronics are described in detail since other hardware might measure the same effect, however, with e.g. a different resolution, which subsequently creates different results.

### *2.1.3 Software*

Any electronics device is running on software that defines the behavior thereof. It is therefore of high importance that eventual self made codes or commercially available programs that were used are uploaded or described in detail in the repository.

### *2.1.4 Individuals*

As mentioned in the introduction, individuals are the most crucial element of uncertainty within an experiment. However, one can still try to capture as much as possible, such as their roles within the experiment. Furthermore, especially the recruiting background of the participants is of high interest.

## **2.2 Interactions**

Furthermore, and most importantly, the user can define the distinct interactions that are part of the experimental protocol. One can define at what time the supervisor instructs the participant to engage in a certain action and one can subsequently define with what elements of the setup these interactions takes place.

One can distinguish between three classes of interactions:

- Individual-Individual Interactions: e.g. Receiving instructions orally
- Individual-Object Interactions: e.g. Entering a number in a computer
- Object-Object Interactions: e.g. A sensor sends data to a computer

### *2.2.1 Individual-Individual*

Throughout a design study the participating individual may get information or objects through another individual. However, it is important to distinguish between interactions that are a result – such as interactions between individuals that work as a team – and actions that are an input – such as helping a participant to place a sensor on their body.

### *2.2.2 Individual-Object*

Part of the experimental protocol can be that the participant has to interact on demand with an object, such as a computer or a sheet of paper. It is important to capture the moment of when such an interaction is demanded. The way it is then performed is not part of the controllable experimental setup.

### *2.2.3 Object-Object*

A typical object-object interaction is the connection between a display and a computer. While the software defines the outcome, the physical connection and subsequent display of information can be considered as an interaction. This type of interactions also lets the user define the exact setup, e.g. which display stands on what table but also how a sensor is transferring its data, e.g. Bluetooth or direct cable connection.

## 2.3 Data Handling

Creating a repeatable and robust experiment does not stop when the protocol is done, as one of the most important parts is still outstanding: Post-processing. By defining the interactions on the timeline as described above, the user also has the chance to define what data streams are measured, stored, and used as results. They can then define the software used for eventual calculations and statistical analyses and subsequently upload relevant code structures.

## 3 Exemplary Application: ‘Mockpit’-Experiment

In order to highlight the potential complexity of design studies, we present an excerpt from a study that was conducted by our research group. The forth-following details of the experimental setup are taken from Kittilsen et al. 2016. As mentioned above, we used our own experiences as initial steps in order to create the logic behind the repository that is described in section 2.

The general context of the experiment was addressing the interaction design in a ship bridge scenario, with the focus on stress level reduction. As ‘stress’ is a very vague concept, our group set up an experiment in order to better understand the phenomenon and figure out how stress can be triggered within a controlled environment, and how one can subsequently measure the effects thereof. Throughout the experiment, the participants were hooked up to various physiological data sensors recording, amongst other things, their heart rate, breathing rate, and skin conductivity. They were then asked to solve three different tasks within a mock-up ship bridge (‘Mockpit’) where they controlled various vessels throughout three situations within a boat-simulator:

- Trial run: The participant is given five minutes to get used to the software mechanics by moving a vessel around an environment provided by the software.
- Cruising task: Similar to the trial run, the participant has no other objective than enjoying a cruise-ride within the simulation for five minutes. The aim of this task is to create a non-stressful environment for the participant.
- Race task: In order to provoke stress in the participant, they are asked to conquer an increasingly difficult racetrack on a vessel that is highly difficult to control. Furthermore, the participant is simultaneously asked to solve calculations shown on a second screen.

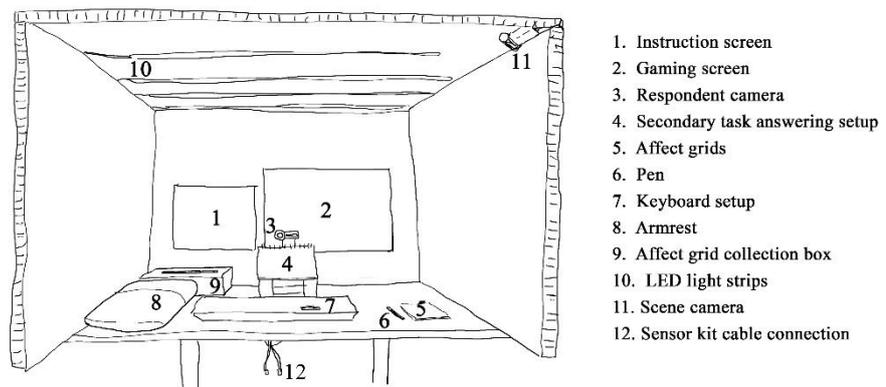
Since this experimental setup involved multiple rooms, supervisors, sensors, computers, as well as a complex physical setup itself, we use this example to highlight the level of details of what is normally described in a publication, and the shortcomings thereof. More specifically, the race-task is analyzed on two different levels, in the context of a descriptive text as it can be found in a publication, as well as the EIMR table that aims to comprise all (detailed) experiment information along the logic of the repository.

The aim is not to discredit any published setups or question their validity. The goal is to show that due to limitations, such as maximum page numbers, and the high complexity, one cannot fully capture all interactions and details. Furthermore, this analysis shows that the combination of text, with its sequential structure, and the level of details contained in a table can create a powerful repository that is easy to use.

### 3.1 Descriptive Text

For the race task the participant, who was wearing earmuffs in order to reduce influences from ambient noise, was seated in the Mockpit, which was set up according to their dominant hand. Centrally in the Mockpit there were two screens, where the smaller one was used to display instructions. Throughout the task, a total of five physiological data sensors recorded

various types of signals. Before and after the race, the participant was asked to fill out a questionnaire. Figure 1. shows the setup for a right-handed participant. The task for the participants was to conquer a premade course within the software called ShipSimulator 2008 by VSTEP (VSTEP, Rotterdam, NL). The course itself took place in the “Atlantic Ocean” environment provided by the software. Various obstacles, such as ships, ramps, and icebergs created a challenging obstacle course. Additionally, the weather conditions became gradually worse whenever the vessel passed checkpoints that were placed in constant, fixed distances within the course. Simultaneously, the participant was instructed on the smaller screen to perform simple calculations and write down the results. These calculation tasks appeared for six seconds and in intervals of 24 seconds. In-between the screen was blank. In order to further stimulate a competitive mind-set of the participant, the winner was promised a 500NOK gift card for the university cafeteria. Furthermore, the instructions said that the scores would be published on the class homepage, once everybody has participated.



**Figure 1. Sketch of the Mockpit setup as it was used in the experiments (setup for right-handed participants). From Kittilsen et al. 2016 (in press)**

### 3.2 Listings

While the descriptive text above gives the reader a good idea about the setup and the types of interactions, many crucial elements can simply not be described in high enough detail. Repeating the experiment along these instruction could potentially result in a very different outcome. Below are two tables, Table 1. and Table 2., that contain a detailed list of the objects according to the classes defined in section 2, as well as a list of interactions. While the descriptions may refer to drawings or code files, this is purely for descriptive purposes within the context of presenting the general idea. Furthermore, the list of interactions is incomplete due to space restrictions.

#### 3.2.1 Objects – Race Task Phase

**Table 1. List of objects during the race task according to the four classes of the repository (incomplete)**

<b>Physical Environment: Room 2</b>	
Area	<i>3rd floor within university building</i>
Background Noise	<i>None in particular</i>
Climate	<i>Thermostat controlled 22°C, humidity of 46%</i>
Floor Plan and Layout	<i>Room 2 is separated in the experimental ship bridge replica, made of cardboard and a supervisor area that is not visible to the participant.</i>

	<i>The drawings give a detailed overview, as well as the pictures.</i>
--	--

### **Hardware**

Earmuffs	<i>Industrial grade earmuffs that are used for noise dampening</i>
Screen 1	<i>Secondary screen of the Mockpit, 17" LCD monitor, 1280x1024</i>
Screen 2	<i>Primary screen of the Mockpit, 24" LCD monitor, 1920x1080</i>
Pen 1	<i>Standard BIC roller pen, blue</i>
PC 1	<i>2x3.60 GHz CPU 16 GB RAM</i>
PC 2	<i>2x3.60 GHz CPU 16 GB RAM</i>
Keyboard 1	<i>Logitech Keyboard, only arrow keys visible, rest covered</i>
Microcontroller 1	<i>Arduino UNO, ATmega328P chipset, version of 2015</i>
Ethernet Shield	<i>Arduino Ethernet Shield R3</i>
Sensor Shield 1	<i>Libelium eHealth shield, v2.0</i>
Sensor 1	<i>Libelium ECG, version of 2015</i>
Sensor 2	<i>Libelium airflow sensor, version of 2015</i>
Sensor 3	<i>Libelium temperature sensor, version of 2015</i>
Sensor 4	<i>Libelium skin conductivity sensor, version of 2015</i>
Sensor 5	<i>Libelium Accelerometer, version of 2015</i>
WebCam 1	<i>Creative 73VF068000001, 1080p resolution</i>
WebCam 2	<i>Creative 73VF068000001, 1080p resolution</i>
LED Strips	<i>IKEA dioder</i>

### **Software**

Arduino Code	<i>Arduino Code used, reading sensor inputs.</i>
ShipSimulator	<i>ShipSimulator 2008 v. 1.4.2, VSTEP, configuration according to settings file</i>
iMotions	<i>iMotions Attention Tool v 5.4</i>
OS	<i>Windows 7, SP3</i>

### **Individuals**

Participant	<i>Recruited amongst mechanical engineering students (n=34) and employees within engineering department (n=6), 26 male, 14 female, age 23-28</i>
Experimenter 3	<i>Third experimenter, male, age 24</i>

#### 3.2.2 Interactions – Race Task Phase

**Table 2. List of interactions identified within the experimental task phase (incomplete). Types of interactions: ‘I-I’ – Individual-Individual; ‘I-O’ – Individual-Object; ‘O-O’ – Object-Object**

<b>Type</b>	<b>Location</b>	<b>Objects</b>	<b>Description</b>
I-O	Mockpit	Participant; Earmuffs	<i>Participant is wearing the earmuffs</i>
I-O	Mockpit	Participant; Sensors 1-5	<i>Participant is wearing the sensors 1-5 according to positions shown in drawings</i>
I-O	Mockpit	Participant; Keyboard	<i>The participant only has access to the arrow keys and uses them to control the vessel within the program</i>
I-O	Mockpit	Participant; Screen 2	<i>Screen 2 displays the ship simulator to the participant</i>
I-O	Mockpit	Participant; Screen 1	<i>Screen 1 displays instructions to the</i>

			<i>participant</i>
I-O	Supervisor Zone	Experimenter 3	<i>The experimenter 3 is responsible for triggering the instructions / slides at the right time</i>
O-O	Mockpit	Sensors 1-5; Sensor Shield	<i>The sensors are wired to the sensor shield according to the diagram</i>
O-O	Mockpit	Sensor Shield; Microcontroller 1	<i>The sensor shield is attached to the microcontroller</i>
O-O	Mockpit	PC 1; Microcontroller 1; Ethernet Shield	<i>The microcontroller is connected to PC 1 via the Ethernet shield and a LAN cable</i>
O-O	Mockpit	PC 1; Screen 1	<i>PC 1 is connected to Screen 1, running it at 60Hz, resolution of 1280x1024</i>
O-O	Mockpit	PC 2; Screen 2	<i>PC 2 is connected to Screen 2, running it at 60Hz, resolution of 1920x1280</i>
O-O	Mockpit	PC 2; ShipSimulator	<i>ShipSimulator is running on PC 2</i>

### 3.3 The Temporal Dimension

A static list such as the tables above increases the level of detail when it comes to describing objects and experimental setups. From a usability standpoint of view, however, it is not a suitable solution. The repository needs to follow a timeline that lets the user create two types of interactions along the timeline: On-going interactions, such as the constant wearing of sensors, or the permanent connection between a computer and a screen; And momentary interactions, such as displaying a mathematical problem for six seconds or an input from the experimenter at a specific point in time.

## 4 GUI/Interaction Design

Based on the conclusions from the case study in section 3, we present a mock-up graphical user interface for EIMR. It is by no means the final design, nor is it a functional software yet.

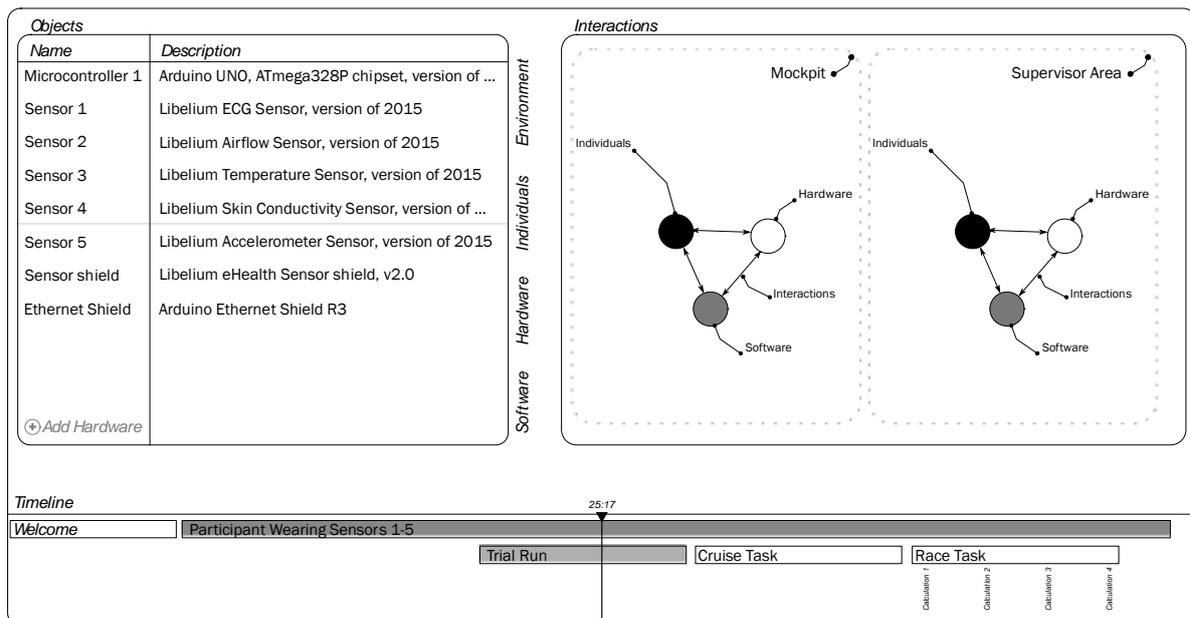


Figure 2. Conceptual GUI of the repository.

It is merely to give the reader an idea of the concept. One concept is shown in Figure 2.

#### 4.1 Key elements

The following key-elements create the main user interface of the repository:

- Timeline: Similar to the timelines that are known from e.g. video editing softwares
- Objects: As defined in section 2, the repository defines four classes of objects. One can easily define and add objects to a list and search for pre-defined objects.
- Interactions: By connecting two objects one can create interactions along the timeline, similar to the block diagrams known from e.g. SimuLink (MathWorks, MA, USA)

### 5 Sharing the experimental item-mining repository with and for the community

With this paper we propose an answer to the design science's call for the systematical implementation of research methodology, in order to empirically study the principles, practices and procedures of design. The special focus lied herein on the research method of experiments, which aim to generate quantitative (objective) outcome in order to test and derive the impact and effect of the independent variable onto the dependent variable. Prominent fellows of the community criticize the lack of information given in many publications, "such as data collection context and data analysis methods, and validation of the results is rather uncommon" (Blender et al. 2002). In the field of experimental design and engineering design research, we are additionally facing the risk that context and participants introduce many unexpected independent and dependent variables to what otherwise is a robust experimental protocol. Due to the lack of given information, as well as the lack in detail of given information, fellow researchers are consequently unable to entirely comprehend and validate the data outcome and proposed interpretation. The basis for empirical scientific discussion and the conduction of comparable confirmatory studies and/or meta-analysis in particular, are ergo severely endangered. Given these points, the *experimental item-mining repository* aims to interrogate all crucial information about the experiment. The targeted outcome is a standardized document including *all* information of the conducted experiment. We propose to submit this standardized document supplementary to the scientific paper in order to make information accessible for fellow researchers.

In this paper, we proposed the concept of the *experimental item-mining repository*, initially structured based on the decomposition of an experiment example from our own studies. The next step is to systematically integrate more experimental setups, as well as focusing on external user inputs, and adapting the EIMR structure accordingly. Subsequently, the actual online *experimental item-mining repository* will be constructed, generating the standardized EIMR document. Following the example of IBM, using the method of Creative Commons (<https://creativecommons.org/>) to crowd-source the adaption and improvement of JAVA scripts yet with super-visionary approval rights, we aim to share the EIMR platform for dynamical improvements in the future with, and for the community. We call for transparent science and aim to support the scientific discussion in design science and engineering design science by means of the *experimental item-mining repository*.

#### Acknowledgement

This research is supported by the Research Council of Norway (RCN) through its user-driven research (BIA) funding scheme, project number 236739/O30.

## References

- Balters, S. & Steinert, M. (2015). Capturing emotion reactivity through physiology measurement as a foundation for affective engineering in engineering design science and engineering practices. *Journal of Intelligent Manufacturing*. DOI 10.1007/s10845-015-1145-2
- Balters, S. & Steinert, M. (2014). Decision-Making in engineering—A call for affective engineering dimensions in applied engineering design and design sciences. *International Conference on Innovative Design and Manufacturing (ICIDM)*. IEEE; Montreal. 2014-08-13 - 2014-08-15.
- Bender, B., Reinicke, T., Wünsche, T., & Blessing, L. T. M. (2002). Application of Methods from Social Sciences in Design Research. *International Design Conference. Dubrovnik*. Croatia.
- Blessing, L. T. M., & Chakrabarti, A. (2009). DRM, a Design Research Methodology. *Springer Science & Business Media*.
- Cross, N. Cross, Nigel (2001). Designerly ways of knowing: design discipline versus design science. *Design Issues*, 17(3) pp. 49-55.
- Cross, N. (1984). *Developments in Design Methodology*. Wiley, Chichester
- Denzin, N.K. (1978). *The Research Act*, 2d ed. McGraw-Hill, New York.
- Descomb, M., (1998). *The good research guide*. Open University Press, Ballmore
- Fuller, R. B., McHale, J., (1963). World design science decade, 1965-1975; five two-year phases of a world retooling design proposed to the International Union of Architects for adoption by world architectural schools. Carbondale: World Resources Inventory Southern Illinois University.
- Gasparski, W, and Strzalecki, A. (1990). Contributions to Design Science: Praxeological Perspective. *Design Methods and Theories*. 24, 2
- Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on ‘Estimating the reproducibility of psychological science’. *Science*, 351(6277), 1037–1037.
- Gregory, S. A. (1966). S A Design Science, in Gregory, S A (ed) *The Design Method*, Butterworth, London (1966).
- Jung, M. F., Leifer, L. J., Gross, J. J., Hinds, P., & Steinert, M. (2015). Displayed Uncertainty Improves Driving Experience and Behavior: The Case of Range Anxiety in an Electric Car. In CHI 2015, April 18 – 23, 2015, Seoul, Republic of Korea.
- Kahneman, Daniel. (2011). *Thinking, Fast and Slow*. Macmillan.
- Kuhn, T. S. (2012). *The Structure of Scientific Revolutions: 50th Anniversary Edition*. University of Chicago Press.
- Lakatos, I. (1970). History of Science and Its Rational Reconstructions. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1970, 91–136.
- Leikanger, K. K., Balters, S. & Steinert, M. (2016, in press). Introducing the Wayfaring Approach for the Development of Human Experiments in Interaction Design and Engineering Design Science. *International Design Conference Design 2016*, Dubrovnik, Croatia.
- Popper, K. (2005). *The Logic of Scientific Discovery*. Routledge.
- Reime, T., Sjöman, H., Gerstenberg, A., Abrahamsson, P., & Steinert, M. (2015). Bridging Tangible and Virtual Interaction: Rapid Prototyping of a Gaming Idea. In K. Chorianopoulos, M. Divitini, J. Baalsrud Hauge, L. Jaccheri, & R. Malaka (Eds.), *Entertainment Computing - ICEC 2015* (Vol. 9353, pp. 523–528). Cham: Springer International Publishing.