Teaching Energy Storage Systems in Laboratories:

Hands-on versus Simulated Experiments

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy

Fabian Steger

B. Sc. (UAS Regensburg, Germany)
M. Eng. (UAS Regensburg, Germany)

School of Engineering
College of Science, Technology, Engineering and Maths
RMIT University

in cooperation with
Faculty of Electrical Engineering and Computer Science
Technische Hochschule Ingolstadt

March 2020

Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

Fabian Steger
Ingolstadt, 31 March 2020

Acknowledgements

I would like to thank everyone who has contributed to the success of the work:

Firstly, this research would not have been possible without the students who generously agreed to take part in the study in order to improve the learning of future colleagues.

I would like to express my sincere gratitude and heartfelt thanks to my senior supervisors *Iouri Belski* and *Lasantha Meegahapola* at RMIT for their trust, motivation, and their insights and continuous inputs during the work, and to the associate supervisor *Hans-Georg Schweiger* at THI for providing the environment to conduct the research, his guidance, and help throughout this PhD project.

Many thanks go to *Alexander Nitsche* for creating the Java program to control the student battery cell test system, for being ready to enhance his software with the necessary simulation model, and for supporting the "hidden simulations" mode with his software.

I would like to thank *Alexander Arbesmeier* for a lot of interesting discussions, and help with statistics and language, *Katja Brade*, *Bhavya Kotak* and *Iuri Ferreira* for investing time to read and revise drafts of different publications and *Moritz Igel* for taking time out of his schedule to review the language of the final work.

I wish to thank laboratory engineer *Sönke Barra* for his assistance in planning, preparing, and conducting the laboratory experiments.

Thanks to the dean of the Faculty of Electrical Engineering and Computer Science *Wolf-Dieter Tiedemann*, the study program managers *Robert Hermann* and *Tobias Schrag*, and the remainder of my colleagues in the faculty of THI for always being supportive and granting me the freedom necessary to conduct the research.

Parts of the study would not have been possible without the support of the THI administration, especially *Michaela Regler*, Vice President for Teaching, Students and Alumni, *Katrin Michl*, *Paola Frignani* and *Kathrin Hahn*.

During the writing of the thesis, and next to the conducting energy storages laboratories, I also performed technical research, and offered a lot of project-based learning to students of several programs. For this I want to thank *Walter Strasser* and *Sönke Barra* for letting me loose on their laboratories, their support of (most of) my projects, and being tolerant of smoke, annoying sounds, and crowds of students. *Adriano Peniche dos Santos*, *Alexandra Hendry*, *Camilla Bendsen*, *Cayler Miley*, *Christoph Nebl*, *Despoina Pourtoulidou*, *Enrique Machuca*, *Jonas Lausch*, *Jonathan Krogh*, *Peter Knudsen*, and *Woong-Ki Kim*: It was a pleasure to be in contact with people from so many different countries and to perform research with you.

Last but not the least; I would like to thank my wife, son, family and friends for their patience and supporting me throughout the duration of the project.

Funding

As part of the German Federal Government's Showcase Regions for Electric Mobility, the *Federal Ministry of Education and Research* provided funding for the project "Academic Education Initiative for E-Mobility Bavaria - Saxony" which was used for the development of the prototypes of the devices.

The devices used in the laboratory were built with financing from the *Faculty of Electrical Engineering and Computer Science* at Technische Hochschule Ingolstadt

A pilot of the English version of the laboratory was partially financed by the *EU "Erasmus+ staff mobility for teaching" program* and the *Faculty of Electrical Engineering and Computer Science* at Technische Hochschule Ingolstadt.

The guest laboratory trial included in the study was partially financed by the *IEEE Student Branch of Chemnitz.*

The *IEEE Transactions on Education* publication was supported by the Post-Grand-Fund Program of the *German Federal Ministry of Education and Research* under grant 16PGF0164.

Many thanks for the support through a scholarship I received in the first years under the *Collaborative PhD Agreement* between THI and RMIT.

Table of contents

List of figures

List of tables

Terminology and abbreviations

Table 1: Educational terminology and abbreviations

·	Abitur	Germany's standard secondary education degree (after 12 or 13 years of school), for details see section 3.1
APE	Amount of Practical Experience	as defined in section 4.3
ECTS	European Credit Transfer and Accumulation System	[1]
STEM	Science, technology, engineering, and mathematics	German: "MINT"
SLM	Superior learning mode	variable to compare the effectiveness of two learning modes regarding a single student, as defined in subsection 4.2.7
TP	Test performance	standardised test result of a student (relative to the students writing the same test in the same run), as defined in subsection 4.2.7
THI	Technische Hochschule Ingolstadt	University of Applied Sciences Ingolstadt, Germany
TUC	Technische Universität Chemnitz	Technical University Chemnitz, Germany
UAS	University of Applied Sciences	A German type of university, see page 38 for details
VET	Vocational Education and Training	see page 36 for details
R	Study run	A sequence of sessions with the same participants
CA	Content area	Grouped topic on energy storages, taught in one or two sessions
·	Laboratory session/lesson	A meeting in the laboratory / computer-lab (e.g. 3 h)
·	Student laboratory experiment	An experiment to investigate a specific topic (e.g. internal resistance)
·	Experimental sequence	A current/voltage profile programmed by students and automatically executed by laboratory equipment or simulation

Table 2: Electrical terminology and abbreviations

AC	Alternating current	
ADC	Analog-to-digital converter	
CID	Current interrupt device	pressure valve at a battery, which permanently disables the cell if its internal pressure is too high
DAC	Digital-to-analog converter	
DC	Direct current	
DUT	Device under test	
FFT	Fast Fourier transform	
$LiFePO_4$	Lithium iron phosphate	
$LiMn_2O_4$	Lithium-ion manganese oxide (spinel)	
LTI	Linear time-invariant	
OCV	Open circuit voltage	Voltage under no-load condition
PCB	Printed circuit board	
PTC	Positive temperature coefficient (resistor)	A resistor increasing the resistance when heated
R_I	Internal resistance	
SoH	State of health	
SoC	State of charge	

Abstract

Background

Engineering courses often complement lectures with laboratory classes to optimise student learning outcomes and further develop valuable skills for future employment. Computer simulated experiments for conducting laboratory exercises have become increasingly popular in higher education and vocational training institutions to replace traditional hands-on laboratories. Reasons for this include for example, cost efficiency and repeatability.

Research question

There has been a wide array of discussion on the efficacy of the two laboratory modes in teaching, both in general and for students in engineering fields (for example, chemical engineering or electrical engineering). However, many previous studies on this question did not reach a universally valid conclusion. The used methodologies mixed other influences with the impact of the investigated learning modes. These influences include for example accompanying lectures, experimental instructions, teachers, learning objectives, tests, working teams, and many more. Thus, the differences in results of these studies cannot be attributed to the laboratory mode only.

The study conducted for this thesis investigated differences in learning outcomes of students in higher education when comparing two laboratory modes in the local domain:

- In-person hands-on laboratories allow students to directly interact with the subject at hand, although this interaction might be mediated through technology or a user interface.
- In-person simulated laboratories moderate all student interactions through a user interface. The properties of the investigated effect are simulated by computer software. The students work in a classroom equipped with computers on which the simulations are running.

Since this study was focused solely on comparing different learning modes, all other aspects were held as constant as possible. Improvements that were theoretically possible in only one of the teaching methods (e.g. time-lapse in simulations) were not implemented in order to keep the surrounding conditions as equal as possible. Thus, the aim of the research was *not* to determine which of the investigated laboratory modes

would be best for teaching a specific topic, but rather to investigate whether or not there are discernible differences in teaching success when conducting the *same* experiment in hands-on and simulated laboratories. The ultimate goal was to establish more reliable and generalisable insights into the influence of a particular laboratory mode on learning.

The study did not include a remote laboratory condition; the comparison was only made between in-person laboratory teaching with proper laboratory equipment and simulations conducted in the local domain. An important note on demographics: a third of students at universities of applied sciences have completed apprenticeships in the German vocational education and training programs (VET) before enrolment. These VET programs mainly consist of practical on-the-job learning and aim to directly prepare apprentices for entering the job market. Due to the large size of this demographic and their previous experiences mostly with hands-on learning, it was of additional interest to see if VET-participants' results differed significantly from those of their peers when confronted with the two laboratory modes. It was also of interest to see if the *perception* of the learning modes influenced the outcome.

Methodology

This study was conducted in *two* consecutive phases on the example of a practical course teaching the basics of batteries (not related to physical manipulation of the batteries). A counterbalanced within-subject methodology was employed with German and international participants in nine study runs. The laboratory modes alternated, while the learning objectives and the experimental approach of laboratory exercises remained practically identical.

In the *first* phase, the objective was to compare students' learning success when working with hands-on laboratories and with *overt* computer simulations, respectively. The *second* phase was conceptualised to give insight into possible subjective influences of students' perception of the two laboratory modes. In this phase, the simulation condition was *hidden*. Participants used hands-on equipment in both conditions. In the first condition, real measurements were shown; in the second condition, hands-on devices displayed simulated battery behaviour to investigate the influence of students' perception. The participants were *not* aware of the differences in data sources.

Besides the comparison of knowledge test results, questionnaires were employed to correlate prior, specifically technical, practical experience and previous apprenticeship training with the success of the knowledge transfer in both of the compared modes. Well-known personality tests were also employed in order to provide further insight into the subjects.

The study collected subjective opinions regarding the laboratory modes in two ways:

- Participants of the main study were asked to provide feedback after conducting a laboratory experiment. This method allowed for the *indirect* gathering of information about the difference in perception towards the two modes.

- Persons who had either not yet started the laboratory or weren't participating in the laboratory were asked to fill out a general questionnaire distributed amongst different universities in different countries. This method asked *directly* for subjective opinions regarding the learning modes.

Finally, the THI university database was analysed to extract objective information about students with and without vocational training degree to gain broad background information about the compared groups.

Outcomes

In the first phase, it was found that there were statistically significant differences in learning outcomes favouring the hands-on mode. When the simulation condition was overt, students with a background in vocational training before enrolment showed statistically significant trends towards better learning with hands-on experiments. Students in the international runs and Germans without a VET background performed similarly in both modes.

In the second phase, when students were not aware that they were using simulations, both modes showed similar student learning across all student groups.

Generally, simulations were reported as less relevant and their authenticity was called into question.

A VET background seems to determine whether or not students had different levels of success in hands-on and simulated laboratories. As hidden differences in the simulations could be excluded from having been the reason for inferior learning results, psychological effects needed to be considered to comprehend the different laboratory modes' effectiveness.

The study outcomes lead to the conclusion that students' personal perception of the laboratory modes, particular simulations, can have a significant impact on laboratory learning.

Chapter 1

Background

To optimise outcomes of student learning and to develop valuable skills for future employment, engineering courses often complement lectures with laboratory classes [2–4]. Linking theoretical learning (based on teacher-centred lecturing approaches) with laboratory experiments is particularly relevant at German Universities of Applied Sciences, which attach great importance to practice-guided learning [5]. The particular educational value of experimentation lies in involving the students actively.

Equipment for hands-on laboratory experiments as well as laboratory supervision of classes that require physical equipment can be costly [6, 7], especially when dealing with potentially dangerous materials such as lithium-ion battery cells [8]. Simulated experiments have gained popularity for laboratory learning in universities for purposes of education (e.g. possible time manipulation), logistics (e.g. risk minimisation, repeatability) and cost efficiency, see [9], [7, p. 332]). A direct comparison of the teaching efficacy of simulated and hands-on experiments is necessary to avoid a deterioration in learning quality.

Over the years, there has been a wide variety of research into the efficacy of laboratory modes and advantages of different laboratory modes for teaching in general and for teaching engineering students in particular [9, 10].

A recent trend towards virtual and remote laboratories can be seen in publications on laboratory learning [9, 11].

One goal of implementing laboratories is to meet various learning objectives (e.g. instrumentation, experimental approaches, data analysis, safety and teamwork). The exact objectives are usually set by study program or accreditation rules, with teachers being afforded varying degrees of freedom.

In many of the compared studies, the research objectives were not clearly formulated. In other cases, comparisons were drawn between studies where research goals did not match [12]. Most studies on the subject of knowledge and understanding gained from laboratory work concluded that student learning was either constant or improved when hands-on laboratories were replaced by or complemented with computer-based laboratories [10].

However, in many previous studies the number of analysed participants was too small or no statistically significant result was found. Additionally, the relative effectiveness of different kinds of laboratories was seldom explored. Many studies in this field did not use strong methodological approaches, making it difficult to attribute their results to the influence of laboratory modes alone [10, 13].

Some scholars did not regard different learning modes as directly rivalling solutions for the *same* educational objectives, but instead tried to achieve different study goals, thus developing and optimising each mode independently [14–21]. Doing so, other important influences on student learning were often ignored in the analysis. Participants from experimental and control groups learned under considerably different conditions (see [14] and section 2.6).

The study conducted for this thesis follows a different methodology. The research questions and methodology were developed based on the literature research presented in chapter 2.

1.1 Statement of problem

A direct isolated judgement of the influence of learning mode seems to be difficult without keeping all other influencing factors constant – for such a comparison, purely focused on the modes, all these other blurry influences need to be optimised to a similar extend for both laboratories. This defined similar extend is difficult to describe and even more difficult to achieve. Thus, it is often difficult to judge the influence of these other influences on the results.

If these interfering aspects have stronger influences, the studies are unable to specifically identify the difference in learning effectiveness of one aspect [13]. This may explain the inconsistent outcome of present research [10, 14, 22].

The goal of this study, therefore, was to validate the possibility of comparing laboratory modes by keeping constant as many influencing factors (such as learning objectives, teachers, cooperative learning, learning synchronicity, guidance, etc.) as possible. To provide reliable insights, this study employed an optimised research methodology to avoid other effects on student learning during laboratory work, all excluded influences are listed in the appendix, see page 239. The experiments were designed based on the same learning objectives in a way that they could be conducted in both modes in the same manner. Thus, only a single set of instructions, used in both laboratory modes, was necessary for each experiment.

1.1.1 Learning modes compared in this research

This research focuses on differences in student learning by comparing hands-on and simulated laboratories.

- In *Practical hands-on exercises* (see Figure 1.1), students directly interact with the subject at hand and check the equipment, although this interaction might be mediated through technology or a user interface. Results/values of the experiment are real measurements derived directly or indirectly from a physically existing specimen.

- In a *Computer-based simulated laboratory* (see Figure 1.2) on the other hand, all of the students' interactions are moderated through a user interface (including all interactions students use to create their understanding of the experimental hardware). The properties of the investigated effect or sample are simulated by computer software. Participants work in a classroom equipped with computers, which process the simulation model.

In this research, both modes are compared in the local domain (in-person laboratories). The physical experiments (using a real battery and a real battery test bench) performed by the students in traditional way were called *hands-on*, while the virtually performed experiments were called *simulations*. In the second research phase, the simulation's performance was hidden, and students thought they were performing real experiments. This mode was called *hidden simulations*.

Remote experimentation was not investigated

Since the study has focused on a strict methodological approach, a *remote* laboratory condition was not included, even though many students favour online education [23] and recent literature reports equal or better learning with remote laboratories [9–11, 24–26].The study was designed to compare in-person laboratory teaching with/without proper laboratory equipment solely in the *local* domain.

1.1.2 The potentials of a strict methodology

On the one hand, one needs to look at the fact that teachers work under given circumstances. Accreditation requirements are most times clear as to what knowledge or capabilities laboratory classes should transfer.

Laboratory classes themselves will operate in a constrained environment, which may imply many named confounding factors. As the proposed research methodology targets a cleaner study environment, one can criticise that some of the compared arrangements are pure of artificial nature. So it might be not clear why such a constrained comparison is helpful.

On the other hand, teachers have many degrees of freedom. For example, in the module the main study relies on, the curriculum specifies the learning objectives, the access (local) but not details the methods used during laboratory teaching.

A higher number of free decisions (differences between conducted modes) makes it impossible to clarify the influence of one of them [27, p. 395f].

Even when judged entirely artificial, much research is done in unnatural environments to gain insights. For example, a batteries capacity is usually investigated by a constant current discharge, to be comparable – even when there are no real-life devices consuming constant current.

That research might be not in the same intense directly applicable as other comparisons, but targets for more confound, generalisable and comparable insights.

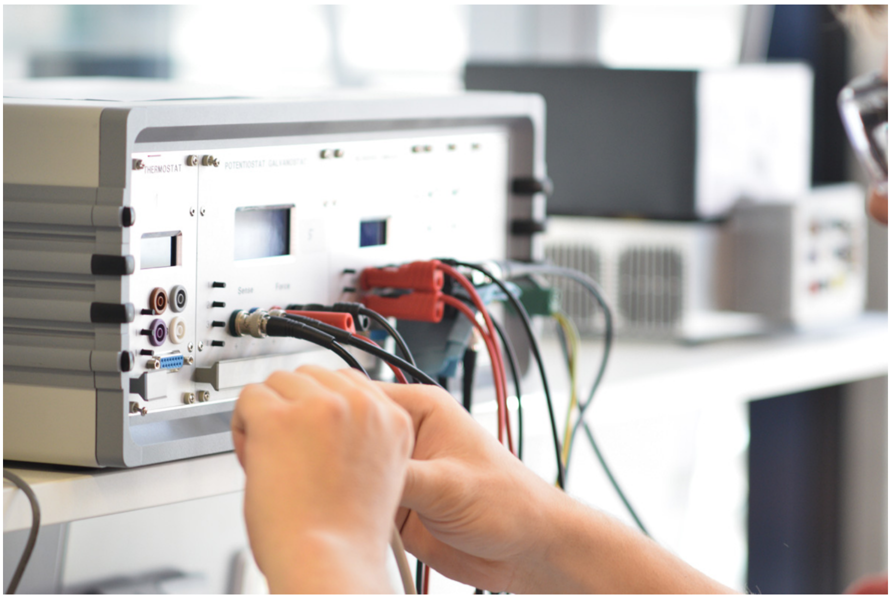

Figure 1.1: Hands-on exercises

Figure 1.2: Simulated experiments in a computer laboratory

Nevertheless, both potentially artificial modes can find some applications in real life:

Application of local simulation in the computer pool
Local simulation in a classroom is applicable under two different aspects:

First, teachers can employ them to *train how to simulate* (e.g. using Matlab/Simulink, as described in appendix Workshop E). That white-box usage is not correlated with the present research.

Secondly, in the main study, black-box simulations emulated physical laboratory experiments to *transfer knowledge about* the simulated object. Advantages compared to hands-on experiments were presented before, like potential cost-savings and safety aspects. Even if there are many advantages of remote teaching, some circumstances may lead to preferring to perform these black-box simulations *locally* in a computer pool at university. For example, when not using web-applications, pre-installed software on a computer-lab PC is functional directly from the start and can be tested before. Students do not waste time installing run-time environments and application at home, as well as making them running. Secondly, the teacher gets real-time feedback about the students' process, including face-to-face interaction. He/she can arrange a synchronous process of all groups (if desired, for example, instructive phases are planned between steps) and adjust the intensity of supervision while the experiments accordingly.

Application of a local hands-on laboratory with simulated equipment
Performing simulations in the hands-on laboratory environment can contribute to students learning success, assuming the number of available genuine stations is limited. The number of devices might be planned before the number of students rose unexpectedly (for the example the reduction of one year in secondary schooling by law leading to higher number of students in one year), or exceptional circumstances force to reduce working team size (for example, the corona pandemic). In such cases, an available simulated copy of the hands-on experiments enables the teacher to react quickly and flexible in that single run without the effort to change the experimental procedure or instructions. Some groups can use the simulated stations as proxies while being instructed in the same room.

1.2 Research questions

Q1: Is the strict methodology proposed suitable to compare different laboratory modes?

The methodology described in the thesis compares the effectiveness of two learning modes during laboratory work by solely focusing on the learning results of each learning mode. The target is to develop a methodology which allows the determination of systematic effects (e.g. that computer-based simulations are as effective as traditional practical hands-on exercises). In the literature, other aspects (e.g. distance learning, instructions, and learning objectives) are often mixed with the learning mode. The proposed system will reduce the degrees of freedom. Part of the goal of this thesis is therefore to evaluate whether this methodology is appropriate.

Q2a: Are computer-based simulations as effective as traditional practical hands-on exercises in understanding battery laboratory topics?

One of the aims of this study is to evaluate whether knowledge and understanding gained as a result of laboratory work using traditional practical training (using the self-developed battery test system) is equivalent to simulation-based training. The outcome of this study provides the scientific community with more information on whether the increased effort for practical hands-on training is justified by better student learning.

As this study is based on a specific case study with certain fixed parameters (learning objectives, battery topic, culture and discipline of the participants), the results cannot be freely generalised, but provide teachers in comparable teaching situations with a guideline to determine the better learning mode for practical learning.

Laboratories in higher education can have several objectives (e.g. promoting teamwork, encouraging problem-solving and critical thinking), for this study the focus lies on knowledge transfer and understanding of the subject matter.

During the study runs on this research question, students are aware of the mode they used, i.e. whether the experiments were physical or simulated.

Q2b: Perception: Do the student learning results remain the same when students are not aware that they are using simulations?

With the increasing complexity in teaching practical engineering skills via laboratories, borders between the learning modes often become blurry. Therefore, the *perceived* laboratory condition is also of interest when comparing laboratory modes.

The second research phase was added first to determine whether or not there are any significant differences between the two modes if laboratory conditions themselves are indistinguishable and second, to verify the first phase results (e.g. wrong outcomes caused by a lack of essential details in the simulation yet undetected by the researchers).

More study runs of the mode-comparing experiment (simulations vs. hands-on) were performed. In contrast to the first research phase, students working with simulations in the second phase also thought they were performing traditional hands-on experiments (so-called “hidden simulations”). In these simulations, students used custom-made devices meant to look like hands-on devices but modified to display the results of simulations. Thus, the simulations were framed as hands-on experiments.

Q3: Is there a relationship between students' individual qualities, attributes, and educational background, and the more successful learning mode?

This research question was posed to verify the assumption that individual students benefit from hands-on mode in contrast to the simulations mode. Focus was laid on three aspects: the Amount of Practical Experience (DS-B), whether or not a student had participated in the German Dual System of Vocational Education and Training (see subsection 3.2.1) before enrolling, and students' educational background, both before enrolment (how they earned their university entrance qualification) and during their studies (e.g. age of enrolment, marks, duration of studies).

Vocational training
There is a particular demographic which is inherently interesting based on the choice of subjects: approximately a third of students at Universities of Applied Sciences have served apprenticeships in the German Vocational Education and Training programs (VET) prior to enrolment [28, p. 5]. These VET programs prepare apprentices for entering the job market and therefore mainly consist of practical learning. Hence, it was of additional interest to compare the results of these particular participants to those of their peers when confronted with both laboratory modes.

Q4: Do students view the compared learning modes differently?

Besides students' knowledge gain, student opinions and satisfaction are also important. In this study, due to the cross-over-like methodology, students experienced both learning modes and were able to judge them directly. The study asked for students' opinion after conducting the laboratories. Are there differences in student satisfaction? Which learning mode dependant factors are judged as advantageous or disadvantageous by students? Following up, another aspect is of interest: Are there any preconceptions regarding both laboratory modes which influence the performance?

1.3 Thesis structure

The remainder of the thesis is structured as follows:

Chapter 2 provides necessary background information about *laboratory learning and teaching*. Chapter 3 introduces essential background information on the study participants' *educational pathways*.

Chapter 4 describes the employed *research approach and methods*, while the *findings* are presented in chapter 5.

The results are *discussed* and put into context in chapter 6. The *conclusions* drawn from the study are listed in chapter 7; Chapter 8 provides proposals for *further research*.

Finally, the appendices (starting from page 228) describe the *devices and simulations* used to teach the *battery laboratory contents* and *the tests used to examine the knowledge gain*. The employed statistical methods are also explained therein.

Chapter 2

Review on laboratory teaching and learning

This chapter provides background information from educational research studies relevant for the current study.

For a short introduction into the technical background regarding energy storage systems as well as a description of the technical set-up of the laboratory experiments employed in the study, please refer to the appendices starting from page 277.

2.1 The role of laboratories

"Nullius in Verba" ("on the word of no one"), the motto of the British Royal Society (founded in 1660) can be understood as "swear by no one's words" or "take nobody's word for it". The founding of the institute marked a clear break with the philosophy of science that had prevailed up to that point: The motto stands for the declared will to establish an experimentally based science that is not satisfied with quoting authorities and verifies statements through experimentally determined facts. [29]

Following in this tradition, a large number of articles emphasise the importance of laboratories for engineering students' learning, as teaching laboratories develop students' skills and knowledge as well as impart an understanding for the importance of empirical evidence in scientific work [2, 9, 12, 14, 30–33].

In 1985, Faucher [3] summarised the different opinions of all affected persons (University administrators, academic staff, students, industry) towards laboratory experiments in undergraduate courses for students of engineering (see also [34]). He concluded that laboratories (independent of the mode) cannot be replaced as they introduce the student to experimental methods. He furthermore underlined that a laboratory is well designed if it teaches experimental methods in a way that enables the students to plan the experiment and critically judge the obtained data.

Faucher also claimed that, for the professional life of graduated engineering students, it is important to be familiar with experimental work, as supervising or conducting tests independently may be part of their future responsibilities. In this context, being able to decide if the results obtained by others are trustworthy and applicable to the case at hand is mandatory. [3]

In 2000, Soysal named interactivity and hands-on laboratories as the most critical components to reach active experimentation in engineering courses. His survey concluded that the participating engineering and physics students considered laboratory experiments more effective teaching tools than lectures, readings, or homework exercises. [30]

In 2005, Feisel and Rosa set out to define the fundamental set of (possible) objectives for laboratories. Their study was motivated by the observation that the actual learning objectives of laboratories are often not stated, which causes problems: First, it is difficult to design an experiment without clear learning objectives (similar opinion to [10, 35]), and second, clearly stated objectives can be the basis for modifying and innovating an experiment. [12]

In 2002, Feisel and Peterson [31] set out to collect a clear general set of learning objectives for distance-delivered engineering laboratories, and found that this list was not available in literature, even independent of the method of delivery. They formulated a list with 13 objectives, all starting with [12, p. 127]: "By completing the laboratories in the engineering undergraduate curriculum, you will be able to..."

- Cognition
  - *Instrumentation.* Apply appropriate sensors, instrumentation, and/or software tools to make measurements of physical quantities.
  - *Models.* ...identify the strengths and limitations of theoretical models as predictors of real-world behaviour. This may include evaluating whether a theory adequately describes a physical event and establishing or validating a relationship between measured data and underlying physical principles.
  - *Experiment.* ...devise an experimental approach, specify appropriate equipment and procedures, implement these procedures, and interpret the resulting data to characterise an engineering material, component, or system.
  - *Data Analysis.* ...demonstrate the ability to collect, analyse, and interpret data, and to form and support conclusions. Make order of magnitude judgements and use measurement unit systems and conversions.
  - *Design.* ...design, build, or assemble a part, product, or system, including using specific methodologies, equipment, or materials; meeting client requirements; developing system specifications from requirements; and testing and debugging a prototype, system, or process using appropriate tools to satisfy requirements.

- Psychomotor Domain
  - *Psychomotor.* ...demonstrate competence in selection, modification, and operation of appropriate engineering tools and resources.
  - *Sensory Awareness.* ...use the human senses to gather information and to make sound engineering judgements in formulating conclusions about real-world problems.
- Cognitive/Affective
  - *Learn from Failure.* ...identify unsuccessful outcomes due to faulty equipment, parts, code, construction, process, or design, and then re-engineer effective solutions.
  - *Creativity.* ...demonstrate appropriate levels of independent thought, creativity, and capability in real-world problem solving. (compare [36])
  - *Safety.* ...identify health, safety, and environmental issues related to technological processes and activities, and deal with them responsibly.
  - *Communication.* ...communicate effectively about laboratory work with a specific audience, both orally and in writing, at levels ranging from executive summaries to comprehensive technical reports.
  - *Teamwork.* ...work effectively in teams, including structure individual and joint accountability; assign roles, responsibilities, and tasks; monitor progress; meet deadlines; and integrate individual contributions into a final deliverable.
  - *Ethics in the Laboratory.* ...behave with highest ethical standards, including reporting information objectively and interacting with integrity.

The objectives spread across all domains of knowledge and were meant to serve as a framework, making it possible for instructors to identify the specific learning outcomes for a laboratory. In a survey, engineering educators questioned reported that the list was judged to be complete and applicable, but that not all points might be essential. [12]

In 2017, Sullivan *et al.* [37] named

- designing experiments
- collecting and analysing data and
- using evidence to justify claims

as general learning objectives for laboratories, all of which covered by the above *cognition* category.

Feisel and Rosa [12] suggested fields which require further research:

- Methods of assessing laboratory effectiveness.

- The effectiveness of remote laboratories. Here they mention that comparative assessments require an agreement of learning objectives for both modes and correct assessing methods.
- The effectiveness of simulations vs. remote access of real equipment.
- Laboratory simulations that include "noise" (see appendix, page 273).
- Novel approaches to meet laboratory objectives.

Feisel and Rosa formulated two provocative questions (in terms of distance education), which were addressed by the present research in the local domain:

First, is it possible to make a simulation so realistic that the student cannot distinguish it from a set of control/measurement equipment controlling a real system? Feisel and Rosa propose a study on students working over the internet by instructing them to complete a remote experiment and then ask whether they thought they were dealing with real equipment or a simulation, in order to test whether they could distinguish between the two. For this type of experiment, Feisel and Rosa state the necessity of user interfaces which claim to control real equipment but in reality provide access to simulations [12, p. 128] (see subsection 4.2.9).

Second, they asked whether an educator should care about what a student perceives, as long as he/she meets the learning objectives of the laboratory. [12, p. 126]

2.2 The physical nature and location of laboratories

This section clarifies some definitions encountered in the literature on the research topic. Some terms/categories are used by different researchers to mean different things. Thus, these terms need to be clearly defined for the context of this research.

The term "online (laboratory)" is used in two different ways: It can mean "using computers" generally, or more specifically employing the internet for remotely conducting the experiment (off-classroom) in contrast to a locally conducted laboratory (which might also use computers). To avoid confusion *online* was not used in this research.

The term "virtual laboratory/manipulative" refers to

- remotely conducted physical laboratories, [38]
- remotely conducted simulated laboratories, [9]
- recorded physical laboratories, [38]
- local laboratories where the data are obtained through simulation, [22, 38–41]
- and laboratories basing on a fixed data set which is handed to the students. [38]

To avoid confusion *virtual* was not used in this research.

Dormido proposed in 2002 [42, p. 78] to characterise the different modalities of experimentation environments following two criteria: the way of access to the laboratory resources, and the physical nature of the laboratory [43].

2.2.1 Physical nature of the laboratory

Both physical/real experiment and virtual/simulated experiments are acknowledged in laboratory teaching to engage students in science. Both environments can offer different prompts for learning, frame student activities around scientific concepts, and expose them to scientific experimentation and the necessary skills to do so. [2]

While working in these laboratory environments, students need to consider the differences in the type of measurements that can be drawn from these resources: model results from simulations and real experimental results from physical experiments. [44]

2.2.2 Access to resources

The location/domain a laboratory experiment is conducted in can be categorised into two classes:

- Remote, off-classroom experiments
- Local, in-person experiments

2.2.2.1 Local experimentation (in-person)

Locally performed experiments can be split into two categories depending on the physical nature:

- *Local access-real resource.* Traditional form of laboratory learning, hands-on labs, the student works in front of a computer connected to the real experiment. The student operates tangible equipment [9]. Also called "proximal" [13].
- *Local access-simulated resource.* The experiment is fully based on software and the interface controls a virtual and physically non-existent (simulated) resource, which is part of the local computer together with the interface. [9]

2.2.2.2 Remotely operated experimentation (off-classroom)

With remote laboratories, the user and the experiment are physically separated. Usually, the user accesses the experiment through an internet or intranet connection. A special user interface is used to operate the equipment.

- *Remote access-real resource.* Real equipment which is accessed through the Internet. The user remotely operates and controls a real experiment through an interface. [9]
- *Remote access-simulated resource.* Like "Remote access-real resource", but substituting the physical system with a mathematical model. [9]

When using experiments based on simulation, the locality where simulations are calculated seems to be less important (*Remote access-simulated resource* versus *local access-simulated resource*). Nevertheless, depending on the learning objectives, differences can be found. Heradio *et al.* [9] identified influences in the perspective of control engineering:

- Decoupling of the model (which can run on the server) and the view (which runs on the client) supports the introduction of control-related experiences, like unknown time-varying delays.
- Possible online collaborative work.

2.3 Kinds of interaction with the experiment

Interaction with equipment can vary depending on the laboratory mode and can require different sets of skills. Especially with remote and/or simulated laboratories, interactions are determined by the offered software or simulations.

In 2008, Harward *et al.* [45] defined three types of online experiments:

- *Interactive* experiments are those in which students monitor and control some aspects of the experiment during its execution.
- With *batched* experiments, the sequence and the parameters of the experiment are specified by the student before the experiment begins and the results become available after the experiment ends. The pre-planned experimental sequence is executed without further influence.
- *Sensor* experiments are also planned ahead, but offer participants an opportunity to follow live data, although they are not allowed to control any aspect of the running experiment.

2.4 Mediation/User interfaces

2.4.1 Mediation of local “hands-on” experiments

“Computer integrated experimentation” refers to a setup in which various elements of the experiment (e.g. equipment control, data collection, and information processing/analysis) are organised on a computer interface [30]. Three of the four possible combinations of access and physical nature are per se technology-enabled. The only setup which can be run without a computer is the combination of local access and physical laboratory.

In 2006, Ma *et al.* [14, p. 10] noticed in their review the aspect that even “hands-on” laboratories become more and more mediated by computers. They claim that the interactive quality of laboratory participation may not differ much in these cases, depending if the student is working with a real apparatus or not. Ma *et al.* stated that most laboratories are a mixture of hands-on, computer-mediated, and simulated.

Tuttas *et al.* [46] identified the lack of haptic experience as obvious for technology-mediated laboratories in their experimental design.

Elaborating on that aspect, Lindsay *et al.* [47] stated that students in such a mediated hands-on laboratory are still able to inspect the hardware they are engaged with, even when the measurements and control signals are mediated.

2.4.2 User interfaces

When computers are involved in experimentation, the user interface can vary widely under different aspects, for example:

- usage of simple command line interfaces
- usage of the original user interface of hands-on devices (just extending the data connection to another place, or to simulations) [48]
- virtual 3D presentation of the experimental apparatus [15]
- usage of fully interactive 3D virtual reality laboratory environments to reproduce simulated hands-on experimentation [49–51]

Additionally, the application of augmented reality can support local hands-on learning by providing a better understanding when dealing with natural objects [52].

In virtual simulations, the level of reality and freedom influences the learning focus. Kangasniemi *et al.* found positive effects of a high degree of photographic realism with sufficient freedom of operation on student motivation. Contrary to Helmer [53, p. 8, p. 25], who lists distraction by too much freedom as a drawback, they are of the opinion that a simulation which is too straightforward to pass disables the need of reflecting one's actions. [50]

When planning to isolate the influence of the laboratory mode, varying user interfaces confound results.

Shimba *et al.* compared the outcomes of virtual and hands-on laboratories in teaching computer networking. They reported that students performed better after hands-on experiments. However, the user interfaces of the laboratory modes differed significantly and might have influenced the study outcomes. [54]

2.4.3 Enrichment of computer supported laboratories

A computer user interface can show animations, 2D and 3D renders which are not directly visible in the real world. Simulations can give enhanced insights into specific topics. The selection of the mode should depend on the desired learning outcomes (compare [45]).

Certain phenomena, such as the movement of individual electrons and ions in a battery while discharging, cannot be sensed directly or even indirectly via devices.

However, a computer program running a simulation can display these phenomena. Thus, if the desired learning outcome includes these aspects, the simulated mode provides clear advantages over the hands-on condition.

Virtual laboratories are also less influenced by time constraints, as participants can slow down, pause, or speed up a process. Experiments can be developed more freely, and essential points can be emphasised more accurately; whereas in a real experiment, time cannot be manipulated.

Corter *et al.* compared learning outcomes and student preferences for traditional hands-on labs, remotely operated labs, and simulations in a mandatory undergraduate engineering course. The authors conducted a very detailed literature research, showing that the reaching of learning outcomes seems to be similarly effective in all modes. For their study, students conducted two cantilever beam experiments (physics). The second experiment was more complex. In the remote mode, the student set experimental parameters on a website and received the result a short time after the experiment had been run, via either the website or email (asynchronous interaction). There were no pictures or videos. In the simulation mode, a 3D-view including a freely chosen point of view was provided. It was a real-time simulation enhanced with colour coded stress and strain values, change of material and geometry parameters of the beam. The students worked in teams of three or four. The content-related instructions were identical in all modes, but supplemented with some additional mode-specific instruction. The researchers compared the learning based on four different combinations (Experiment 1/more complex Experiment 2): hands-on/simulated, simulated/hands-on, hands-on/remote, and remote/hands-on. Learning outcomes of 306 students were measured with a multiple-choice test which was conducted directly after the laboratory session. These knowledge scores were equal or higher after doing remote or simulated versus hands-on laboratories. Students additionally completed a laboratory preferences questionnaire. Based on this data, Corter *et al.* concluded that, even though students mentioned advantages from technology-enabled laboratory formats, they still generally prefer traditional hands-on laboratories. The authors were additionally interested on the social and collaborative working patterns: Students had less face-to-face interaction when engaged in remote or simulated laboratories, compared to hands-on laboratories. In the semester before, information on individual students' visualisation skills had been collected. The students' college admission test results were also available for a correlations analysis. In both cases, the authors found no significant correlation with the learning outcome measures. [15]

Stern *et al.* enriched 7th grade lessons regarding kinetic molecular theory/particle movement/particulate nature of matter (contents difficult to demonstrate otherwise) with simulations. The objective was to evaluate the effect of that additional simulation lessons on the understanding. The researchers compared learning results of two groups. The first attended a lesson but no simulations, the second additionally received supplementary computer lessons using a software simulation. Knowledge gain was measured by a pre- and post-test. The influence of the teacher was eliminated by distributing equal numbers of classes to each participating teacher, so that each taught one class in each of the modes. Teacher dependent data is well recorded in the study. As expected, both groups improved their understanding of the theory. The results indicate that students using the simulations scored statistically significantly higher than those in the control group. Stern *et al.* found that students who had not used simulations had a more simple concept of particle movement, as they did not understand that particles are in constant motion, even at room temperature. [55]

An investigation conducted by Sarabando *et al.* targeted an improvement in student understanding of weight and mass. The researchers discovered that students involved in simulation-based learning performed better than their peers who participated in either a mix of simulation and hands-on or only hands-on learning. Unfortunately, these results cannot be generalised either. Firstly, different teachers taught classes utilising different laboratory modes. Secondly, students who learnt with simulations had more aspects of weight and mass to reflect upon than their peers who engaged in the hands-on mode (e.g. weight and mass on the moon). [56]

Hannel and Cuevas compared computer-based simulations (University of Colorado's PhET Simulations) with traditional hands-on methods in teaching 176 middle school students (6th grade) the concepts of A) density and B) the greenhouse effect. The teaching content was fitted to use each of the modes to the fullest: Using the simulation, students were additionally able to investigate materials of the same mass, but different volumes (A), or display visible and IR photons (B). Three written tests (a pre-, post-, and delayed post-test) were employed to measure knowledge gain and retention. With (A), students of both groups gained statistically significantly in knowledge (post vs. pre), but the gains did not differ statistically significantly between both groups/modes ($p = .064$). With (B), students of both groups again gained statistically significantly in knowledge (post vs. pre). The knowledge gain after hand-on teaching was statistically significantly stronger than after simulations ($p = .045$). Evaluating the delayed post-test (delayed post vs. pre), Hannel and Cuevas found that while both teaching methods helped to increase students' knowledge, no significant difference between both learning modes could be determined. [57]

The study of Corter *et al.* [15] had the goal of comparing the same experiment in different modes. However, features (e.g. colour coded stress values, material change) were available in simulation-mode but not in hands-on mode. It cannot be ruled out that the slight improvement in students' learning was caused by this factor or the additional mode-specific instructions. Many examples demonstrate that choosing the optimal mode correlates to the learning objectives: In the case of Stern's laboratory [55], particle movement is challenging to show in a hands-on experiment, as it happens at an atomic level. Also, Domínguez *et al.* [58] used simulations that allowed for well-teaching conditions that the available hands-on equipment could not reproduce. Hannel and Cuevas [57] visualised single photons to conduct the laboratory.

For these learning objectives, simulations can provide valuable insights that are impossible to make visible in hands-on experiments.

The choice of the optimal mode depends on the learning objectives. The comparison of the learning modes used through studies is reasonable and logical, while the drawback is the lack of possible generalisation of the results on the success of the learning modes in general.

2.5 Students' perception of laboratory modes

In 2005, Lindsay published results on perceived learning. Experiments were conducted in three perceived modes:

1. Hands-on experiments in the local domain.
2. Simulated experiments in the local domain.
3. Remote experiments (physically present experiments in the next room).

The user interface was the same for remote and simulated experiments, but not for the hands-on condition. In the experiment, students calibrated a piezoelectric accelerometer using a laser-Doppler and a spectrum analyser. It was attempted to make the remote and simulated laboratory conditions as similar as possible. For example, in one of the experiments, the simulation group heard recorded sound from a real experiment, while the remote group heard sound from the actual experiment. Both the remote and the simulation group conducted the experiment from the same room. Lindsay reviewed laboratory protocols and thereby determined the quality of the work and focus of the students in the respective conditions. He observed that the laboratory mode (hands-on, remote or simulated) had an influence on students' learning. The distance between students and equipment made the remote and the simulation groups more reflective, since the two groups did better at noticing and adapting to unexpected results than the hands-on group. The learning outcomes from the remote group and the simulation group, who had almost identical experiences, were not equal either: The simulation group paid more attention to the theory behind the experiment, and less attention to hardware and how it influenced the experiment. Lindsay noticed that the focus of students changed, because they knew that they were not using real hardware. He concluded that the students' attitude towards the laboratory was the deciding factor that affected the learning. The students' focus, and thereby the learning outcomes, were affected by whether students believed instruments were physically present in another room or merely simulated. [13]

Lima *et al.* investigated how deeply students realise the differences between two learning modes: remotely conducted laboratories and simulations. The researchers tried to answer the question if students really understood the difference between simulation and remote laboratories – and the different type of results obtained from them. Students from two universities and three study programs took part in the study: System Engineering (2nd semester) at Polytechnic of Porto and Computer Engineering

(4th semester) and Energy Engineering (4th semester)) at the University of Santa Catarina. 4th semester students task was to determine the electric current and the voltage drop across the inductor in a RL (Resistor-Inductor) circuit, using first remote experimentation, secondly a simulator and thirdly calculus. They were requested to compare the results of their experiment using the three methods in a report. 2nd semester students constructed a voltage divider in the remote mode and compared the results with calculus. The remote laboratory "Virtual Instrument Systems in Reality (VISIR)" (started in 2004 by the Blekinge Institute of Technology) was used for the remote experiment. In all cases, it was students' first contact with electric circuits and remote experimentation. Quantitative and qualitative data were collected and analysed: Students' final grades, their number of accesses to remote experimentation, a satisfaction questionnaire and an interview. The researchers found that the more often students accessed the remote version, the better their performance in the course test. Nine students were interviewed in detail to investigate the main research question. Here, students were especially asked for their reactions to deviating results in the modes and if measurements were repeated. The students' comments made clear, that most of the interviewed students had *not* understood the difference between simulation and remote experimentation (e.g. "in my understanding, we used the remote laboratory to perform a simulation"). Even students with good grades (or considered as having developed higher order thinking skills), did not always understand the difference or were able to explain it in clear words. This was true independently from teacher's mediation, students' maturity and academic experience. Lima *et al.* concludes that teachers must prepare some kind of simple (demonstrative) activity to make the difference obvious for the students. [59]

In summation, students' perception of the laboratory mode changes their attitudes and focus. Depending on the environment and modes used, it may be difficult (or not essential) for students to identify the actual data source.

2.6 Success of different laboratory learning modes

Several researchers have reviewed articles on laboratory learning and summarised outcomes on the comparison of laboratory modes.

Ma and Nickerson created in 2006 a comparative literature review on hands-on, simulated, and remote laboratories. It summarises different points based on twenty articles: First, many universities conduct laboratories as computer-based simulations or/and remotely if hands-on laboratories are not possible. Secondly, in many of the examined studies, the number of students was small or the researcher did not get significant results. Thirdly, the relative effectiveness of different kinds of laboratories is seldom explored. Lastly, based on their literature review on the comparison of different modes of laboratory work, Ma and Nickerson concluded that the reviewed studies do not allow for drawing a universally valid conclusion on the superiority of any laboratory mode [14].

Brinson reviewed literature on 56 empirical studies which directly compared student learning after traditional laboratories (hands-on) and non-traditional forms (virtual and remote). He found that most studies concluded that learning outcomes results either are improved (65%) or stay the same (24%) when hands-on laboratories are replaced by computer-based laboratories. The majority of studies (95%) examined how knowledge and understanding was transferred (The focus lay on dimension *K*, not *IPPAS* [60]: inquiry skills, practical skills, perception, analytical skills, and social and scientific communication). In many studies, the desired learning outcome to be researched was not specified clearly, and thus could not be categorised. The majority (71%) employed quizzes and tests for assessment. Other assessment methods such a scientific inquiry skills (7%), laboratory reports (9%) and practical exams (9%) were only used sparingly. The assessment method was not clearly described in many of the investigated studies. Brinson stated that the results of comparisons may change, as virtual laboratory technologies improve in terms of quality. These laboratories will become *more real* with time (3D, haptic feedback). Based on his research, Brinson made recommendations for the design of future studies. He stresses the importance of clear categorical boundaries, which – if mixed – are made transparent, as meaningful and unambiguous comparisons cannot be made otherwise. [10]

It needs to be noted that Brinson excluded all studies where the same students participated on both learning modes (e.g. crossover-like). Thus, the present study would have been not included in his review, since he doubted the employed sequence's influence on the measured efficacies (compare [39, 40, 58] vs. [41]).

In 2018, Tsihouridis *et al.* investigated virtual (simulations) and real (hands-on) laboratories in science education by reviewing research papers (articles, doctoral dissertations, and reviews) related to the comparison of real and virtual laboratories. The literature was analysed regarding result (most effective environment), if the trend varies over time, and the level of education in which the laboratories were employed. Before doing so, the authors filtered the literature according to three criteria: comparable "groups of participants", "methodological teaching approach", and "duration" of teaching interventions. 106 publications met the criteria and were further evaluated. It was observed that 52% of studies resulted in similar learning after virtual to real environments. Among the studies that found advantages for either mode, simulations (32%) outnumbered hands-on laboratories (16%). No statistically significant correlation was found between the year a study was conducted and the trend it reported. It was found that the effectiveness of the two different experimental approaches are influenced by the educational level of the students. 45% of studies regarding tertiary level education found better learning with simulations, compared to only 31% for secondary level and 14% for primary level students. According to the authors, a trend through all the relevant studies is that researchers often directly asked students about their preferred mode of laboratory learning, and students preference generally preferred a combination of the two laboratory modes over one alone. [22]

From the literature cited by Tsihouridis *et al.*, it is clear that criteria regarding the methodological approach were not interpreted as strictly as in the present research.

In sum, literature reviews found that learning outcomes after computer-based laboratories are at least equivalent if not superior to those after hands-on experiments. Research studies often identify no significant differences. The lower the educational level of participants, the more the hands-on condition is needed.

After conducting broad literature reviews, it has been recommended that studies on the relative effectiveness of learning modes with a relevant number of participants should be conducted. All categorical boundaries should be made transparent to allow for comparisons of study results: participant groups, teaching approach, experiment duration and assessment methods should be comparable and clearly stated in the publications.

2.7 Influences on laboratory learning

This section describes which intentional or unintentional factors influence the learning success. Scholars who compared the educational effect of different laboratory modes often omitted important influences on student learning.

These researchers all compared different forms of a laboratory, where important outside influences on student learning were not necessarily excluded. This approach is clearly justified to compare the best approach to teaching a certain topic (in certain environment), as idealised teaching in one mode might not be ideal in another. However, since the change of modes was usually associated with other factors which influenced students' learning, many studies benchmarked a mix of all of these aspects. Such factors include adapted learning objectives and tests, scope and type of supervision, distance learning vs. learning at university, customised experimental approach or different teaching materials.

Thus, the majority of these studies cannot pinpoint the effectiveness on learning of one influencing factor. As this also affects the judgement on the learning mode, this may explain the uncertainty in the findings of recent reviews examining laboratory learning [10, 14].

As this study's aim was to determine generalisable results purely focused on the learning mode, all other factors were held as constant as possible. In the following different examples illustrate the difference in approach.

Edward focused his study on problems with teaching aspects of mechanical engineering (behaviour of centrifugal pumps) to part-time undergraduate students working overseas. The various shift patterns of students working abroad led him to develop a substitute for laboratory experiments. He compared multimedia packages, including a computer-based simulation, with a conventional laboratory. The learning objective was to understand the operating characteristics of a centrifugal pump. Compared groups were built by two classes separated into two equal sections. Learning efficiency was measured by comparing pre- and post-test results, the report (required for summative assessment) and students' responses in questionnaires and interviews.

Edward determined that the students showed a preference for hands-on modes when asked directly. However, the achieved results were overall similar in both modes. Unfortunately, the details of group creation were not provided. The multimedia package consisted of texts, videos and computer applications, preventing a clear-cut comparison between simulations and hands-on mode. The number of participants was rather limited (56) and some of the tests were only carried out by half of the students. [16]

This study is an example of the difference discussed above: Two complete different solutions, each relevant for different situations, were compared. The found differences may be caused by specific details in one or the other solution (mode-specific instructions, different media) or by the different environment offshore (distance learning, cooperative learning, and supervision).

2.7.1 Intensity and type of guidance

While working on laboratories, students are guided in different ways. Lecturer, teaching assistance and instruction manual set goals for students to achieve and offer assistance towards reaching those goals. Guidance can greatly affect student learning and determine the success of the laboratory.

In 1984, Andrews [61] investigated the effects of two different laboratory styles in an introductory chemistry course: discovery (encourages the learner to generate conclusions inductively from ambiguous materials) vs. exposition-application (begin with an organised presentation of the material, and than ask the students to learn with the material). He employed two groups of students with different learning styles (“independent” and “dependent” Grasha-Reichmann Student Learning Styles Scales [62]), based on a learning style pre-test. It was found that the discovery laboratory style led to the best results for both groups of students. Considering the learning styles separately, it was found that the students categorised “dependent” performed better in the exposition-application style, while in the discovery laboratory the “independent” group was superior.

Domin [63] presented a review on laboratory instruction styles based on the chemistry laboratory. He distinguishes between three descriptors: outcome, approach, and procedure; and four instruction styles: expository, inquiry, discovery and problem-based (shown in Table 2.1). The named categories are usually valid for research in undergraduate settings [64].

Table 2.1: Descriptors of laboratory instruction styles [63, p. 543]

Style	Outcome	Approach	Procedure
Expository	Predetermined[a]	Deductive	Given
Inquiry	Undetermined	Inductive	Student generated
Discovery	Predetermined[b]	Inductive	Given
Problem-based	Predetermined[b]	Deductive	Student generated

Note. [a] = student and lecturer are aware of the expected outcome.
[b] only lecturer is aware of the expected outcome.

In *deductive* approaches, students apply a general principle to understand a specific effect, while in *inductive* approaches students derive a general principle by observing particular instances. [63]

Domin [63] states the expository type is the traditional and most widely used type. Here the participants follow the manual or the instructions of the teacher step by step. Usually, they experience the predetermined outcome, which is already known to them before starting the experimentation. Participants' results are then compared against this expected result. Disadvantages mentioned are the lack of emphasis on planning of the experiment or the interpretation of results.

Domin [65] also conducted a comparative analysis on the classification of expository laboratories in the literature with Blooms taxonomy and found that, in most cases, they address the three lower levels: knowledge, comprehension, and application.

The level of inquiry and degrees of freedom for the students differs. Buck *et al.* [64] characterised the intensity of inquiry in undergraduate laboratories in five basic rubrics based on six characteristics. Those characteristics represent areas in the experiments where students could either be expected to act independently or strictly follow the laboratory manual:

- *Problem/question* Does the student formulate the problem to investigate, or does the laboratory manual provide it?
- *Theory/background* Is the prior knowledge necessary for the investigation stated by the laboratory instructions, or are students expected to do research independently?
- *Procedures/design* Do the students execute predefined experimental procedures, or design them on their own?
- *Results analysis* Is the methodology of interpreting and analysing data given in the instructions?
- *Results communication* How free are the options on presenting data and experimental results?
- *Conclusions* Does the manual provide a list of observations and results that should have been obtained?

Based on these six characteristics, Buck *et al.* developed categories ranging from "Confirmation", where all six characteristics are given by the laboratory instructions to "Authentic Inquiry", where all six characteristics are defined by the participants. The researchers found that the first four characteristics are fixed in most laboratory texts. Only in 5 out of 386 investigated cases were *procedures* and the following characteristics free to be decided by the students.

Chamberlain *et al.* [66] investigated how the intensity of guidance affects participants' engagement with an interactive simulation in an undergraduate chemistry laboratory. Three written activities with different guidance level were compared: The light guidance level did not mention any specific controls of the simulation, and participants' goal was to investigate and observe anything in the simulation that they thought was relevant for the experiment. The moderate guidance level contained the goal of answering open response questions about the experiments, but provided minimal instructions on which controls of the simulation to manipulate. The heavy guidance level contained instructions which told participants how to interact with the program (in a specific sequence), with no directions to explore. Chamberlain *et al.* discovered that the intensity of guidance can strongly influence the participant research activity: Mouse click data and classroom field notes were used to assess participants' engagement while the experiment. The number of employed features of the simulation was significantly higher when light or moderate guidance was used with exploitative tasks. One week after the activity participants, were asked to sketch what they remembered from the simulation (arrangement), to measure how intensely the participants had engaged with the simulation. Participants which used light guidance drew more features than participants with moderate and heavy guidance during experimentation.

2.7.2 Teamwork and cooperative learning

The influence of the working team and group environment is an important factor on learning [67] [68]. Students have different abilities and background experiences. When working in teams, they do not have specific roles like in other formats (e.g. tutor/tutee), and they are expected to help each other to learn the materials. Webb [69] discussed small group interaction focused on a setting where all participants were expected to achieve the final test individually, instead of delivering an overall group result. She found that learning depends strongly on how the group is composed, and the abilities of other students working on the same topic. She also formulated a need to consider group development, as the role of individuals in working groups may change over time. She concluded that student learning cannot be understood isolated from group interaction, and that individual students may have different experiences in different groups.

Phillips *et al.* [70] analysed the influence of team work in an investigative biology laboratory. They analysed subjective reports based on group work, and found that team size affects the quality of team work. Based on their experience, better collaborative team work lead to better scientific results. They stated team sizes of three as ideal for laboratory work. If teams were bigger, workload was distributed unevenly, and teams of two students are always a risk for a singleton when someone leaves the course. They recommend allowing students to choose their own working team members to avoid blame for any internal problems falling on the teacher. They also identify a strong psychological component: if students believe not to be able to succeed without their peers, they seem to act more cooperative.

Stamovlasis *et al.* [71] validated the effectiveness of a group-learning approach in physics, focused on the oral interaction process. The researchers investigated group performance (number of correct answers being brought up by the group) and working group activity (measured as number of utterances). They found that a high-active group has high probability to change wrong statements to the right answer, and become a high-performance group. A low-active group has the highest probability to fail to change an initially wrong statement to the correct answer. Stamovlasis stated that *"If a student was lucky enough to be in a group with high performance, he/she would have had the chance to learn more."* An individual pre-test, a group test (solved by the group) after the discussion session, and a individual post-test, one week afterwards, were conducted. Just as with the groups, the results also favoured the more active students. Active group members improved during cooperative learning – while less active "spectators" did not benefit. The more cognitive statements a student made during group work, the better the results in the individual post-test.

Van der Laan Smith and Spindle [72] investigated if heterogeneous groups formed by the instructor lead to a more effective cooperative learning environment than groups self selected by the students. To evaluate the success of groups, individual academic performance and perceptions were considered. To judge the ability of students, prior marks were employed. Heterogeneous groups were formed based on that information. A trend was identified: students with higher ability had statistically better performance in homogeneous groups, while students with lower ability performed (less significantly) better in heterogeneous groups. Thus, Van der Laan Smith conclude that the best group composition may not be the same for all students and there is no overall best group composition.

Tien *et al.* [73] examined how instructors can reduce efforts when generating well-balanced groups based on their characteristics. Their literature research came to the conclusion that the composition of the group influences student learning and constructing inter-homogeneous and intra-heterogeneous group environments is desirable. They proposed a grouping method, based on the equivalence of an "individual" and a group, with students being "chromosomes" of the individual and usage of a survival of the fittest (most heterogeneous individual) algorithm. Their method promises to be very beneficial when trying to solve multiple characteristics grouping problems for cooperative learning environments.

To sum up, interaction in groups has an important influence on student learning. In this study, students cannot be understood independent of their peers. An important aspect in this thesis is the exclusion of effects other than the learning mode (like cooperative learning effects) as much as possible.

2.7.3 Teachers in laboratories

Due to the (usually) smaller class size of the laboratory in comparison to normal lectures, teachers in laboratories usually have more intense personal contact to students [74].

Thus, the teacher's mediation role in the classroom should be taken into account as important [44, 75] and special characteristics are needed to describe it further.

Gobaw [76] performed a regression analysis which showed that teachers' experience had significant influence on undergraduate biology students' laboratory skills, and Nikolnic *et al.* [75] emphasised the difficulty of evaluating the influence of teachers on laboratory learning as the prevalence of team based teaching with multiple teachers in the classroom does not allow for a separate analyses.

Previous studies in the literature often failed to exclude teachers' influence as a factor. This can influence the main comparison focused in a study and lead to skewed results. One example of a study *not* excluding teachers' influence is [56]. To avoid this interference, teachers and laboratory personnel were kept constant for all sessions and runs in the present study.

2.7.4 Sophistication of the compared laboratory courses

In some cases, the researchers tested add-ons to existing courses or compared a newly installed learning mode with an already established course in a different learning mode. From teachers'/researchers' perspective the research question is clear: Is there improvement in students' learning due to the add-on (or mode change) in this particular case? The specificity of the question makes generalisation particularly difficult.

McAteer *et al.* used simulation software in a practical laboratory on life sciences. Two simulation packages were integrated into a third year practical course on animal physiology. The first software package was employed to teach ion mobilities, diffusion, membrane potentials, and the Nernst and Goldman equations. The second package was employed to teach neurobiological processes: membrane and action potentials, threshold and refractory period, voltage and patch clamping, and the effect of neurotoxins. The number of participants was 66 over two years. The students were organised into groups of six or eight students. The laboratory classes were set up in "round robin circuits" of seven stations in a well supervised environment. Six of the stations were traditional hands-on experiments, one the simulation. Simulations were planned to be conducted in three sub groups with two students each, but most students decided to work in bigger groups on a single computer. Two of the "round robin circuits" were conducted in the course, in each of which one of the simulations were tested. The compulsory course ran for five weeks, providing fourteen 3-hour practical labs. To collect study data, the students were asked for before and after each laboratory if they were expecting to meet each learning objective (so-called "confidence log"). A delayed confidence log was done three months later. This delayed confidence log correlated with the overall exam results, which were available for the study. No significant correlations were found on special learning objectives and fitting exam questions. In addition, researches collected a post-course questionnaire and made informal recordings of feedback such as interviews with students and teaching staff. 90% of the students found the use of simulations in the laboratory useful, but many answers stressed that students still want to collect some hands-on experiments in the laboratory. [17]

Domínguez *et al.* changed one of five existing experiments (determination of the different parameters of an alkaline electrolyser) of a laboratory from a fully hands-on experiment to a sequence of a hands-on experiment followed by a virtual laboratory the next day. The grades of the students' reports were evaluated. The new (blended) mode showed a positive effect in contrast to the results of the unchanged experiments without the virtual laboratory. The results of the unchanged experiments decreased over the years, while the results of the changed experiment remained constant. The use of the simulations allowed for well teaching conditions that could not be reproduced by the available hands-on equipment. [58]

Engum *et al.* were faced with the disadvantages of hands-on experiments (such as a necessarily high teacher-student ratio) when training students in the first steps in placing intravenous catheters. Thus, they compared the effectiveness of an interactive virtual reality computer catheter simulator with the traditional laboratory experience (plastic arm training). Participants were nursing and medical students. After the procedure, the students – independent of the learning mode – showed a similar level of ability to correctly demonstrate the skill on a real human. The traditional hands-on learning method was preferred by students. Student satisfaction (Likert-scale questionnaire), and documentation of the procedure (simulated patient chart), in a comparison of results before and after, were better in the traditional laboratory group. [18]

Engum's research was designed for comparing two specific teaching solutions. Engum compared the proven solution (a scripted self-study module with a 10-minute videotape, instructor demonstration, and hands-on-experience using plastic arms) with a new virtual reality simulation. McAteer as well as Stern [55] added a new learning mode to a existing laboratory: simulations enriched learning, but it cannot be judged if hands-on experiments would lead to a better performance regarding the same leaning objectives. Another example is Edward [16] who could not to meet the requirements of offshore students with his well established course and compared the well established in-person laboratory with a newly created online training, conducted offshore.

These methods do not allow generalisable conclusions specific to comparisons of the learning mode, as other factors (e.g. the quality of the VR simulation) might have had a big influence. Additionally, it is problematic to compare experiments in well established laboratories with newly installed modes, since the existing mode has undergone an evolution to be optimised, while the newly installed one may be facing initial difficulties.

2.7.5 Students' time invested in the training

Mathiowetz *et al.* compared the results of an online anatomy software-based pathology investigation with a gross anatomy laboratory. Students were permitted to select their preferred learning mode. The hands-on laboratory group had a significantly higher grade-percentage, showed more self-perceived learning and higher satisfaction with the laboratory than the online software-based group. [19]

This outcome, however, cannot be generalised as the time students spent in the gross laboratory differed significantly from the time devoted to the software-based laboratory. Similarly, during Stern's study [55] one of the student groups invested considerably more time and effort into the topic, so better test results were to be expected.

2.8 Assessment of laboratories

Assessment can be defined as a judgement about achievements which requires evidence. Regarding laboratory learning and teaching, such assessment can regard students' knowledge, skills and abilities [35, p. 547].

The overall picture when investigating how laboratory learning is being measured shows that there is no general and widely used method to assess students' learning. It seems that the method is often decided freely by the researcher, teacher or university. Some methods recur and are often used, but differences were found between every study.

Self-assessments performed by the students seem to be a widespread method, but the questions differ from study to study. The specific questions given are rarely provided proper justification or a validation of the method. An absence of clearly stated learning objectives also often makes it difficult to judge an assessment.

Another widespread way to measure (gained) practical experience is applying an indirect method, evaluating the students on their written assignments [35, p. 547]. Again, the type of assessment varies. Examples are laboratory reports/notebooks (e.g. [13]), and tests with free response questions, multiple-choice questions, completion of diagrams, and Likert scale questions. One of the disadvantages of indirect assessment is the lack of control regarding the students' practical skills [35].

Hofstein and Lunetta [2] emphasised an important point when conducting assessments in laboratories: the need for consistency between desired learning outcomes and chosen assessment methods.

Regardless of the differences from study to study, inspiration could be drawn from the papers to develop a way to assess practical experience.

Abraham, Reiss and Sharpe [77] found that literature on the assessment of practical work is limited, and a definition of practical skills and how they are best assessed is not given. Opinions about which skills are important differ between employers and university departments.

Sanders *et al.* [78] proposed to give the list of learning objectives (combined with performance criteria which students should meet in the assessment) to the participants to guide them in building the right skill sets.

Campbell *et al.* [79] criticised that there is also an absence of detailed documentation of laboratory work done by students.

Studies describing useful learning outcomes in terms of practical skills can be found in the literature. An example is Aziz and Ferris [60], which proposed to expand Bloom's taxonomy through a hierarchy of seven learning levels for the Psychomotor Domain:

- Recognition of tools and materials
- Handling of tools and materials
- Basic operation of tools
- Competent operation of tools
- Expert operation of tools
- Planning of work operations
- Evaluation of outputs and planning means for improvement

Unfortunately, in their publication, no assessment methods were suggested.

Test results are also influenced by the time period between experimentation and testing, an effect which seems to differ depending on the learning mode. Chini *et al.* performed a study on undergraduates' learning from physical (hands-on) and virtual (computer simulated) experiments. The researchers had sequentially combined laboratories on the topic "pulleys" in two following weeks. Depending on the group, students used one of the modes in the first week. In the second week, the students performed similar experiments with the other mode. They report three major outcomes: First, students' performance in a immediate post-test was related to the concept being learned: While no significant difference between both modes was found for the concepts of "force" and "mechanical advantage (how many times a machine multiplies the applied force)", *virtual* experiments supported students' learning about "work and energy" significantly better. Second, the performance depended on whether the post-test was immediate or delayed: When the test was delayed by one week, the advantage of simulations was absent, meaning that advantages of learning modes may disappear over time. Third, at the final test, both sequences of experimentation (hands-on–simulation, simulation–hands-on) resulted in equal student understanding for all three concepts. [39]

Even when no general method of assessing practical laboratory experiences was found, studies about assessing students practical work do exist:

2.8.1 Direct and indirect assessment

Abraham, Reiss and Sharpe reviewed how practical work is currently being graded in England. They categorise two different methods of assessing practical skills: direct and indirect. In a direct method, the student is directly being assessed while performing laboratory work. In an indirect assessment, a report or a questionnaire based on practical work is assessed. The authors argue that by using an indirect assessment, the students are being assessed on what they know about how to perform practical work instead of their ability to actually perform it. They recommend using a direct method if a student's ability to perform a specific task is measured, and an indirect method if the focus is on the student's understanding of a skill. [77]

Baxter, Pine and Shavelson developed and evaluated different methods to assess performance in science. They judged an assessment of students' practical skills by directly observing them as costly and time-consuming. One of the indirect alternatives is making the students fill a notebook while performing the experiment. The notebook is then assessed by the teacher afterwards, without consuming too much time. They name multiple-choice and free response paper exercises are alternatives as well. The authors noticed that some students would do well using one alternative, but do less well using another. [80]

Radin-Salim, Puteh and Daud evaluated students' practical skills in basic electronics, based on nine experiments designed to fit the learning levels suggested by Aziz and Ferris [60]. For assessment, the students performed perform eight different tasks, the first three were of which writing exercises (name and identify types of components, sketch components and identify symbols, and explain functions of components). The other five were practical tasks like construct a circuit, connect instruments, set an instrument to a required value, use an instrument to measure a required value etc. As a result, the authors recommended that assessments solely relying on laboratory reports should be replaced. [35]

2.8.2 Self-assessment

Letting students perform a self-assessment about their abilities and experience in a laboratory has also been used in some studies as a method for assessing practical work/experience:

An example is found in a publication of the assessment office of the University of Colorado Denver [81], which made an outcome assessment report about a electrical engineering undergraduate program. The report states that a self-assessment is used to measure student learning, by making the students rate themselves according to the achievement of learning outcomes. The university performed the self-assessment over a few semesters, and found that student state both their shortages and capabilities. The report contains the learning outcomes for the electrical engineering courses offered by the university, some of which concern the students' ability regarding the practical tasks.

A research group from "The Council on Undergraduate Research (CUR)" [82] developed an online survey: the Undergraduate Research Student Self-Assessment (URSSA). The survey was used to assess what students learned from doing undergraduate research. Several rounds of interviews with students and administrators with experience in undergraduate research resulted in a list of gains achieved by doing undergraduate research. The presented self-assessment tool has been designed as a way to measure those gains.

Anastasio *et al.* [83] altered this URSSA-tool to make it more appropriate for assessing student learning in chemical engineering laboratory courses. The authors do present some modified sample questions, but do not argue exactly why the modifications have been made.

Campbell *et al.* [79] evaluated the validity of self-assessments. The evaluation was performed comparing two different data sources: perceived practical experience, reported in self-assessments, and remembered practical experience, tested by letting the students describe names and purpose of laboratory instruments. First, the perceived assessment collected student answers on how often they had done laboratory experiments. They were also asked if their teacher often demonstrated experiments, and if they had read about instruments. Then, the remembered practical experience was assessed by introducing the student to pictures of chemistry laboratory equipment. They were asked to name each piece of equipment and describe its purpose. In their results, Campbell *et al.* found no significant difference between the perceived and remembered practical experience.

McAteer *et al.* used so-called "confidence logs" to collect study data. The students were asked, before and after each laboratory, if they expected to meet specific learning objectives. [17]

2.8.3 Practical intelligence

Razali and Trevelyan (2008) defined practical intelligence as implicit or tacit knowledge. It can be described as unintentional learning, which comes from personal experience [84].

They argue that students are, in most cases, evaluated on their explicit knowledge, typically by assessing specified learning outcomes, neglecting students acquisition of practical intelligence during laboratory experiments. They investigated a new kind to measure practical intelligence gained by laboratory classes by constructing a problem to solve (wire stripping). The students are presented with a number of response items; pliers, teeth, scissors and professional wire stripping tools. The students rated each item on a Likert scale. Engineers and technicians were asked to do the same. Students were judged by their deviation from the average expert response. The authors claimed that this evaluated students' practical intelligence, since some of the items may be less suitable in theory but, from experience, may be known to work quite well. The test was performed for students actually doing a laboratory experiment and students who did not. The results showed that the practical intelligence gained from the laboratory experience can be measured by this method of setting of practically relevant problems. [85]

Chapter 3

Review on the educational pathways of the study participants

The study conducted for this thesis compared data of the participants based on whether they had completed a vocational education before enrolment at university. This chapter provides information on the different educational backgrounds of these two types of students to readers not familiar with the German education system.

In particular, data was separately collected for two distinct groups of participants:

- Students who had not completed any VET education “non-VET”:

  These students followed two possible paths to earn their university entrance qualification:

  – completion of upper secondary schooling,
  – completion of intermediate secondary schooling, followed by specialised upper secondary schooling.

- Students who had completed a VET education “VET”:

  VET students followed three possible paths:

  – university entrance qualification (e.g. upper secondary school or specialised upper secondary school) was earned prior to VET education
  – extra schooling (e.g. senior vocational schools) after VET graduation to earn university entrance qualification
  – further training qualifications in the VET-system to earn university entrance qualification

The German/Bavarian education system

Germany is a federal republic. Education is based on the law of the individual “Bundesländer”/states. Nevertheless, the general ideas discussed in this chapter hold across all German federal states. A more detailed depiction of the German educational system can be found in [86] and [87, p. XIV].

Figure 3.1 presents the general pathways into the German vocational training system (VET) and university education system. The figure separates the compared groups in this study: Participants who attend university after completing vocational training, and participants who enrolled without such training.

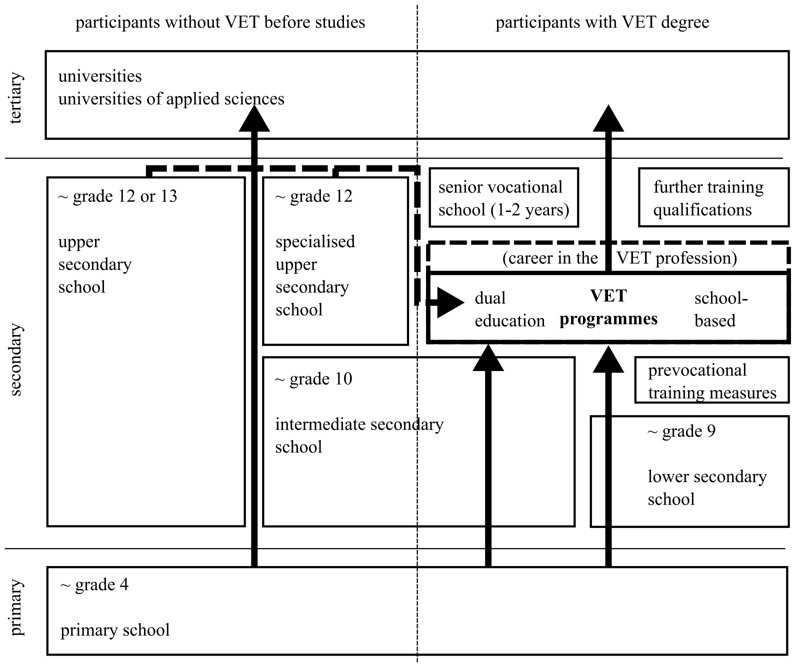

Figure 3.1: Pathways toward tertiary education (simplified). The top left side represents non-VET study participants who enrolled immediately after having finished secondary school, the top right side shows the educational career of students who completed a VET before enrolment.

3.1 Primary and secondary schooling

In Europe, especially in the German-speaking countries' stratified school systems, children are divided in early schooling into separate tracks, often taught in different schools [88, p. 31, p. 34]. In the federal state of Bavaria (where UAS Ingolstadt is located), students are separated into different school types after fourth grade, where they can earn three different degrees:

- First, the lower secondary school degree, which is awarded after ninth grade.
- Secondly, the intermediate secondary school degree, which is awarded after tenth grade. Different types of intermediate schools with accordingly varying curricula exist. As shown in Figure 3.1, pupils are allowed to attend specialised upper secondary schools based on the intermediate degree in order to gain access to tertiary education.
- Thirdly, the upper secondary school degree (Germany's general university entrance diploma), which can be obtained after twelfth or thirteenth grade. It is required to access tertiary education at universities.

3.2 Education and training in Germany

"Education" and "training" are often distinct in society's expectations. "Education" usually emphasises the understanding of general conceptual issues and "training" refers to the development of expertise in the use of specific tools [89].

After finishing school, youth continues education. Bavarian education law mandates a *compulsory education*, which lasts twelve years [90]. It is divided into compulsory full-time schooling (9 years) and compulsory vocational schooling (3 years). A higher secondary school can replace vocational schooling. Depending on their school degree, young people have two ways to continue after secondary school: vocational education and training (emphasis on "training"), or enrolling at a university (emphasis on "education"). Both options are described in detail in the following paragraphs.

3.2.1 The German Vocational Education and Training (VET) system

A VET program can be completed after getting any of the school degrees [91, 92]. For some occupations, upper or intermediate school degrees are mandatory, others do not require any school degree. As shown in Figure 3.1, there are two pathways to a fully qualifying occupation-specific VET certificate that is acknowledged in all of Germany:

- The first is taking part in the dual system of vocational training [92]. Herein, a company based training is combined ("dual") with 8 to 12 hours a week of state school-based education. [93, p. 579]. The school-based part provides general upper secondary education (German language, mathematics, etc.) and theoretical knowledge on the particular profession [94]. These schools also teach basic knowledge and special information not every company is able to instruct and demonstrate [92]. In the dual system, company training including the apprentices' wages are financed by the employers. These wages increase over the years of training and differ between professions, with an average amount of 908 €/month in 2018 [95, p. 264]. The vocational schools are financed by the state. In 2017, 29.2% of VET apprentices held an upper secondary school degree, 42.3% held an intermediate secondary school degree, and the remainder either held a lower secondary school degree or no school leaving certificate at all [95, p. 139]. VET apprentices with upper secondary school degrees would have had the option to enrol at universities instead of applying for a VET.

- The second way to get a VET certificate is to attend a school-based vocational education program. Here, the students do not have a contract with a company and therefore, do not get paid during their training [94, p. 6]. This category of training regards other types of occupations – mainly "white-collar". It cannot be seen as an alternative for occupations that require training of the dual kind [94, p. 3]. Only 28% of VET participants are in this category [28, p. 2].

> Because of this low percentage and the non-technical focus, this type of VET-participants is very rare in engineering studies compared to graduates of the dual branch.

The more than 300 occupation-specific training curricula (3 or 3.5-year programs) cover a broad array of competencies in order to avoid skills that are too specific to the company they are trained by. Chambers/Trade Unions (industry organisations overseeing their trade-specific training programmes) observe and organise the knowledge that trainees have to acquire and perform quality control. In addition, these chambers define and manage the duration of training as well as the examinations [96, p. 10]. After the VET program is completed successfully, former trainees obtain their occupation-specific certificate [94, p. 3]. The system delivers highly standardised qualifications/degrees in the German workforce [88, p. 35]. Usually, people only find employment in the occupation they have been trained for or in very similar positions [94, p. 7].

When it comes to the availability of VET positions, two opposing trends can be witnessed. On the one hand, the number of offered training positions in the dual system depends on the needs of the companies. 54% of German companies are allowed to offer VET-training, but only 29% do so [97, p. 2f]. Additionally, most people apply for a dual VET apprenticeship at 16 or 17 years of age. Since many of them live with their parents, options are locally limited [94]. Thus, esteemed professions are competed for heavily [98]. On the other hand, craft businesses (such as carpenters or plumbers) face increasing difficulties to fill their open VET positions. While only 15% of companies reported not having enough trainees in 2007, the share of companies with open positions increased to 34% in 2017 [99, p. 2]. The main reasons given were that no (26%) or no fitting (70%) candidate had applied for the position [99, p. 6].

Mostly lower secondary school graduates are trained in the lower skilled and less well-paid service professions, while in well-paid professions, most trainees have an intermediate school degree [94]. Career chances differ notably between different VET professions [98, 100]. Candidates with an upper secondary school degree attend higher-skilled and more esteemed VET programs [94].

Germany has a strict qualification structure, separating employees without degree from those with VET degrees and again from those with a completed tertiary education [94]. The education system is also sharply divided between VET and higher education [101, p. 7]. Skills gained in the VET sector are rarely recognised by the tertiary sector. A VET-upper-secondary-education-degree is not regarded as equivalent to a general upper secondary education degree, as learning objectives differ between vocational and general competencies and the instruction methods differ between practice and theory. VET participants in the technical sector work at least three days per week. Training is usually hands-on, consists of learning by doing, and therefore practical knowledge is gained, whereas education in secondary schools focuses on theoretical knowledge.

Each year, more than 520,000 new contracts in the Vocational Education and Training (VET) system are signed between trainees and companies in Germany [95, p. 15].

3.2.2 Higher education institutions in Germany

Germany's institutions of higher education are categorised into universities, colleges of art and music, and Universities of Applied Sciences (UAS).

The majority of higher education institutions are financed and regulated by the state. There are also privately operated institutions (by for-profit and non-profit establishments and churches) that are officially recognised by the state (these are mainly UAS). [102]

3.2.2.1 Universities (and institutions with the same status)

Universities are research-oriented and offer a wide range of subjects. Some of them focus on particular fields (for example medicine). Only these approximately 120 traditional/full universities have the right to grant doctorate degrees [102].

Nevertheless, the Bologna Process (with different accreditation procedures) results in more similarity between UAS and universities [103].

Tenured professors are usually recruited based on their scientific career in academia and post-doctoral lecture qualification.

Approximately 60 German "colleges of art" and "colleges of music" have equivalent status to universities. Some traditional universities also have faculties offering these subjects. [102]

3.2.2.2 Universities of Applied Sciences (UAS)

The approximately 210 Universities of Applied Sciences (German: HAW = "Hochschule für angewandte Wissenschaften") usually offer study programs covering a limited range of subjects (mainly engineering, business, and social sciences) [102].

The main difference to traditional/full German universities is the more applied / practical orientation [5].

At UAS, professors are recruited based their experience and career in the industry.

Currently, more than thirty study programs concerned with electric automotive engineering are offered at UAS in Germany [104]. The bachelor program "Electrical Engineering and Electric Mobility" at UAS Ingolstadt contributed most participants of the present study.

3.3 VET-graduates in tertiary education

Completion of a VET plus at least one year of additional school (in Bavaria: senior vocational school) entitles graduates to enrol in a University of Applied Sciences in accordance with the field of their educational training [91].

Approximately 13% of successful VET graduates later take part in tertiary education at universities or UAS after VET graduation [28]. 35% (SD = 6%, N = 7 bachelor study programs) of the bachelor students in the Faculty of Electrical Engineering and Computer Science at the Ingolstadt University of Applied Sciences (Technische Hochschule Ingolstadt, THI) completed a VET before enrolment.

It seems plausible that only VET candidates who perform well regarding the VET schooling component tend to later go on to university, as they make a conscious decision to rejoin a formal school environment. More than 70% of these former VET trainees choose Universities of Applied Sciences over traditional universities [28]. In 2012, 40% of all newly enrolled UAS students held a VET degree, while the ratio was 11% for students who enrolled in universities [105, p. 33].

In most of the federal states of Germany (including Bavaria), senior vocational schools were implemented to provide VET participants with an upper secondary education degree (alternatively, five years of work experience in an occupation) which allows enrolling in tertiary education institutions. In senior vocational schools, students are assigned to a schooling program based on their completed VET profession. The duration varies between one to three years, depending on the completion of an intermediate secondary school degree before VET and the type of educational institution the students want to gain access to [106].

The share of UAS-enrolled students who are VET-graduates and did not have the qualification to enrol before starting their VET education is shrinking. These students had to, for example, attend senior vocational schools to be able to enrol. In contrast, the share of UAS-enrolled students who already held a school degree allowing them to enrol, but chose to do a VET first and enrol later has increased from approximately 33% in 1990 to 43% in 2012 (shown as stroked way from left to right in Figure 3.1). [105, p. 34].

There are further training qualifications in the VET-system (e.g. master craftsman, or technician courses) which include the right to enrol at universities, but the share of students based on that education is minor. Less than 3% of graduates of these courses later enrol at universities [105, p. 38].

3.4 VET: Social background and mobility

This study does not investigate the topic of social mobility per se. Nevertheless, statistics regarding social background and mobility can give supportive information about the most likely participants' environment while growing up, which differed between compared groups.

The German school system and the German labour market are strongly regulated through official certificates, which limits the ability to freely choose jobs [94]. On the one hand, the gap of reputations between jobs requiring a university degree – and those who do not – is less significant as in many other countries, as the German VET-system partially fills this gap with its state-wide recognised degrees. On the other hand, the stratification of the German school system (which incorporates the VET system) tends to conform to parents' expectations. Thus, particularly underprivileged children are often lead to VET apprenticeships and thus prevented from accessing or seeking out higher tertiary education [88, 107]. Even with the German state financing the majority of universities in Germany and costs for studies being comparatively low (no tuition fees for Germans, EU citizens, and foreigners), it is overall less costly and risky to go for a VET compared to a tertiary education [94]. The existence of German VET education may cultivate social difference across generations [88, p. 46].

There are two major cost advantages to a VET program over a university-level education. Firstly, VET trainees are paid wages which allow young people to finance their life independent of their parents. Secondly, participating children may leave school after grade 9 or 10 instead of grade 12 or 13. Both cost aspects may lead underprivileged families with low income to push their children into VET instead of enabling tertiary education. As shown below, in Germany, the educational background of parents strongly correlates with the educational career of their children.

Higher secondary school education

When looking at the distribution of pupils among various types of school, strong connections to the respective social background become clear. The higher the parental education, the more likely students attend upper secondary school. In 2016, 58% of German children whose parents held a university degree attended upper secondary schools, whereas only 24% of children of parents who held a VET as their highest degree did the same [87, p. 54].

Access to tertiary education

Even when children from non-academic families attended upper secondary school (graduation from which directly qualifies to enrol at universities), they more frequently proceeded with an education at vocational schools than children of academics [108].

The participation rate for higher education depends strongly on the school degree of the parents: Only 12% of children whose parents do *not* have either a vocational qualification or higher degree enter higher education. If at least one parent finished a VET education, 24% of children enter higher education. For children where at least one parent held a school degree which included a university entrance qualification (upper secondary school or equivalent), the university participation rate is 48%. [108, p. 5]

Similarly, children of parents who are academics are approximately three times more likely to attend university than their peers: 79% of children whose parents hold a university degree attend tertiary education, compared to only 27% of children from non-academic families [108, p. 5].

3.5 VET: Wages and career perspectives

Career perspectives were analysed to understand the consequences of choosing a VET or tertiary education at university: Ignoring regional discrepancies in Germany, both educational ways assure a safe future. In Germany, the average unemployment rates of qualified workers has been decreasing for several years [109, p. 5]. Only 3.3% of VET certificate holders and 2.2% of university degree holders are searching for work, while 18.3% of persons without either degree were unemployed [109, p. 28]. 74% of VET apprentices are hired directly after completing their training by the companies that trained them [97, p. 3].

The average German employee with VET degree earns 36% more than employees without such a certificate. However, this is only 59% of the average wage of a university graduate [110, p. 22]. Looking at the trend over a career, the wages directly after finishing the degrees are similar: At age 25, holders of bachelor and VET degrees earn the same, on average 2,750 € [111, p. 4]. However, for advanced careers, the difference increases significantly: In 2014, university degree holders above 60 years of age earned, on average, +66% more than their colleagues below 30. For VET participants the wage increase over age is quite lower, only +14% more[110, p. 23]. At age 50, the difference averages about 30,000 € per year [112].

These numbers show that, while VET candidates do not face disadvantages when entering the work force, their career options above a certain level tend to be very limited.

Chapter 4

Research approach and methods

This chapter details the general research methodology of the study presented in this thesis.

As the overall study consists of nine study runs (R1-R9), different approaches were combined to answer the research questions. This section sums up the employed methods/data sources and their application to the respective study runs.

This chapter is structured as follows: After an overview about the employed data sources, the methodology for collecting and evaluating the data for each data source is provided. Then follows a description of each individual study run (section 4.8). The chapter ends with a short discussion of the research ethics.

The analysis of the collected data is presented in chapter 5.

4.1 Data sources

The following data sources were available during the two research phases:

- (DS-A) Section 4.2 discusses the main study which uses an approach similar to a medical cross-over study to investigate the influence of the compared laboratory modes on student learning. The first research phase compared hands-on and simulated experiments, the second compared hands-on experiments to hidden simulations. Tests on the knowledge gained during the previous laboratory experiment or (e.g. in case of summer schools) final test regarding all laboratory experiments (see section 4.2) were employed to compare student learning in both modes.

- (DS-B) Section 4.3 presents the introductory questionnaire, which included questions regarding prior (specifically technical) practical experience (APE), and prior apprenticeship training (VET). In German runs, the enrolled students were split into two groups based on their prior practical experience to ensure a similar group environment (see section 4.3).

- (DS-C) Section 4.4 details the collection of additional information about the personality and background of the participants based on the well known personality types defined by John L. Holland (RIASEC). As the employed question sets are commonly known in German-speaking countries, the results serve to enable other researchers/teachers to compare their participants/students with those from this study.

- Subjective opinions regarding the laboratory modes were collected from both participants and non-participants via two distinct methods:
  - (DS-D) Participants of the main study were asked to provide feedback on a specific experiment after its completion. This method allowed for indirect gathering of information about the different perceptions of the compared laboratory modes (section 4.5). The students were asked for their thoughts on *each* laboratory session in a short survey. This allowed analysis regarding subjective opinions after experimenting (see section 4.5). In order to allow for unbiased results, students were not asked to directly evaluate the respective learning modes.
  - (DS-E) Persons who either had not yet started the laboratory or were not participating at all were asked to fill out a general questionnaire, distributed amongst different universities in different countries (section 4.6). This questionnaire, created using using Qualtrics, asked for subjective opinions regarding the compared learning modes of *hands-on* and *simulation*.
- (DS-F) Section 4.7 presents the methodology and findings of an analysis of the THI university database to extract objective background information about students with and without VET degrees.

Table 4.1: Data sources: participants and frequencies

Data source		Data collected from	Frequency
DS-A	Test results	Laboratory participants	each test
DS-B	Amount of Practical Experience (APE) and VET	Laboratory participants	once per run
DS-C	RIASEC/Personality	German lab. participants	once per run
DS-D	Feedback regarding previous experiment	German lab. participants	each test
DS-E	Questionnaire on opinions on learning modes	Persons from different universities	once
DS-F	University database	all actual and former THI STEM students	once

4.2 Comparing student laboratory learning using different learning modes (DS-A)

This section discusses the general research methodology to collect data for DS-A, including the conduction of the laboratories in both research phases and usage of the knowledge tests for comparing the employed learning modes. The research findings can be found in section 5.2. The outcomes are linked and discussed in chapter 6.

4.2.1 Counterbalanced methodology / crossover

In order to minimise the influence of various study factors on knowledge acquisition as a result of conducting hands-on and simulated laboratories, a methodology *similar* to a crossover trial in medicine was applied. In crossover trials, subjects receive a sequence of different treatments. For example, in the first phase of a study, participants of Group 1 receive treatment A, and participants of Group 2 receive treatment B. In the second phase, treatment B is administered to the subjects of Group 1, while treatment A is administered to the subjects of Group 2. [113]

Grouping and working teams
In order to implement a similar methodology in this study, participants were divided into two comparable (inter-homogeneous and intra-heterogeneous) groups. Both groups had the same learning objectives, which were clustered around either two or four content areas depending on the study run. Each group was divided into working teams of three to five (very seldom) students to conduct laboratory experiments together.

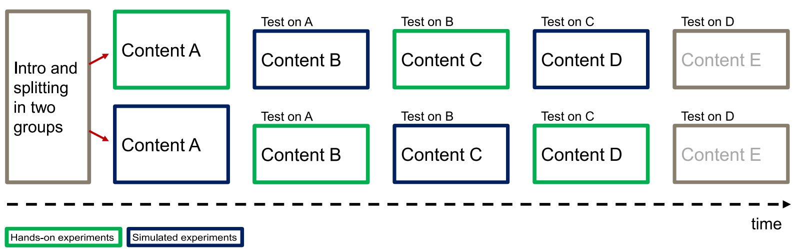

Figure 4.1: Counterbalanced within-subject methodology, similar to a crossover trial

Flip-flop / ABAB pattern
Both groups worked on content areas in the same order but switched laboratory modes between sessions (see Figure 4.1).

In odd weeks of the semester, working teams of the first group conducted simulated experiments, while the second group conducted hands-on experiments. During even weeks of the semester, the groups switched their respective modes: the first group ran hands-on experiments, the second group ran simulated experiments.

All participants did the full crossover every session (ABAB, instead, for example, AABB), to minimise "Carry-Over-Effects" [114]. Given differences between the learning modes (and following content areas based partially on knowledge accumulated before), the study targeted an overall equal knowledge protrusion of all students over the semester. It made individual knowledge tests for single content areas better comparable and reduced ethics concerns. As the potential number of participants was unknown (individual students were free to choose to participate, and the number of possible study runs was unknown initially), the study was not enhanced with other additional designs.

In contrast to medical crossover studies [113], participants in this counterbalanced within-subject study were never exposed to both learning modes with the same learning objectives. In the first research phase, at the beginning of the session each week, students were notified of the mode they were to conduct the laboratory experiments in.

Assessment of student learning

A 10-minute test was held to assess the influence of the learning mode on knowledge acquisition during a laboratory. In the German runs, these tests took place at the beginning of each subsequent laboratory session, while participants of the remaining runs were tested two weeks after the laboratory session (in both cases a posttest-only design). The study was fully anonymous, and participants did not receive incentives for taking part.

Study runs

The study was conducted in two consecutive phases over four years (2016 to 2019) in nine study runs (R1 – R9) (see Table 4.2). Study runs of both phases were carried out in the *local access* domain. An overview of demographics can be found in Table 4.3.

- In the first phase, the objective was to compare student learning through hands-on laboratories with results gained from simulated laboratories (R1 – R5).
- The second phase was conceptualised to verify the quality of simulations used in phase one and at the same time give insight into possible subjective influences of the laboratory mode itself by comparing the results of hands-on laboratories with laboratories in which hidden simulations were conducted (R6 – R9).

In each phase, two runs were conducted in German with 2nd-year electric mobility bachelor students, and the other runs were conducted in English with students of mixed backgrounds (see section 4.8 for a description of all study runs).

In total, 268 students participated in this part of the study. All students were asked to join this fully anonymous study, and all agreed to participate. Participants did not receive any incentives for taking part in the study, monetary or otherwise.

Local domain comparison

Since the study at hand focuses on a strict methodological approach, a *remote* laboratory condition was out of the scope of this study and has not been investigated, even though many students favour online education and recent literature reports equal or better learning with remote laboratories [9–11, 24, 26]. The study was designed to compare in-person laboratory teaching with/without proper laboratory equipment solely in the *local* domain (subsection 2.2.1).

Kind of interaction with equipment

Interaction with equipment can vary depending on the laboratory mode and can require different sets of skills. In this study, three out of four content areas of laboratory exercises were designed as *sensor* experiments (see section 2.3) of different duration. The procedure was planned and configured ahead (e.g. a current profile) by the students. After starting the procedure, participants were not allowed to control any aspect of the *running* experiment (besides stopping the procedure). The participants got the opportunity to follow live data.

The experiments were planned and executed by the students. In contrast, a *demo* would be a video (or simulation) of an experiment *someone else* planned.

The duration of individual programmed laboratory experiment procedures in terms of this paragraph differed and was given by the technical battery experiment. Determining one DC internal resistance of a battery (see page 307), for example, only took minutes to complete, while some exercises that require battery discharge (see pages 301, and 322) took one hour to finish. After running one of these pre-programmed procedures, students had to configure the following procedure.

Duration of these procedures and interaction with hands-on and simulated experiments were practically the same.

Avoiding other influences

To establish reliable and universally applicable insights on the influence of a particular laboratory mode on learning, this study aimed to exclude other influencing factors on the student learning than the learning modes themselves (such as accompanying lectures, experimental instructions, teachers, learning objectives, tests, working teams, and many more; please see chapter B). In both modes, the learning objectives and experimental approach of the exercises were practically identical. Student interactions with equipment/simulations remained the same.

Besides the learning modes, all other factors were held constant, to minimise differences between sessions. Thus, even when there was a possibility to shape the curriculum according to the learning modes (for instance, a time-lapse with simulations or adding different aspects about connecting the cells during hands-on experiments), the content and experimental procedures were *not* altered. Favouring comparability over pure learning success meant choosing procedures that can be taught similarly in both learning modes over those that fit only one learning mode.

As a result, this study does not determine which of the modes is a better fit for teaching batteries specifically, but the influence of the laboratory modes on student learning in general.

Case study
Even though the methodology was designed to produce generalisable results, a case study had to be performed. Case studies are specific instances which have to be frequently used in order to illustrate/investigate a more general principle [115]. The selected parameters are easy to describe:

1. The study is based on teaching the topic of batteries, where most learning objectives (such as current and voltage) are not tangible or audible. In both modes, students need to believe data displayed on computer screens and displays of devices. The intensive use of device mediated experiments was identified as a general trend for hands-on laboratories [14].

2. The period between laboratory session and test was chosen to be one or two weeks, which includes the knowledge loss in that time.

3. The participants came from study programs taught by the researcher. Thus age, background, and discipline of the participating students were determined by the study programs included in the study.

Please see section 6.2 for all identified study limitations.

4.2.2 Learning objectives and content areas

The laboratory program followed main learning objectives that aimed to provide students with a sound understanding of energy storage systems. Upon completion of the course, students were expected to comprehend the most important parameters (e.g. internal resistance and open-circuit voltage) and the characteristic behaviour (e.g. temperature dependencies of parameters) of battery cells. They should have also determined these parameters independently through suitable experimental setups. The students were expected to comprehend the effects when connecting battery cells to build energy storage systems. Emphasis was placed on the design of cell-type independent experiments, as cell types are expected to change over a student's career as an engineer.

Laboratory experiments covered four content areas:

- (A) contact and isolation resistance (for details see section G.5),
- (B) open-circuit voltage (see section G.6),
- (C) internal resistance and power (see section G.7 and section G.8), and
- (D) the energy content of cells (see section G.9).

Within these areas, seven laboratory experiments in both modes were developed:

- (A1) Low Resistance Measurements: on procedures for low ohmic measurements (Kelvin measurement);

- (A2) Contact Resistance: experiments with a variety of typical electrical connections in battery systems to determine exemplary contact resistance values;
- (A3) Isolation Resistance/Flash-over Voltage: handling appropriate measuring equipment to determine insulation resistance;
- (B) Voltage of Lithium-ion Cells: the dependence of voltage on the state of charge using two different types of cells;
- (C1) Internal Resistance: the influence of internal resistance on the efficiency of a battery system (covering AC- and DC-methods to measure internal resistances including the temperature dependence as well as applying industry standards like ISO 12405-1);
- (C2) Power: the maximum discharge rate of battery cells (covering dependency of the maximum discharge power on the state of charge, pulse duration, and temperature);
- (D) Energy and Capacity: the capacity of lithium-ion cells and factors influencing capacity (including calculating the efficiency of charge and discharge cycles).

For each of the content areas, a session with practical experiments and a session with computer-aided simulations were created. Content areas, experiments and learning objectives are presented in the appendix; see page 277.

4.2.3 Design of teaching experiments and instructions

Each laboratory experiment was developed to cover both modes.

All experimental procedures were created so that each individual step (e.g. starting a discharge current of 1 A for 2 s, followed by a rest phase to record the voltage response) was the same for both users of simulated and hands-on equipment. Since Chamberlain *et al.* [66] discovered that the intensity of guidance could strongly influence students' research activities, only a single set of written instructions was used in both modes. The instructions included preliminary questions, guidelines for the experiments, and advice on collection and analysis of data.

To facilitate critical thinking and to reduce the influence of teacher supervision, all experiments were designed to allow students to execute tasks independently in a supervised environment. All learning objectives could be met by following the set instructions without further help from the instructor. Moreover, laboratory instructions explicitly encouraged learning through trial and error (compare [116]).

Instructions were provided to the participants of each study run as a group in order to minimise the influence of individual students' prior knowledge on the effectiveness of laboratory learning. All laboratories were supervised by the same instructor.

The research study consists of two versions of study runs. An extended version, called in the following text "German" and a shortened one, allowed for additional research runs entitled “English”. It is essential to mention that the coherence of the language with the version of the study run was not intended – but caused by students’ nature in the available study programs.

Full 4 content-area study runs
Students who participated in one of the study runs in German conducted laboratory experiments that covered four content areas (content area C was covered over two sessions). Laboratory experiments took two hours and fifty minutes each. After the completion of all laboratory experiments, students used their own data in a workshop to parametrise a battery simulation model. All necessary basic theoretical knowledge was gained from a lecture course on battery systems that was conducted in the same semester (four hours per week). During the study runs in German, participants were asked to prepare a written laboratory report for each content area prior to the following session. Submission of all laboratory reports was required to be admitted to the final examination (written) of the lecture course.

Shortened 2 content-area study runs
Shortened English versions of the laboratory experiments were designed for participants of the summer school runs (R3, R7), the guest laboratory trial (R4), and the master’s degree program (R5, R8). These shortened versions covered two content areas (B* and C*) and took approximately two hours each. As preparation for these experiments, participants received a short introduction to lithium-ion cells.

Laboratory sessions
Each session was conducted as follows: After meeting in the laboratory room (hands-on experiments) or the computer lab (simulated laboratories), that session’s experiment was introduced and the students had the chance to ask questions regarding the experimental procedure. Afterwards, the students were grouped into small working teams and connected the prepared cell to the device (hands-on) or started the simulations (simulated).

Every team worked autonomously in a supervised environment following written instructions. To start an experiment, students defined a current and voltage sequence (including a temperature profile in some cases) for each measurement. All laboratory experiments consisted of a series of measurements to collect data (current, voltage, and temperature over time), which were evaluated before the next measurement or later at home, to produce the requested graphs (e.g. internal resistance over temperature) or conclusions.

The parameters of the battery cell during simulations and hands-on experiments were displayed on the computer screen as graphs and values. *Hands-on* and *hidden simulation* laboratories were equipped with tangible devices that additionally displayed the momentary values of current and voltage.

All learning targets were addressed in the proposed experimental procedure and did not require essential explanations from the instructor.

4.2.4 Laboratory environment, hands-on devices and simulations

A safe and easily manageable battery test system was developed for the hands-on laboratory experiments (for details see section F.2). It supported temperature-dependent experiments with different cell types including lithium-ion cells and incorporated a redundant safety shut-off module that protected students from being injured.

All measurement equipment was controlled by a computer program (Control-GUI), which was the same in both modes (Figure 4.3). The program allowed for the design and execution of all test sequences for battery analysis. Therefore, exercises in three content areas (B, C, and D) did not require any physical interaction with the measurement devices or batteries while the experiments were running.

The simulation mode was designed to achieve experimental results equivalent to those of the hands-on mode. To guarantee this, battery cells were analysed in order to parametrise the underlying simulation model (for details see section F.3). As a result, simulations closely imitated the actual behaviour of battery cells. Only input and output data were visible to students. The model itself, as well as the cell parameters and internal computed values of the simulated cell, have not been released to students (black box model). Such an arrangement ensured that the participants involved in simulations had the same information as their peers working in the hands-on condition. The simulation of cell behaviour always started after opening a graphical user interface regardless of whether experimental procedures were running or not. Thus, simulated cell behaviour also included aspects regarding the global design of the experiment (e.g. showing the cell cooling down for an appropriate time between two experiments).

4.2.5 Group formation

Conducting laboratory experiments in groups comes with its own set of opportunities and challenges.

Cooperative learning effects
A significant advantage of a counterbalanced study, is the possibility of compensating differences in the average performance of student groups through statistics/biasing. However, since students with more practical experience may behave differently than their colleagues with less practical skills, group interactions must also be considered [72]. Students' learning depends strongly on the other students working on the same topic [67, 69]. Individual students had previously gained different practical experiences which influenced their learning and interaction with their team and group members during the laboratory sessions.

Grouping method
The following method was used to create two groups (inter-homogeneous and intra-heterogeneous) of participants with similarly distributed *Amount of Practical Experience (APE)* (see section 4.3) in the German runs (R1, R2, R6, and R9): Participants

were sorted in descending order according to their respective APE. The first and fourth on the resulting list of participants were assigned to group A, while the second and third were assigned to group B. This procedure was repeated until all participants were assigned to either group A or B.

Unfortunately, organisational constraints only allowed for this approach to group-formation in the German bachelor's degree program. Due to their diverse backgrounds, participants of the international summer school (R3, R7) had to be manually assigned to their groups, based on their respective field and year of study, in order to create two homogeneous groups. For the guest laboratory trial (R4) and the master's program (R5, R8), group formation was arranged randomly.

Please refer to Table 4.4 for an overview of the employed grouping methods.

After being divided into two groups, students of each group were allowed to select partners in order to conduct laboratory experiments in small *working teams* of three to five. To ensure the same cooperative learning conditions, working teams remained together throughout the whole duration of laboratory work.

Conducting laboratories

For the simulation experiments of the content area A, a special web-based application was used, while the aforementioned black box simulation model (see section F.3) was used for all other content areas.

Groups in the study runs, which were running a full semester with 4 content areas (all in German), performed the same experiment on two different weekdays. In the shortened runs with 2 content areas (summer school, guest laboratory trials, and master's degree program; all in English), both groups conducted the laboratory simultaneously and in the same room, with one group using simulations and the other group using the hands-on equipment.

Please refer to Table 4.2 for an overview of the study runs.

As in the short runs both laboratory experiments were conducted on the same day, the groups switched to the other mode for the second content area. In the second research phase, it was possible to conduct both modes mixed in one room also in the German runs (R6, R9). The conducted experimental procedures are described in detail in chapter G.

4.2.6 Assessing student learning

To analyse the effect of the learning mode on student learning, their knowledge about the relevant topics was examined using written tests.

These tests were used to assess the effect of the learning mode on student learning using a mixture of item formats (descriptive, single-choice, multiple-choice, drawings, and graphs).

- The following is an example of a simple single-choice question: "The state of charge is stated in percent. Which physical dimension does it relate to?

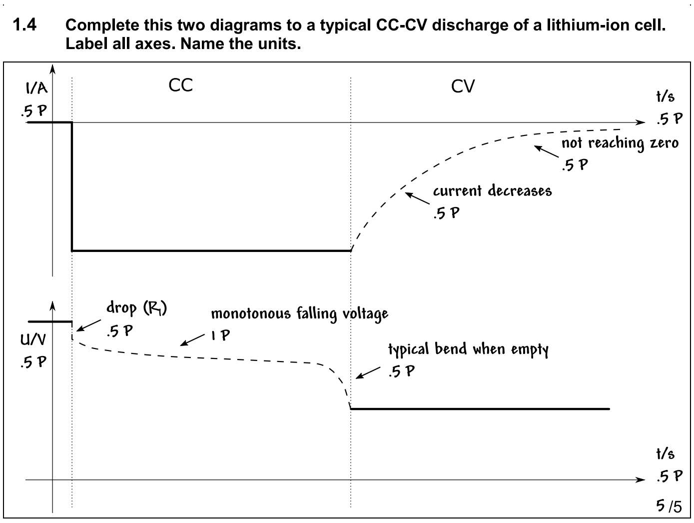

Figure 4.2: Example task of the written tests. Students had examined Constant Current/Constant Voltage (CC-CV) charge and discharge in their experiments. The text CC and CV and the straight lines were given. The students were expected to identify the two graphs and label the axes. Additionally, they were asked to add the missing cell behaviour.

(Which quantity is stated in percent?)". Different physical dimensions (voltage, current, charge, impedance, energy) were given.

- Graph tasks involved describing test procedures or battery cell behaviour (e.g. internal resistance dependent on temperature). A typical graph task is shown in Figure 4.2.

- Other tasks required students to explain relationships based on sketches, for example, arrangements of measurement equipment or the physics behind temperature gradients in a cell. Some tasks also included calculations, for example the discharge time depending on a given C-rate. Here, students were expected to carry out calculations before answering the question.

- In addition, the test included questions on typical values covered in the respective laboratory experiments, e.g. the allowed voltage range of a lithium manganese dioxide cell. For these questions, a reasonable tolerance was given for right answers; it was not necessary to remember the values precisely.

The employed tests are described in detail in the appendix starting from page 358.

The participants were given 10 minutes to complete each test. Since the test environment influences test results [117], tests took place in the same environment regardless of the experimental condition: To achieve such uniform test environment in study runs R1 and R2, all tests were written in a computer pool before the students changed rooms to conduct the next experiment.

Participants of the full semester study runs held in German took the tests one or two weeks after the respective laboratory exercises were conducted, prior to the next laboratory session. For both groups, equal periods of time between experimentation and the associated groups were allowed to pass in order to equalise the influence of time on the ability to remember (compare [39]).

Students that participated in the short study runs conducted in English were handed a single test between one and two weeks after completion of the laboratory work. These tests covered all the materials studied during the laboratory sessions. Tests submitted by students who did not attend the respective laboratory experiment were excluded from the experimental data.

In the full semester German runs (R1, R2, R6, R9) and the IEEE guest laboratory (R4), the test results did not have any influence on the participants' grades. In the other runs the test questions contributed approximately 20% to the overall grade for the class.

Tests submitted by students who did not attend the appropriate laboratory experiment were excluded from the experimental data. All tests were graded by the same person using a positive point system. Percentages of scored points relative to the maximum score were calculated per content area. Partial credit was awarded. The test questions/tasks are described in detail in section H.1.

In order to determine which laboratory mode produced better student learning, the test results in the two experimental conditions, hands-on and simulated, were compared after the test results were normalised for each single tested topic (see section 5.2).

4.2.7 Evaluation

Test performance

Data transformation was used in order to account for different group sizes, and to minimise the potential influence of differences in the difficulty of the content areas. Slight improvements of the lessons between the research runs might have also influenced students' success in different content areas. To counterbalance this, the difference between a student's individual scored percentage and mean percentage of all participants in the same test (e.g. mean result of all students of R1 in the content area A) was calculated and divided by the standard deviation of the respective content area from that run. This method is called *Studentisation* ([118], section I.5), and its result is considered the *student's performance* on a single test.

A value above zero indicated that the student's performance was above the average of all participants in the respective content area, a value below zero indicated a below-average performance.

Average test performances and respective standard deviations are presented in Table 5.3 to Table 5.12.

In order to determine which laboratory mode produced better learning, the averages of all students' test performances were compared according to the respective experimental condition.

Students' superior learning mode
While test performances were weighted per test, it was interesting to evaluate the mode which leads to better learning for an individual participant.

To determine the *superior learning mode* of a student, the differences between a participant's average test performance after hands-on experiments and his/her average performance after simulated experiments was calculated. If the value was positive, the participant performed better after hands-on experiments, if it was negative, the participants learned better with simulations (1st phase) or hidden simulations (2nd phase).

Test statistics
The data were analysed from three different angles:

- Test performance weighted per test results.
- Students' average test performance.
- Students' superior learning mode.

Independent-sample t-tests were employed to compare the performance of the different modes or groups of participants.

Shapiro-Wilk tests for normal distribution were applied to all compared subgroups. In case of significance (data of a group was *not* normally distributed) the regarding group data were marked with ^{s}. Non-parametric Mann-Whitney U-tests were performed, and the result of the Mann-Whitney test was considered instead of the t-test result.

As explained on page 364 in the appendix, the Mann-Whitney U-test analyses the difference of shape of the subgroups without focusing on the median. In cases where the U-test was significant, the data were analysed further to isolate the reason of significant difference.

The data analysis is presented in section 5.2. A summary of the findings is presented in subsection 5.2.4.

4.2.8 First research phase (R1–R5): hands-on experiments vs. simulated experiments

As shown in Table 4.2, the first phase was carried out between 2016 and 2017 in five study runs (R1 to R5). Participants in R1 and R2 were German bachelor students of "Electrical Engineering and Electric Mobility", a major with a focus on electrical engineering and electric mobility. These participants were in their 2nd-year and were instructed in German (R1 and R2). Participants of the other runs were instructed in English (R3 to R5). The educational research was explained to the participants in detail during the first session. Due to meeting in different rooms (laboratory vs. computer pool in R1 and R2) and due to the absence/presence of measurement devices and batteries, students were always aware of the experimental condition.

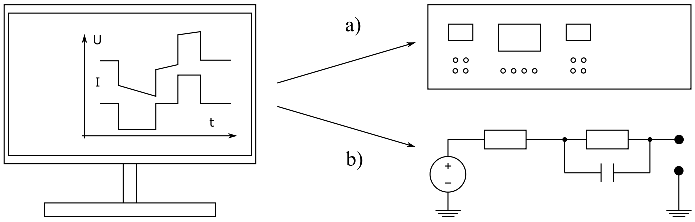

Figure 4.3: To avoid influence on the learning, the hands-on device (a) or the simulation model (b) was controlled by the same graphical user interface. In the first phase of the research, the students were clearly *aware* of the learning mode.

For the simulated experiments, a black-box simulation model that imitated hands-on battery behaviour was designed and calibrated (Figure 4.3). See page 272 in the appendix for a description of the technical details of the employed black-box simulation model.

Figure 4.8 shows the working environment during simulations. The setting for the hands-on experiments is presented in Figure 4.5.

4.2.9 Second research phase (R6–R9): hands-on experiments vs. hidden simulations

The first phase of the study found statistically significant differences in student learning in favour of the hands-on mode (see subsection 5.2.1).

Reasons for the second phase

Since laboratory experiments in both modes were created to be very similar, the reason for this difference needed clarification. It was hypothesised that the difference in learning originated from different perceptions on the efficacy of hands-on and simulated experiments. In order to check this hypothesis, a modified second phase was conducted. It was performed in four study runs in 2018 and 2019 (R6 to R9), see Table 4.2.

Although the simulation model for the first phase was created to be as authentic as possible, it was impossible to definitely exclude unrecognised weaknesses of that model as an influence on student learning in the first phase and a disadvantage of learning with simulations.

Modes of the second phase

Phase two followed the same main concept as phase one. This time, in order to ensure that the perceived laboratory mode would not influence the results, all participants used hands-on equipment in a laboratory environment in *both* conditions. Subjects from one group used unmodified hands-on equipment (*"hands-on"*). Subjects from the other group only thought that they were conducting hands-on experiments (*"hidden simulations"*). In fact, their equipment was modified to display the results of simulated experiments (*"hidden simulations"*).

In the *hands-on* condition, real measurements were shown.

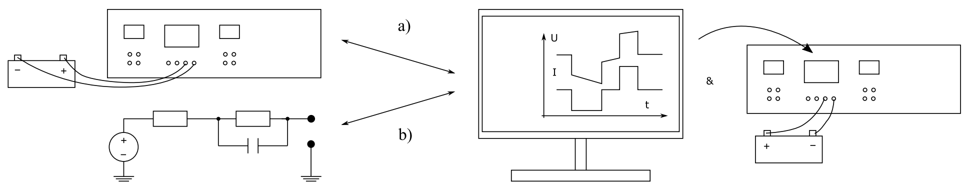

Figure 4.4: “Hands-on” vs. “hidden simulation”: (a) The hands-on device (with a real battery cell) or (b) the simulation model was controlled by students, using the same graphical user interface. The results were displayed both on the GUI and the test bench’s displays while a battery cell was actually connected to the equipment. This way, students confronted with b) were *not aware* that they were facing a simulation. It was not possible to discern student groups using simulation from hands-on groups when entering the laboratory.

Figure 4.5: Laboratory environment/setting during *hands-on* experiments in the first phase and during *hidden simulations* as well as *hands-on* experiments in the second phase.

In the *hidden simulations* condition, the results of simulated battery behaviour were displayed. Participants of hidden simulations were presented with data calculated by the same simulation model as in the first study phase. In order to convincingly imitate the hands-on condition, the output was not only displayed on the computer screen but also on the measurement devices (see Figure 4.4).

Participants were again aware of the ongoing engineering education study “hands-on vs. simulations”. But in the second research phase, students were *not aware* of differences in data sources during their technical experiments. It was expected that all students would perceive that they are conducting hands-on laboratory.

Much effort was necessary for arranging these laboratories. Simulations (model, parameters) stayed the same as in previous runs, but computer software and device firmware (see section F.2) were modified to show simulation results in real time on the device displays. A battery was connected, but no actual current was applied. Even the open/close actions of the attached safety box were triggered to create the same audible sound of the opening/closing relay.

For a visitor, it was not possible to distinguish the hidden simulation set-ups from the hands-on set-ups on entering the laboratory, or even to determine that laboratory conditions differed between stations (see Figure 4.5). The number of participants in the German runs in the first research phase required both modes to be conducted on different weekdays. In the second phase it was possible to perform the crossover in the same room for both groups, further excluding the influence of a possible difference in teachers comments on the two week days.

By collecting tests of the two laboratory rows (employing the different modes) separately, code words could be attributed to the learning modes during all runs of the second phase.

Reasons to replace specifically the simulation mode with ***hidden simulations***
It would have also been possible to replace the real *hands-on* mode by *perceived simulations*, which would present the participants' results of real running experiments in the laboratory next door covered as simulated results while participants work sitting in a computer pool (hidden hands-on). Presenting data sources of both modes as simulated data would also allow for checking the the simulation model.

The first hurdle to the hidden hands-on approach is practical: Lindsay [13] used recorded experimental data to ensure equal results in his study, which was impossible here as students were able to directly influence the experiment, e.g. freely choose the actual employed current profiles.

Finally, the reasons to continue the second phase with the *hidden simulations* variant were:

1. In the first phase, no differences were identified after hands-on experiments (see subsection 5.2.4); the differences between subgroups regarding simulations were the determining factor on the overall outcome.

2. In the hands-on mode, former VET participants were performing similar to students who had no VET education. Former VET participants had, in contrast, less good results using simulations. Based on the first phase results, the usage of perceived simulations may disadvantage former VET participants for the rest of the ongoing study. The chosen modes (present as hands-on results) for the second phase allowed both groups to perform equally well.

3. Usage of hands-on experiments increased the chance for more research runs, as program managers generally prefer hands-on experiments in a laboratory environment over simulations in a computer pool.

4.3 Amount of Practical Experience (DS-B)

This section discusses the general research methodology for collecting and evaluating data regarding DS-B. The results can be found in section 5.3, with analysis and discussion in chapter 6.

Personal interest can be seen as a prerequisite for learning success since it has a profound effect on cognitive functioning. Students with a personal interest tend to be more attentive and focused, and thus have greater success in gaining knowledge [119].

A widely renowned concept to describe interest on a general level is Holland´s RIASEC-typology, in which six personality types are distinguished (Realistic, Investigative, Artistic, Social, Enterprising, and Conventional) and connected to a wide array of attitudes, values, and self-beliefs [120]. Two well-known test sets for the RIASEC-typology were used in later study runs to allow for comparison, see section 4.4.

However, this concept might be too broad to capture students' interests in a science-based topic [121], as saturation effects of the standard RIASAC scales were possibly observed among STEM students. Therefore, a questionnaire comprised of questions about prior experiences in practical tasks as well as self-beliefs was created. The items were designed to resemble questions used by inventories based on the RIASEC-model, but with higher specificity towards technical practical experiences.

4.3.1 Assessing the *Amount of Practical Experience (APE)* of the participants

Questions used to assess the APE started with "I have..." or "I am able to..." and were for example "...made a thread on a bore hole." or "...realised a function using a self-made circuit diagram." (for the full list, see Table 5.13). The questions were rated on a 4-point Likert-type scale (yes (1.5) / rather yes (.5) / rather no (.5) / no (1.5)), and based on the author's "get-to-know-questionnaire" which had been in use in the years before the thesis to fit practical project topics to student groups.

This questionnaire was distributed among the participants of six study runs (German Electrical Engineering and Electric Mobility (B. Eng.) and English Renewable Energy Systems (M. Sc.) runs, see section 5.1) during the introductory session.

The average value of a student's responses was considered their *Amount of Practical Experience (APE)*.

Definition of APE

A value above zero indicated that the participant had prior practical experience with most of the items listed in the questionnaire, while a value below zero indicated little prior experience. As the Likert scale values were chosen in equal steps, a full "yes" contributed double in comparison to "rather yes". The same applies to "no" and “rather no”. E.g. an APE value of 1.5 meant that the participant experienced everything in a certain dimension, and a value of -1.5 meant a participant had no experience with any of the items.

Enhancements during the study

Starting with the second research run (R2) the survey was enhanced to collect additional information about students’ possible completion of a German Vocational Education and Training (VET) program (see subsection 3.2.1) before their studies.

From here on out, students/participants who completed a VET degree before enrolling at university will be referred to as *VET,* the other students as *non-VET.*

Usage of the APE dimension for the research

The APE was employed for three purposes:

- To create similar group environments for the main study, participants were distributed in accordance with their level of practical experience. Details are given in section 4.2.
- To investigate the background of specific subgroups (VET, non-VET, German, international).
- To investigate the influence of APE on student laboratory learning.

4.3.2 Evaluation

The data regarding the Amount of Practical Experience was evaluated as follows:

A factor analysis was used to check the responses to the employed items regarding possible sub-dimensions. The reliability of main and sub-dimensions was validated.

Difficulties were reported, in order to allow other researchers to make comparisons with their results/student groups.

Then, the APE dimensions were used to characterise three types of participants: German VET, German non-VET and international participants.

Finally, the APE and the sub-dimensions were correlated

- with the overall test-performances and
- with the superior learning mode

of the participants in both research phases. These correlations were then compared between the three different types of participants.

The data analysis is presented in section 5.3. A summary of the findings is presented in subsection 5.3.10.

4.4 Personality/RIASEC (DS-C)

In this section, the general research methodology for collecting and evaluating data regarding DS-C is discussed. The results are presented in section 5.4 and discussed in chapter 6.

A widely renowned concept that describes interest on a general level is Holland´s RIASEC-typology, in which six personality types are distinguished (Realistic, Investigative, Artistic, Social, Enterprising, and Conventional) and connected to a wide array of attitudes, values, and self-beliefs [120].

Personal interest can be seen as a precursor for learning success, since it has a profound effect on cognitive functioning, including but not limited to focus and attention as well as knowledge gain [119].

- The usage of well-known test sets allows for the comparison with published difficulties.
- It may help to understand the reasons for differences in learning of VET and non-VET students in the laboratory of the main study. Even when the RIASEC concept might be too broad to capture students' interests in a science-based topic [121], results can contribute to investigating the different backgrounds of German VET and non-VET UAS students further.
- Also, it may enable an explanation of what the APE-dimension signifies.

4.4.1 Methodology

Two established personality inventories based on the RIASEC model were used: "Allgemeiner Interessen-Struktur-Test" (AIST) [122], and "Explorix – das Werkzeug zur Berufswahl und Laufbahnplanung" [123]. Both inventories were already adapted to German culture and were available in the German language.

The test sets were transferred to a Qualtrics online questionnaire [124] and participants of the last German runs were asked to participate.

In the following, a summary of John L. Holland's RIASEC-dimensions of and the two employed test sets are presented.

4.4.2 RIASEC: Six personality types according to John L. Holland

The development of Holland's theory and the six dimensions began in the 1950s [125]. Holland's theory plays a central role in occupational psychology, because it is one of the best-researched theories with a solid empirical foundation [126, p. 12].

Holland described six interests and personality types and, by analogy, six types of work environment. The personality types are described in detail in "Making Vocational Choices" [120, p. 21-28]. Holland refined his work to include gradual and structural differentiation.

The six categories cannot only be seen as primary types, but instead span a six-dimensional space. Nevertheless, there are relations and correlations between the dimensions. Holland found the strongest relationships between neighbouring dimensions in the circular RIASEC arrangement, but this has not been unequivocally confirmed in later studies. [126, p. 14]

According to theory, the six types characterise ideal types which serve as basis for comparison [126, p. 12]. They are summarised in the following paragraphs.

Code R (Realistic): manual-technical, "Doers"
People of this type like to work with their hands and things/objects, and they are interested in tools and machines [123, p. 15]. They like to be outdoors and are physically active [123, p. 15]. In typical cases, they are characterised by the following characteristics: natural, down-to-earth, practical, self-confident and somewhat conservative [123, p. 15]. They prefer to approach problems by doing something rather than talking/sitting/thinking about them or trying to apply an abstract theory [120]. Their interests tend towards mechanical/scientific topics rather than aesthetic or cultural areas [120].

Recommended professions involve concrete objects that are worked on by hand or with tools. They often include the use of machinery and technical equipment. Such professions require manual dexterity and an understanding of technology. Some professions also take place outdoors and require physical robustness and endurance. Occupational areas: Craft, technology, agriculture. Examples include carpenters, mechanics, farmers, and electricians. [123, p. 23]

Code I (Investigative): investigative-researching, "Thinkers"
People of this type like to work with data and to immerse themselves in mental or scientific problems. They analyse, study, learn, read, write and calculate with passion rather than to act [123, p. 15]. In typical cases, they can be characterised as curious, inventive, intellectual, accurate, logical, rational and performance-oriented [123, p. 15].

Recommended professions involve problems that are examined with the help of logical thinking, new ideas, precise observation and scientific methods. These professions require a high level of thinking ability and curiosity as well as the willingness to familiarise oneself with a subject in detail. Occupational areas: Science, research. Examples include physicists, researchers, laboratory assistants. [123, p. 23]

Code A (Artistic): artistic-creative, "Creators"
People of this type like to express themselves creatively or linguistically. They like to occupy themselves with unusual ideas and unique materials, music or culture. Aesthetics are important to them. In typical cases, they have the following characteristics: imaginative, creative, expressive, intuitive, open, sensitive, unconventional, idiosyncratic and idealistic. [123, p. 15]

Recommended professions involve opportunities for artistic design with the help of materials, musical instruments or one's own body. They often aim to express ideas or ideal conceptions, design and prettify something or enrich society culturally. These professions require high artistic, creative, musical or linguistic talent as well as tireless practice and improvement of artistic abilities. Occupational areas: Art, music, theatre, writing. Examples include musicians, actors, designers, and writers. [123, p. 23]

Code S (Social): educative-nursing, "Helpers"
People of this type like to take care of other people. They educate, teach, advise, care, heal and care for physical, mental or spiritual well-being. In typical cases, they are characterised as friendly, helpful, warm-hearted, compassionate, understanding, sociable, idealistic and partly instructive. [123, p. 15]

Recommended professions are about helping other people – caring for them, advising them or training them. The focus is on the mental, spiritual or physical well-being of children or adults. These professions require a great willingness to help, a high level of empathy and a way with people. Occupational areas include education, counselling, and health care. Examples include teachers, nurses, psychotherapists, and social workers. [123, p. 23]

Code E (Enterprising): leading-selling, "Persuaders"
People of this type like to motivate, convince and lead. They like to take care of economic planning and financial goals. Typically, they are characterised as follows: self-confident, motivated, stirring, success-oriented, ambitious, dominant, responsible and sociable. [123, p. 15]

Recommended professions focus on economic goals, managing and selling. They want to convince and motivate others (e.g. to buy a product or perform a service). Enterprising professions require economic thinking and a convincing appearance, often incorporating organisational and administrative skills. Occupational areas: Management, sales. Examples include hotel managers, politicians, salespersons, and advertising agents. [123, p. 23]

Code C (Conventional): ordering-managing, "Organisers"
People of this type like to work neatly, accurately and well organised in an office. They work, control and transmit numbers or texts. Clear rules are important to them. In typical cases, they are characterised by the following characteristics: careful, precise, detail-oriented, persevering, neat, practical, adapted, conscientious. [123, p. 15]

Recommended professions consist of the orderly and systematic handling of numbers, data or information. The focus is on reliable execution, administration or bookkeeping, as well as excellent organisation and control. Occupational areas: Commercial, office and visual professions. Examples include secretaries, commercial employees, cashiers, correspondents. [123, p. 23]

4.4.3 Explorix-Test

Explorix is a questionnaire designed for the career planning of apprentices, technical students, high school students or adults. It is used for career guidance. The original purpose of the procedure is to show participants an extended range of career choices. At the same time, the procedure helps to focus on an individually adapted section of the professional world.

The original German-language Explorix can be carried out, evaluated and interpreted both online and on paper by the participants themselves, which requires a certain degree of independence and confident use of the German language. The Explorix is based on four different types of questionnaires: activities (“Tätigkeiten”, capabilities (“Fähigkeiten”), professions (“Berufe”), and self-assessment (“Selbsteinschätzung”).

Two components – activities and capabilities – were used.

*Activities, Exp_Inte_**

A list of activities (occupational and spare time) is given. Participants are asked to decide if they (1) are interested in doing them or (0) not (binary choice). There are 11 items per dimension.

*Capabilities, Exp_Capa**

A list of capabilities is given. Participants are asked to decide if they (1) are able to perform them or (0) they are not able to perform these or never performed them (binary choice). There are 11 items per dimension.

4.4.4 AIST-Test

The German-language AIST-R records school and professional interests based on the model of J. L. Holland. It consists of 60 items measuring the six dimensions of interest (RIASEC). It targets adolescents and adults. It is also intended for career orientation, career decisions, and internal career and personnel decisions. The degree of interest is measured on a 5-point Likert scale ranging from (1) “That does not interest me at all, I do not like doing that” to (5) “That interests me a lot, I like doing that a lot”. Independent of the proposed scales in the tests, for this publication, all item scales were mapped to the range 0 to 1. [122]

4.4.5 Evaluation

Independent of the proposed scales in the original test sets (Explorix, AIST), all item scales were mapped to the range 0 to 1 for reasons of comparability.

The study participants were measured against the published difficulties of the test sets (in general, regarding age, educational level, and job).

VET and non-VET subgroups were compared regarding their absolute RIASEC category values (using independent samples t-tests and non-parametric U-tests) and the level of correlation between these categories.

Finally, a possible correlation between the APE dimensions and the RIASEC categories was investigated using Spearman rank-order calculations.

The data analysis is presented in section 5.4. A summary of the findings is presented in subsection 5.4.5.

4.5 Survey for participants' subjective opinion after conducting the individual experiments (DS-D)

In this section, the general research methodology for collecting and evaluating data regarding DS-D is discussed. The results are presented in section 5.5 and discussed in chapter 6.

DS-D compared self-reported opinions after conducting experiments in both modes. These opinions were reported on the previous laboratory session in general. Students were asked only to evaluate their experience in the last session, not specifically for feedback on the laboratory modes. Thus, with DS-D, the comparison was achieved *indirectly* and unbiased, by comparing the respective feedback given after simulations and after hands-on experiments, without directly involving students' preconceptions on the modes.

Only data of bachelor students of one study program at one University of Applied Sciences was available, as DS-D is only based on answers of students from THI (R1, R2, R6, and R9).

4.5.1 Methodology

In the German Runs, the content area and the teaching method changed each week.

In 2016, a standard feedback form regarding the previously conducted laboratory experiments was installed at different laboratory classes in the study program "Electrical Engineering and Electric Mobility (B. Eng.)". The initial installation of the default feedback form was independent of the research. The study program leader aimed for iterative improvement of the recently installed laboratory experiments (Physical chemistry laboratory in semester 3 and Energy storages laboratory in semester 4).

All participants of the German runs were asked to express their opinions on the laboratory learning during the previous lesson. The same questions were asked regardless of the learning mode used in the previous experiment. The questionnaire was conducted fully anonymously using the university's digital learning environment Moodle [127]. As the two compared groups were recorded separately (two subgroups in the Moodle course), it was possible to correlate the answers to the learning mode.

Responses to the following survey questions were compared in order to evaluate the laboratory learning in both learning modes:

- (a) "By experimenting, I gained new insights/comprehension." (German: "Ich habe heute durch den Versuch neue Erkenntnisse gewonnen.")
- (b) "At which point in the experiment did you encounter the biggest problem proceeding?" (German: "An welcher Stelle im Versuch hatten Sie am meisten Probleme voranzukommen?") This item was a free text question, which was not compulsory. For data evaluation, the information was coded to 1 if any

problem was mentioned or to 0 if students wrote nothing or expressed they had no problems.

- (c) "The procedure of the experiment is quite difficult/doable/easy" (German: "Die Versuchsdurchführung ist recht schwer/machbar/leicht") The answers were coded in a 3-point Likert scale.
- (d) "The content of the experiment is also relevant to me outside the university; I can imagine that it will be beneficial for my future professional life." (German: "Der Inhalt des Versuchs hat auch außerhalb der TH Relevanz für mich; ich kann mir vorstellen, im Berufsleben Gewinn aus dieser Versuchsdurchführung zu ziehen.") The answers were coded in a 5-point Likert scale: fully agree / somewhat agree / maybe / somewhat disagree / disagree.

Enhancements between the runs

Starting from R6, a question regarding the participants' code-words was added to the form to allow for comparison with the test results and the info about VET.

4.5.2 Evaluation

In both research phases, students' answers regarding the learning mode were compared using non-parametric Mann-Whitney U-tests. Additionally, the correlations between the four test items were investigated (in sum, for the modes separately, and for differences between both modes).

In the second phase, the answers were additionally investigated with regards to the completion of a VET before studies and the student's performance in the test regarding the experiment referenced in the respective questionnaire.

The answers concerning hands-on experiments in both phases were compared using Mann-Whitney U-tests to detect whether students perceived the hands-on mode in the second phase as identical to that of the first phase..

An analysis of the data is presented in section 5.5. A summary of the findings is presented in subsection 5.5.5.

4.6 Survey for participants' and other persons' subjective opinions on the learning modes (DS-E)

In this section, the general research methodology for collecting and evaluating data regarding DS-E is discussed. The results are presented in section 5.6 and discussed in chapter 6.

An online questionnaire was developed to collect a broader view of opinions on the learning modes. The participants were asked to answer *direct* statements regarding the learning modes.

The participants of this survey were partly comprised of participants of the main study and partly of students and employees of other universities without connection to the laboratory experiments, both in Germany and abroad, in order to gain a general view.

4.6.1 Methodology

The additional questionnaire was developed to provide a broader database regarding subjective opinions on the learning modes. It was designed to test whether the general opinion on the laboratory modes differs between university types, countries, and students/lecturers.

The questionnaire, described in this section, primarily aimed to discover differences between students of UAS and "full" universities. Thus, it provides insight if the results of the main study allow for generalisation.

Next to questions discovering the background of participants, paired single-choice questions, standard single-choice questions, and free text questions were asked.

The background information covered undergraduate vs. postgraduate status, years of practical experience, Bachelor or Master studies, the semester of study, and completion of a VET degree before enrolment.

It was distributed online (via Qualtrics [124]) amongst several universities in Germany (incl. THI), Europe and the RMIT in Australia. Thus, this database included persons who did not participate in the battery laboratories.

Paired single choice questions

Five pairs of questions were used to compare opinions on the laboratory modes directly. All used a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree. The neutral value for the answers was coded 4. The questions were asked in a group for each mode. The order of questions within the groups, as well as the order of both groups, were randomised for each participant.

- Pair 1: I learn extremely well when I conduct *hands-on laboratory* experiments. — I learn extremely well when I conduct *computer-simulated* experiments.

- Pair 2: *Hands-on* laboratory experiments offer authentic laboratory experience. — *Computer-simulated* laboratory experiments offer authentic laboratory experience.

- Pair 3: The outcomes of *hands-on* laboratory experiments allow learners to accurately familiarise with the actual behaviour of rechargeable batteries. — The outcomes of *computer-simulated* laboratory experiments allow learners to accurately familiarise with the actual behaviour of rechargeable batteries.

- Pair 4: The majority of students/participants believe that conducting *hands-on* laboratory experiments provide them with the best opportunities to study the real behaviour of rechargeable batteries. — The majority of students/participants believe that conducting *computer-simulated* laboratory experiments provide them with the best opportunities to study the real behaviour of rechargeable batteries.

- Pair 5: If I had the opportunity, I would always participate in a *hands-on* laboratory experiment. — If I had the opportunity, I would always participate in a *simulated* laboratory experiment.

Single choice questions

- Q1: I always simulate technical problems by myself (e.g. using SPICE or Matlab) in order to understand them better and/or to check my assumptions. (7-point Likert scale from 1 = strongly disagree to 7 = strongly agree)

- Q2: Students use simulations in their studies far too often, instead of trying things out. (7-point Likert scale from 1 = strongly disagree to 7 = strongly agree, the neutral value for the answers was coded 4.)

- Q3: Which learning mode for laboratories would you recommend for university students?

  - Simulations, alone, at student's preferred time, online (remote, from home or any place)
  - Simulations, in small working teams, at a specific time, at the university/company
  - Hands-on, in small working teams, at a specific time in the lab
  - None of these, my proposal is instead (in that case a free text response was provided)

Free text questions

1. Describe in one short sentence your general opinion on simulations in teaching laboratories.

2. Describe in one short sentence your general opinion on hands-on experiences/ experiments in teaching laboratories.

4.6.2 Evaluation

The overall answers regarding the paired questions were analysed using a Wilcoxon signed rank test. One-sample Wilcoxon signed rank tests were employed to check for statistically significant deviation to the neutral value of the unpaired single choice questions.

Free-text responses were categorised in positive, neutral, and negative responses for both modes separately.

Then, possible correlations of answers with participant characteristics were investigated. To support this comparison, also the trends of the question pairs were calculated. Mann-Whitney U-tests (as well as paired-samples t-tests in case of normal distributed data) were computed to compare the participant characteristics regarding their statements. Answers were compared according to these participant characteristics: Student/graduate, German/international, academic level, years of practical experience, for students: bachelor/master, VET/non-VET, type of university.

An analysis of the data is presented in section 5.6. A summary of the findings is presented in subsection 5.6.6.

4.7 Objective data – THI university database (DS-F)

In this section, the general research methodology for collecting and evaluating data regarding DS-F is discussed. The results are presented in section 5.7 and discussed in chapter 6.

4.7.1 Getting objective data from the university administration

German participants that completed a Vocational Education and Training (VET) before enrolling at THI (University of Applied Sciences Ingolstadt) performed the same as the rest of the students after hands-on laboratories, but under-performed after conducting simulated experiments compared to their non-VET colleagues (see section 5.2). It was therefore interesting to closer examine the differences between VET and non-VET students that resulted in different learning after simulated laboratories.

This main study was conducted fully anonymously using code-words. All correlated background data on participants in the study were self-reported / subjective. The research had not collected objective data to compare the background of participants with/without VET. This section analyses objective data to investigate the reasons for the above-mentioned significant difference between VET and non-VET-students in learning with simulated laboratories.

4.7.2 Methodology

The following information from the university administration (THI) was requested to get objective data. Please note that the THI Vice President for Teaching, Students and Alumni willingly gave permission to access the data and an Ethics Amendment was granted by the RMIT (see page 242 in the appendix).

Students

- year of birth,
- (optional) date of completing a VET (to determine first, if any VET was completed, and secondly, how much industry experience the student was able to collect before enrolling as a student at university),
- kind of school degree/university entrance qualification,
- date of enrolment,
- study program,
- (optional) date and reason of ex-matriculation,
- ECTS-points achieved in each individual semester and
- average grades of each semester, as well as the overall averages for the first 4 semesters and the first 7 semesters.

The database included data from all current and former students. The evaluation was focused on the STEM study programs (mechanical and electrical faculties, ignoring data of the business school). DS-F was not collected correlating to a specific study run. The data was anonymous (no clear names and no code words of the main study available), and it was not possible to correlate it to the participants of the central part of the study.

The data was also used to approximate any possible bias of students participating the main study by approximating the data sets corresponding to the study runs, using the study program, the cohort/year of matriculation and VET-information. This analysis can be found on page 96.

4.7.3 Evaluation

The data from the university database was analysed to assess potential differences between VET and non-VET students that could have resulted in differences in learning success (Non-parametric Mann-Whitney-U-tests).

Based on the data, the time between last school degree and enrolment at university was calculated as an indicator of industry work experience. The age of enrolment was correlated with the reason for un-enrolment (successful degree/failed studies). The reasons for un-enrolment were plotted dependent upon the semester for VET and non-VET students.

An analysis of the data is presented in section 5.7. A summary of the findings is presented in subsection 5.7.4.

4.8 Study runs and participants

Different types of study runs were used to answer the research questions. This section describes the different types of runs to allow the reader to understand the scope and kind of students participating in the runs.

A summary is given in Table 4.2, while demographics are shown in Table 4.3. Table 4.4 shows the methods used to create the compared groups.

Generally, with this type of research, it is difficult to increase the number of participants, as efforts to conduct study runs are high. Usually, a modification of the curriculum of study programs is necessary. Further, it is challenging to convince unrelated universities to include compatible guest lectures in their study programs. Therefore the outstanding share of study participants derived mainly from the university, which employed the author as a lecturer (THI). The most significant share of participants (R1, R2, R6, R9) and laboratory time derived from one study program: "Electric Mobility (B. Eng.)". On the one hand, this might be negative for generalisation. On the other hand, this program allowed for comparing VET data as the program is taught in German at UAS and thus, selected by many German learners with VET background. The participants are most likely comparable between years through the usage of data of four following years of the same study program.

Table 4.2: Study runs

Study Run		Year	Background	Location	CA	LT	Part.
R1	Electric Mobility (B. Eng.)	2016	German	THI	4	14	40
R2	Electric Mobility (B. Eng.)	2017	German	THI	4	14	30
R3	International Summer School	2017	International	THI	2	4	32
R4	IEEE guest laboratory	2017	International	TUC	2	4	[a]9
R5	Renew. Energy Syst. (M. Sc.)	2017	International	THI	2	4	18
R6	Electric Mobility (B. Eng.)	2018	German	THI	4	14	29
R7	International Summer School	2018	International	THI	2	4	39
R8	Renew. Energy Syst. (M. Sc.)	2018	International	THI	2	4	45
R9	Electric Mobility (B. Eng.)	2019	German	THI	4	14	26
All	All Study Runs						268

Note. CA = number of content areas. LT = Laboratory time in hours equivalent to 60 minutes
THI = Technische Hochschule Ingolstadt. TUC = Technical University Chemnitz
R1-R5: hands-on vs. simulations R6-R9: hands-on vs. hidden simulations
CA 2: (B*) open-circuit voltage, and (C*) internal resistance and power, 2:00 h each.
CA 4: (A) contact and isolation resistance, (B), (C1), (C2), and (D) energy of cells, 2:50 h each.
[a]: 18 participants while conducting the laboratory and lectures, 9 participated in the test

Table 4.3: Study runs: Demographics, VET and Amount of Practical Experience

Study Run		Female	Age / Years M	Age / Years SD	Semester	APE M	APE SD
R1	Electric Mobility (B. Eng.)	< 5%	23.3	2.2	4	.49	.33
R2	Electric Mobility (B. Eng.)	< 15%	23.4	2.5	4	.44	.58
R3	International Summer School	< 15%	23.4	3	3-9	n/a	n/a
R4	IEEE guest laboratory	n/a	n/a	n/a	n/a	n/a	n/a
R5	Renew. Energy Syst. (M. Sc.)	< 20%	25.8	2.4	8	-.08	.49
R6	Electric Mobility (B. Eng.)	< 10%	22.8	2.3	4	.51	.59
R7	International Summer School	< 10%	24.0	3.5	3-9	n/a	n/a
R8	Renew. Energy Syst. (M. Sc.)	< 15%	26.7	4.3	8	.10	.52
R9	Electric Mobility (B. Eng.)	< 5%	23.0	3.3	4	.56	.51
All	All Study Runs	< 15%					

Note. APE = Amount of Practical Experience as defined in section 4.3.

Table 4.4: Study runs: grouping method and test environment

Study Run		Grouping method	Test environment
R1	Electric Mobility (B. Eng.) 2016	APE	Computer-Lab
R2	Electric Mobility (B. Eng.) 2017	APE	Computer-Lab
R3	International Summer School 2017	Fields and year of studies	Classroom *
R4	IEEE guest laboratory 2017	Randomised	Classroom *
R5	Renewable Energy Systems (M. Sc.) 2017	Randomised	Classroom *
R6	Electric Mobility (B. Eng.) 2018	APE	Laboratory
R7	International Summer School 2018	Fields and year of studies	Classroom *
R8	Renewable Energy Systems (M. Sc.) 2018	Randomised	Classroom *
R9	Electric Mobility (B. Eng.) 2019	APE	Laboratory

Note. APE = Amount of Practical Experience as defined in section 4.3.
* = one test covered all content areas.
Computer-Lab = Students which performed on that day hands-on experiments changed the room for the test.

Table 4.5: German study runs: share of former VET participants

	University database (DS-F)						Survey (DS-B)		
	Beginning of studies			Fourth semester[a]					
Run	Students	VET	Share	Students	VET	Share	Students [b]	VET	Share
R1	65	36	55%	40	24	60%	[c]0		
R2	61	21	34%	34	14	41%	27	16	59%
R6	53	26	49%	30	19	63%	29	15	52%
R9	59	22	38%	36	18	50%	26	16	62%

Note. [a] Extrapolated by assuming the student did not repeat semesters.
[b] Not all participants answered the question.
[c] The question was introduced starting from R2.

Preliminary testing/piloting the laboratory experiments

Two preliminary runs were conducted to validate the quality of the laboratory experiments described in chapter G. In these runs, no data for the educational study was collected.

Institute of executive education, Electric Mobility students (M. Eng.), THI, 2016
The instructions and exercises in German language were piloted with students of a part-time master program in 2016 to validate the laboratory experiments before conducting the educational study.

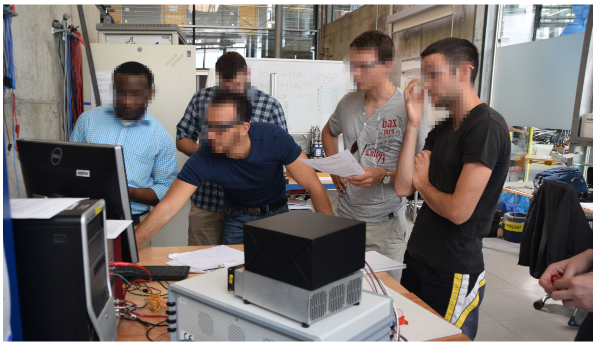

Figure 4.6: German pilot run: Institute of executive education, Electric mobility students (M. Eng.), THI, 2016

Erasmus+ STA staff mobility teaching, Ferrol, Spain, 2017
After translation of the instructions into English, and receiving additional funds for the transport of devices, the laboratory was conducted in the study program "Master Efficiency and Energy Use" at the University in Ferrol, Spain, in two weeks of April and May 2017. The students of a class of fourteen students were very interested in the battery topic. As it was the first run in English, it helped to improve the translation of the instructions (and thereby the comprehensibility) and prepared the later international runs in English.

Figure 4.7: English pilot run: Erasmus+ STA Staff Mobility Teaching, Ferrol, Spain, 2017

German runs, 4 content areas

R1, R2, R6, R9: Electric mobility (B. Eng.), summer semester, THI, 2016-2019
Most of the research data is based on the undergraduate students of the electric mobility bachelor program at THI for three reasons:

- The study program starts each winter semester. Thus, a cohort for educational research was available every year from 2016 to 2019.
- Cohort size usually ranges from 30 to 40 students.
- The laboratory is distributed throughout the semester, and includes several sessions. In these runs, 4 instead of 2 content areas contributed to the educational research.

As these runs dominated the research, the particulars of that study program are described in more detail in the appendix (see page 279).

The course "Electrical Engineering and Electric Mobility" represents the common electrical engineering bachelor programs at THI in terms of the average age of the students, which is 23 years. As the program is in German, usually locals (Germans) enrol in that program. The proportion of students with vocational training (German VET, subsection 3.2.1) at enrolment in this B. Eng. program (46%) slightly exceeds the average of VET graduates enrolled in bachelor programs at the THI Faculty of Electrical Engineering and Computer Science (M = 38%, SD = 6%, N = 7).

Figure 4.8: German runs: R1, R2, R6, R9: Electric mobility (B. Eng.), summer semester, THI, 2017-2019, the picture shows the simulations setting

As passing the laboratory module depended on the laboratory journal and report, it was possible to write all knowledge tests fully anonymously, using code-words for correlation. Test results did not influence students' marks.

English runs with two content areas

To have better opportunities to broaden the data-base and to include international students, the runs needed to be more flexible to fit the hosts' and study program manager's needs. Also, data sources needed to be excluded, as the lessons were reduced in time to fit these programs. Instead of writing individual tests for each crossover/content area, one joined knowledge test was written after one to two weeks (end

Experiment B (CC-CV discharge and recording the voltage) was conducted

❍ simulated (on the windows side in the lab)
❍ hands-on (using the real battery cell and the test system, inner side of the lab)

Experiment C (determine the internal resistance of a battery cell) was conducted

❍ simulated (on the windows side in the lab)
❍ hands-on (using the real battery cell and the test system, inner side of the lab)

Figure 4.9: In short English study runs of the first phase, the participants had to state the mode they conducted the experiments in.

Keyword (ID)	Please do not write anything in this box RES18-

USE CAPITAL LETTERS here

1-2: first and second letter of your mother´s surname, in upper case letters
3-4: birthday of your mother, two digits number
5-6: the two first letters of your birthplaces name, in upper case letters
7-8: month of birth of your mother, two digits number
To complete your ID: addition of your choice in lower case letters

Example: Your mother´s surname is **Wi**lliams and she was born in 1963 at the **12th** of **March**. Your name is Miller and you´re born in **Ch**elmsford. You are a great fan of **luckyluke** and that is why you can remember it till the end of the training. In that case your ID would be: **WI12CH03luckyluke**. As mentioned this is an example. SC12MA03asterix is not your ID, you have to build your own ID.

The Experiments were conducted

❍ monday
❍ tuesday

Figure 4.10: In short English study runs of the second phase, the participants were attributed to the actually employed learning mode by code-words or the identified subgroup (day/time) they visited the labs.

of regarding program, conducted by the host university), covering all modes. The employed learning mode was either directly asked for on the test (see Figure 4.9), or could be attributed by asking which subgroup the student came from (see Figure 4.10).

In the short study runs, not all data sources were applicable (e.g. VET) or could be exploited as student time was limited (e.g. APE) (see section 5.1). Only two content areas (B* and C*) were conducted (single cross-over).

In some of the short runs, the program required no marking (attendance meant passing the module), and the tests could be written anonymously (using code-words instead of clear names), as in the German runs.

In cases where marking the tests was necessary, the laboratory-related test was combined with the tests for the associated theoretical reading. To guarantee full anonymity regarding the researcher, two title pages (one for the lecturer of the theoretical reading, including the clear student name; one for research including the students' self-created code-word) were employed. A third person removed the title page with clear name before the laboratory was evaluated. Later, it was reattached using a specific order number (see Figure 4.10, right side). This ensured anonymity regarding the research at all time.

R3, R7: International summer school, THI, 2017 & 2018

The Technische Hochschule Ingolstadt traditionally conducts a summer school. It offered an excellent chance to collect more data and include international students. The summer school lasts three weeks: the students have one week German language training, followed by two weeks of laboratory tours, lectures, and company visits. Starting from 2017, a full day on battery teaching in the school was requested, including a laboratory.

Students from Argentina, Brazil, Colombia, Czech Republic, Egypt, Hungary, India, Iran, Mexico, Poland, Spain, Turkey, and the USA participated in the summer school, 35 of whom had a technical background. After a theoretical reading, a two content area laboratory was conducted.

As students arrive immediately before the labs, they can be considered "international" even when staying in Germany for the labs. Nevertheless, the runs might be biased as only international students willing to travel to Germany for a summer school participated.

Regarding the technical background, the summer schools delivered very heterogeneous groups: The students were from Automotive, Mechanical, Mechatronics, Electrical, Aerospace, Automation, and Electric Mobility engineering programs. Some had already finished 12 semesters of study, and some came after the first semester of study.

Good feedback from R3 (Figure 4.11) meant the laboratory could be repeated in the next summer (R4) and also delivered data for the second research phase.

Figure 4.11: R3, R7: International Summer School, THI, 2017-2018, © Press department THI

R4: IEEE guest laboratory trial, TU Chemnitz, European summer break 2017

The chair of sensor and measurement technology invited the laboratory to the University of Chemnitz. The guest laboratory trial was partially financed by the IEEE Student Branch of Chemnitz.

A group of 18 internationally mixed students with a small share of German students participated. Unfortunately – as the knowledge tests were written two weeks after the laboratories, during university holidays – only 50% of the participants attended the test.

Due to very positive feedback regarding that run, the theoretical reading and laboratory sessions developed for the study were included in the "Advanced School on Impedance Spectroscopy" in the following years. Unfortunately, that framework did not allow for gaining data for educational research.

Figure 4.12: R4: Guest laboratory trial, TU Chemnitz, European summer break 2017

R5, R9: Introductory laboratory of the study program renewable energy systems (M. Sc.), THI, 2017 & 2018

The international (English language) master program "Renewable Energy Systems (M. Sc.)" was introduced at the mechanical faculty of THI in 2017. The first run was considerably small (only 18 students), while the second run already included 45 students. The study program contributed to both phases well. The participants can be considered international, as they arrived shortly before the laboratories were performed. The groups were very heterogeneous. The introductory laboratory was one of the first lectures the participants attended after arrival. After good student-feedback regarding the battery experiments in the introductory laboratory in 2017 (R5), the academic advisor of the study program asked for repeating the battery laboratory sessions with the next cohort (and inclusion of a small lecture).

4.9 Research ethics

The presented educational research targeted the improvement of laboratory teaching with the aim that students could profit off more efficient and effective teaching methods.

However, it is possible for the research itself to have adverse effects on the participants. These negative effects have to be justified before conducting a study on human beings/behaviour.

Ethical research demands that methods are always selected carefully under the aspect of research ethics in order to minimise adverse effects on participants.

This is a particular challenge for research in educational fields. It is essential to strike a balance between student learning and research goals. The generally positive feedback of participants supports the view that this was achieved in the present study.

The methods used in this research were approved by the Faculty of Electrical Engineering and Computer Science of TH Ingolstadt and the College Human Ethics Advisory Network of RMIT, Melbourne (HREC/CHEAN approval number ASEHAPP 18-16). The letters of approval can be found in the appendix on page 242.

However, there are some specific ethical concerns to be addressed:

Generally, students depend on their lecturer. That special relationship must be taken into account, especially if the research is connected to the grading of students or any compensation for participation.

Teaching/learning time had to be used instead to fill out the tests on knowledge.

For the first phase, the main concern largely stems from the fact that participants are also students, and the research question suggests that a particular student group might be disadvantaged by the study format.

For the second phase, it was only possible to research perception by letting the students be unaware of the compared modes, especially their employed laboratory mode.

Countermeasures to reduce the possible negative impact of the study on participants

- Using a crossover for all participant with an equal assignment to both modes minimised the possible impact on individual participants (Assuming an improvement of learning with hands-on training, it would have been inequitable for one group to only do hands-on experiments with the other group being forced to only simulate.)

- Using a crossover with the FlipFlop pattern for the compared modes reduced the potential adverse effects on individual participants even more.

- Participants of all runs (including research phase two) were informed about the fact that they contributed to the research. The research goals were explained to all participants.

- Participants were always free to not participate in the research by simply not returning information or not stating code words. Real names were not recorded.

- It was never possible to detect which students did not participate in the research.

- After finding that simulations were disadvantageous for former VET participants in the first research phase, for the second research phase, the equal performing mode (VET/non-VET) hands-on was selected. Specifically it was decided that *hidden simulations*, not *hidden hands-on* experiments, were to be used..

- For the second phase, the participants were told they are part of a "control group".

- Opting out had no negative consequences; and participation in the study was not incentivised, monetary or otherwise. The only exception was a standard questionnaire, participation in which the study program leader encouraged by offering a small grade-point incentive. This questionnaire was *not* devised for the study, nor was it unique to the courses in which the study was conducted. It was a general laboratory feedback form, which was *additionally* evaluated for the present study.

- In the German full semester runs (R1, R2, R6, R9) and the IEEE guest laboratory (R4), test results did not influence participants' grades for the course. In the other runs, the test questions contributed approximately 20% to the overall grade for the class. Otherwise, the conduction of these study runs would have not been possible, as teaching time was limited in these programs and it was mandatory for all subjects to contribute to the overall grade. Still, measures were also taken to avoid using real names in those runs.

- Individual behaviour or results were never made public in front of a group or on paper. Only overall trends were studied.

Chapter 5

Evidence and findings

This chapter presents the analysis of the collected data, following the methodology presented in the last chapter.

The outcomes of the different data sources are linked and discussed in chapter 6.

5.1 Available data sources vs. study runs

During the course of a study of this scope, hypotheses change or are adjusted based on results derived. Accordingly, the types of data sources considered changed/enhanced while conducting the study. Also, some of the data sources did not fit to the circumstances under which a run was conducted (e.g. limited time frame for international runs). Table 5.1 presents an overview of the available data sources versus the study runs.

Table 5.1: Data sources: availability vs. study runs

		DS-A	DS-B		DS-C	DS-D
Run		Tests (CA)	APE	VET	Personality	Moodle FB
R1	Electric Mobility (B. Eng.)	4	×	·	·	× [a]
R2	Electric Mobility (B. Eng.)	4	×	×	·	× [a]
R3	International Summer School	2	· [b]	☐	·	☐
R4	IEEE guest laboratory	2	· [b]	☐	·	☐
R5	Renew. Energy Systems (M. Sc.)	2	×	☐	·	☐
R6	Electric Mobility (B. Eng.)	4 + A	×	×	× [c]	×
R7	International Summer School	2	· [b]	☐	· [b]	☐
R8	Renew. Energy Systems (M. Sc.)	2	×	☐	· [b]	☐
R9	Electric Mobility (B. Eng.)	4 + A	×	×	×	×

Note. CA = Number of content areas.
× = data available. ☐ = not applicable. · = not employed / no data.
[a] = no correlation to test results due to missing code-words.
[b] = not employed due to time constraints/organisational reasons.
[c] = very low response rate.
DS-E and DS-F were collected without correlation to specific study runs.

Table 5.2 presents the available number of data sets correlating the different data sources.

Table 5.2: Possible correlations, number of data sets

	DS-A ho-sim	DS-A ho-hid. sim.	DS-B APE	VET	DS-C Personality	DS-D Moodle FB
DS-A ho vs. sim.	122					
DS-A ho vs. hid. sim.	☐	109				
DS-B APE	71 R1 R2 R5 p. 108	78 R6 R8 R9 p. 111	180			
DS-B VET	25 R2 p. 85	54 R6 R9 p. 94	86 R2 R6 R9 p. 105	87		
DS-C RIASEC Personality	·[d]	24 R6 R9 p. 128	28 R6 R9 p. 128	28 R6 R9 p. 124	29	
DS-D Moodle FB	·[a]	42 (112 fbe) R6 R9 p. 139	42 R6 R9 p. 140	119 fbe R6 R9 p. 140	18 R6 R9 p. 141	42
DS-E Subj. opinion on learning modes	4[e] R2[c] R3[c] p. 157	11 R8 p. 157	16 R8 R2[c] R6[c] p. 160	205 [f] p. 155	1 R6[c]	3 R6[c]
DS-F University data	☐	☐	☐	3745 [f] p. 166	☐	☐

Note. ☐ = not applicable. · = not employed / no data
fbe = feedback on single experiment
[a] = no correlation due to missing code-words
[c] = very low response rate
[d] = DS-C research started with the second phase
[e] = DS-E research started while R8, but many students were invited to participate independent of the main study
[f] = VET data for this analysis was collected independent of DS-B

For this table individual items of DS were grouped. Slight deviations occur on single items.

5.2 Comparing student laboratory learning using different learning modes (DS-A)

In this section, the data and results of the knowledge tests (DS-A) are discussed. A summary of the findings is presented on page 98. The methodology used was outlined in section 4.2.

Test difficulty level

The percentage of reached points (relative to the maximum achievable points in the test) was calculated for each test response.

Throughout all study runs, participants' test results indicate reasonable knowledge retention with mean test scores of the content areas of all runs ranging from 36% to 74% in the first phase and from 47% to 70% in the second phase. Individual scores covered the complete range of 0% to 100%, both for hands-on and simulations. Overall, students met the expectations of teachers; the difficulty of the tests was deemed appropriate to the students' knowledge. Details on the difficulties of individual tasks and content areas are presented in the appendix on page 356.

5.2.1 First research phase (R1–R5): hands-on experiments vs. simulated experiments

In total, 129 students participated in the first phase.

5.2.1.1 Results of individual runs

In the German B. Eng. study run R1 in 2016, a weak effect indicating better learning in the hands-on mode was discovered in three content areas. At the same time, no difference in test results between the two modes was discovered in the fourth content area. Altogether, students' test scores after traditional laboratories were a little better than those after simulated experiments. For R1, data needed further evaluation as no normal distribution was detected regarding the student performance after simulated experiments. A Mann-Whitney U-test showed a trend towards better learning with hands-on experiments ($p = .096$). Table 5.4 shows that the detected difference (U-test) was not caused by the overall shape of the data, as the median of both groups is different.

During the second experimental run with German B. Eng. students in 2017 (R2), an effect favouring the hands-on mode was discovered for two content areas. The third area resulted in a weak effect in the same direction, and the fourth showed no difference between the modes. Overall, students' test performances after hands-on laboratories were statistically significantly above their performances after simulated exercises ($t(146) = 2.01$, $p = .048$, Cohen's $d = .37$).

During the international summer school of 2017 (R3), one of two taught content areas demonstrated a medium effect towards better learning in hands-on mode, the other one was showing no effect. Although test results after hands-on exercises exceeded those after simulations, the difference was not statistically significant (t(146) = .65, p = .521, Cohen's d = .17).

The overall test results of the laboratory trial at Chemnitz (R4) did not show a statistical significance (p = .943). Content area B resulted in better learning with simulations, while in content area C the hands-on mode was more successful. These tests results suggested that, in spite of the random formation of groups, one of them consisted of better-performing students.

In the master's study program's preliminary laboratory (R5), the result was also not statistically significant (p = .822). This time, content area B was showing better results with hands-on, while in content area C simulations were more successful. Again, one randomly created student group outperformed the other.

5.2.1.2 Analysis grouped by test

In the following paragraphs, the results are weighted per test (i.e. students from double-crossover-experiments, taking four tests instead of two, count twice).

Although in each study run with the exception of R4 (Cohen's d = -.04 < 0) students' performed better on tests after hands-on exercises than after simulations, the differences in most of the individual study runs were not statistically significant (see Table 5.3).

Table 5.3: Comparison of experimental conditions in the first phase: test performance

	hands-on			simulated								Mann-Whitney		
Study run	Tests	M	SD	Tests	M	SD	df	t		p	d	Z		p
R1 B. Eng.	73	.12	.96	75	^{s}*-.11*	*1.01*						1.66	*	.096
R2 B. Eng.	58	.18	.83	57	-.18	1.10	113	2.01	*	.048	.37	1.96	*	.050
R3 SS	28	.09	.94	28	-.09	1.05	54	.65		.521	.17	.62		.538
R4 IEEE	9	-.02	.97	9	.02	1.03	16	-.07		.943	-.04	-.09		.929
R5 M. Sc.	18	.04	.98	18	-.04	1.02	34	.23		.822	.08	.14		.887
R1,2 Ger.	131	.15	.90	132	-.14	1.05	261	2.40	*	.018	.29	2.53	*	.012
R3,4,5 int.	55	.05	.94	55	-.05	1.02	108	.57		.570	.11	.57		.566
All	186	.12	.91	187	-.12	1.04	371	2.33	*	.021	.24	2.45	*	.014

Note. * = p < .05. d = Cohen's d (pos. = adv. of hands-on).
The last row shows the results weighted per test (i.e. students from double-crossover-experiments, taking four tests instead of two, count twice).
^{s} = Shapiro-Wilk p < .05, no normal distribution

Table 5.4: Descriptive Statistics R1: Comparison of experimental conditions

	hands-on		simulated	
	Statistic	SE	Statistic	SE
Median	.02		-.38	
Interquartile Range	1.24		1.44	
Skewness	.16	.28	.54	.28
Kurtosis	-.40	.56	-.31	.55

In the German runs, mean student performance after hands-on laboratories statistically significantly exceeded mean student performance after simulation (R1 and R2, t(261) = 2.40, p = .018, Cohen's d = .29). In international runs (R3 to R5), no significant effect was found, even though means and effect size point towards the same direction.

Combining all five runs of the first phase of the study, a total of 129 students, the 186 returned knowledge tests after hands-on laboratories outscored those taken after simulations. The overall results showed statistically significant differences in learning, favouring hands-on experiments (R1-R5, t(371) = 2.33, p = .021, Cohen's d = .24).

VET

Starting with R2, information on whether or not participants completed a German Vocational Education and Training program (VET) before studies was collected. To evaluate the relationship between the completion of a VET before studies and the test performance of the learning modes separately, independent-samples t-tests were computed based on that information.

Data in Table 5.5 indicates that students who had not completed a VET program achieved similar results in both conditions. By contrast, students who finished a VET program performed statistically significantly weaker after simulated experiments the effect size indicating a medium to large effect (R2 with VET; t(58) = 2.38, p = .021, Cohen's d = .59).

Table 5.5: Comparison of experimental conditions and VET in the first phase: test performance (R2 only, as no info about VET was recorded for R1)

	hands-on			simulated								Mann-Whitney		
VET	Tests	M	SD	Tests	M	SD	df	t		p	d	Z		p
No	19	.23	.80	18	.41	.87	35	-.65		.520	-.22	-.49		.626
Yes	30	.23	.84	30	-.38	1.11	58	2.38	*	.021	.59	2.13	*	.033
n/a	9	-.07	.94	9	-.71	1.04	16	1.37		.190	.63	1.41		.157
All	58	.18	.83	57	-.18	1.10	113	2.01	*	.048	.37	1.96	*	.050

Note. * = p < .05. d = Cohen's d (pos. = adv. of hands-on).

Data in Table 5.6 indicates that after hands-on laboratories, students with and without vocational training achieved practically identical test performances. However, after simulated laboratories test performances of students that had not completed a VET program statistically significantly exceeded the scores of participants with a VET degree (R2; t(46) = 2.74, p = .009, Cohen's d = .72) – indicating a large effect.

The comparatively low performance of former VET participants after simulated experiments leads to a lower performance of former VET students overall. The VET-data based on both modes was not normally distributed. A general difference between VET and non-VET was somewhat visible, but not statistically supported by the U-test (Z = 1.48, p = .138).

Table 5.6: Comparison of VET and experimental conditions in the first phase: test performance (R2 only, as no info about VET was recorded for R1)

	non-VET			VET								Mann-Whitney		
Mode	Tests	M	SD	Tests	M	SD	df	t		p	d	Z		p
hands-on	19	.23	.80	30	.23	.84	47	.02		.988	.00	.27		.789
simulated	18	.41	.87	30	-.38	1.11	46	2.74	**	.009	.72	2.28	*	.023
both modes	37	.32	.83	60	^{s}*-.08*	*1.03*						1.48		.138

Note. * = p < .05. ** = p < .01. d = Cohen's d (pos. = adv. of non-VET).

^{s} = Shapiro-Wilk p < .05, no normal distribution

5.2.1.3 Analysis grouped by participant

The above results are based on individual knowledge test results. In the following the data were analysed grouped by participant. Thus every participant contributes equally to the statistics.

Individual modes

Hands-on

The VET and non-VET students' average test performances after *hands-on experiments* were compared using an independent sample t-test. No statistically significant differences between students with and without former VET education were found (R2; Non-VET: N = 10, M = .16, SD = .80; VET: N = 16, M = .20, SD = .66; t(24) = -.125, p = .902).

Figure 5.1 shows the correlation between the average test performance of a student and his/her performance after *hands-on* experiments.

Simulations

The VET and non-VET students' average test performances after *simulations* were compared using a independent sample t-test. Participants without VET performed statistically significantly better after simulated experiments compared to participants with completed VET (R2; Non-VET: N = 9, M = .41, SD = .62; VET: N = 16, M = -.40, SD = .88; t(23) = 2.42, p = .024, Cohen's d = 1.01).

Figure 5.2 shows the correlation between the average test performance of a student and his/her performance after *simulated* experiments.

Hands-on vs. simulations

A paired-samples t-test was conducted to compare the average test performances of participants after hands-on experiments with those after the simulation condition. A strong and significant difference in the former VET participants' performances after both conditions was found (R2, VET; hands-on: N = 16, M = .20, SD = .66; simulation: N = 16, M = -.40, SD = .88; t(15) = 3.08, p = .008, Cohen's d = .74), while no significant differences in the test performances of participants without VET between both conditions were identified (R2, non-VET; hands-on: N = 9, M = .30, SD = .71; simulation: N = 9, M = .41, SD = .62; t(8) = -.529, p = .611, Cohen's d = -.16). These

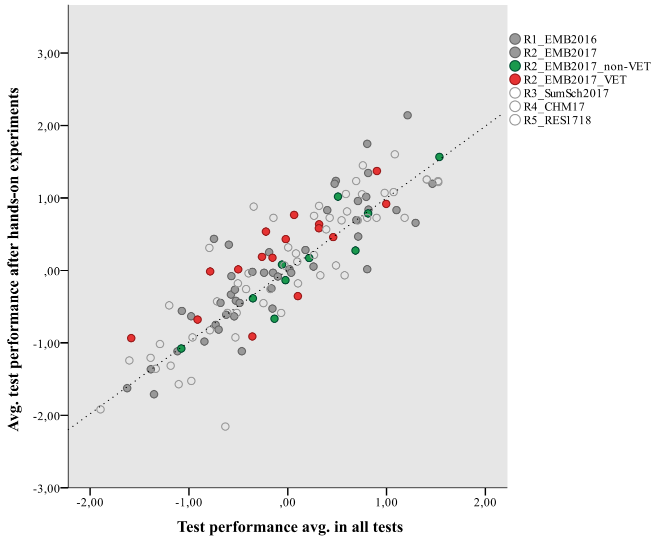

Figure 5.1: Phase 1, all runs: Average test performance of a student vs. his/her test performance in *hands-on* experiments, a circle above the line indicates above-average performance in tests after *hands-on* experiments.

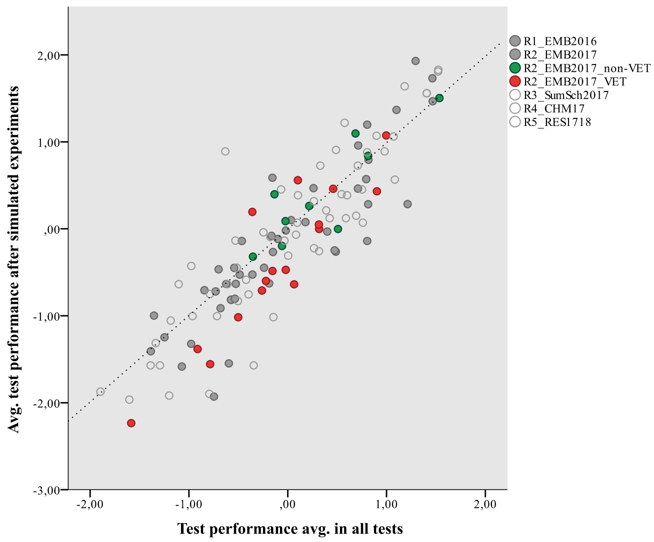

Figure 5.2: Phase 1, all runs: Average test performance of a student vs. his/her test performance in *simulated* experiments, a circle above the line indicates above-average performance in tests after *simulated* experiments.

results are coherent with the results of the results based on individual tests presented in Table 5.5.

Former VET participants perform statistically significantly lower after simulated experiments (compared to hands-on experiments and compared to their non-VET colleagues), while for non-VET no difference between modes was detected.

Superior learning mode

59% of all participants in the first phase of the research showed better learning after conducting the hands-on experiments (superior learning mode > 0). To validate better learning with hands-on experiments in the first phase, a one-sample t-test for a difference from the neutral value 0 was performed. A significant difference was found (R1 to R5: N = 122, M = .23, SD = .87; t(121) = 2.01, p = .004). Participants were on average more successful in the knowledge tests after hands-on experiments than after simulated experiments. Over all study participants, the results suggest that real hands-on experiments lead to better learning. As the significant effect observed only applies to a distinct subset of the participants (German VET), that general outcome derives from the composition of study runs and study participants and cannot be used for generalisation.

Figure 5.3 shows a histogram plot of the distribution of the superior learning mode values of all participants of the first phase. The graph illustrates the results described above well.

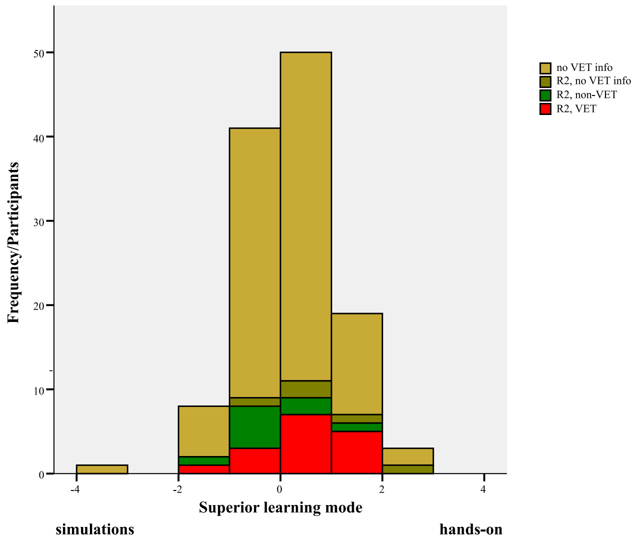

Figure 5.3: First research phase: Histogram of the distribution of the *superior learning mode* of all participants. Significant trends towards better learning with hands-on experiments were found both overall and for the students who had completed a VET.

Superior learning mode – VET

Looking to Figure 5.3, one can conclude that former VET participants tend to learn better with the hands-on approach (R2, VET: N = 16, $^{s}M = .59, SD = .77$). This aspect could not be verified using a one-sample t-test to validate the significance of the trend, as the superior learning mode data of that group was – based on a Shapiro-Wilk test (p = .042) – not normally distributed. Based on the median and percentiles in Table 5.7, it was finally acknowledged that former VET participants tended towards better learning after performing experiments hands-on than after simulations (Mdn = .80).

The somewhat visible trend in the histogram towards better learning with *simulation* of *non*-VET students could not be supported by statistical analysis (R2, non-VET: N = 9, M = -.11, SD = .60; t(8) = -.529, p = .611). This might be due to the meagre number of non-VET participants in the second run.

As VET data were not normally distributed, a non-parametric Mann-Whitney U-test was performed to validate the difference between VET and non-VET regarding the superior learning mode of the participants. The U-test showed a statistically significant difference between both groups (Z = -2.20, p = .027). Using data presented in Table 5.7 and excluding differing skewness and kurtosis as reasons for significance, it was concluded that a different median of both groups was the cause of the significant difference between them.

Table 5.7: Descriptive Statistics R2: Superior learning mode, comparison of VET and non-VET

	non-VET		VET	
	Statistic	SE	Statistic	SE
25th percentile	-.52		.13	
Median	-.07		.80	
75th percentile	.17		1.11	
Interquartile Range	.69		.98	
Skewness	.19	.72	-1.14	.56
Kurtosis	1.04	1.40	.66	1.09

Former VET participants' trend towards weaker learning with simulations (when compared to hands-on experiments) is statistically significantly different to the trend of participants without a VET.

Taking all participants of all runs of the first phase into account, no significant correlation between an individual's *superior learning mode* and the individual's overall results of the knowledge tests in the laboratory could be found (see Figure 5.4).

Figure 5.4 presents a scatter plot of a participant's overall test performance vs. his/her *superior learning mode*. The VET and non-VET participants accumulate in specific shapes:

- Students without former VET education belong to the better participants, they show no clear trend for a specific mode.
- Students with a completed VET, both those performing well and those performing badly, learn better using the hands-on condition. In the VET group, a further trend is visible: students who generally perform worse in the knowledge

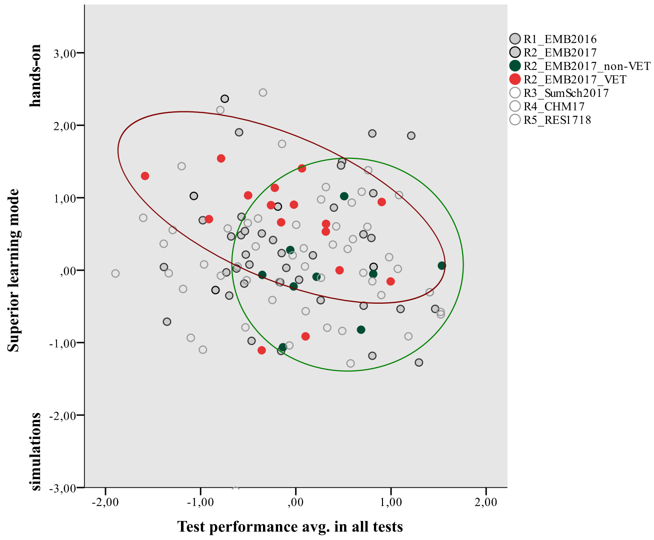

Figure 5.4: Phase 1, all runs: Average test performance of a student vs. their superior learning mode; German Students of R1 and R2 in darker grey, R2 including information about VET

> tests show a clear trend towards better learning with hands-on experiments, while those who performed above average show a less significant trend.

5.2.1.4 Further validation of tests and tasks

To exclude that specific test and task-related underlying trends (e.g. task-type related bias) influenced test results, and lead to overall wrong conclusions, further statistical tests were performed focused on the first phase.

It was found that broadly distributed data (different item formats, learning areas, tasks) contributed to the overall trends.

Influence of individual tasks / difficulty

Table H.5 on page 357 in the appendix presents the difficulties of all tasks employed in the first research phase, R1 and R2.

A Mann-Whitney U-test for differences in the learning modes of the first research phase was calculated for each task separately. Out of the 42 different questions used, 28 showed trends (based on a comparison looking to the difficulties) favouring the hands-on mode, while 14 items pointed in the direction of simulations. Only one of the 14 items pointing to better learning with simulations showed a statistically significant trend ($p < .05$), while five of the 28 items whose trends pointed towards better learning with hands-on experiments were statistically significant.

Influence of the item formats

An analysis was performed to check if the frequency of the chosen task/item format (see page 341 in the appendix) influenced the results regarding the learning modes and vice versa.

A Mann-Whitney U-test was conducted based on the first phase data of R1 and R2 to compare results grouped by the different item formats. The results are presented in Table 5.8. The scores of “Draw/Sketch to explain” and Multiple-choice tasks differed statistically significantly between both compared modes or showed trends for VET. Multiple-choice, and Single-choice tasks were most frequently employed. Nevertheless, all formats which showed statistical significant differences contributed to the general outcome “disadvantages of the simulations condition with VET”.

Influence of the learning objectives

The employed tasks were grouped into four main learning objectives (for a detailed description of the tasks please refer to the appendix, page 339ff).

It was analysed if the different learning objectives correlated with learning modes. Table 5.9 presents the results of a Mann-Whitney U-test comparing the test performance in individual test items of both conditions grouped by main learning objectives.

For non-VET students, no significant differences were identified. For VET students scores in “Battery Behaviour” did not differ significantly between both modes, whereas the scores in “Battery System Design”, “Battery Parameters”, and “Experimental Setup” showed statistically significant differences/trends with medium effects. VET students’ scores were lower after the simulations condition in all learning objectives. Data without available VET-info represents the average results of non-VET and VET.

In sum, this analysis supports the general outcomes of research phase one. Split down to learning objectives, the trend towards better learning with hands-on experiments is visible for VET and absent for non-VET participants.

Table 5.8: Comparison of experimental conditions and VET in the first phase: test performance in individual *tasks* (R1. R2), grouped for item format.

		hands-on			simulated			Mann-Whitney		
VET	IF	Tasks	M	SD	Tasks	M	SD	Z		p
n/a[a]	DG	39	^{s} *-.18*	*.75*	46	^{s} *.03*	*1.18*	-.522		.602
	DS	84	^{s} *.02*	*1.01*	86	^{s} *-.02*	*1.02*	.077		.939
	MC	229	^{s} *.05*	*1.03*	227	^{s} *-.10*	*1.03*	1.583		.113
	SC	257	^{s} *.06*	*1.00*	268	^{s} *-.06*	*1.01*	1.513		.130
	TR	84	^{s} *.07*	*1.08*	83	^{s} *-.15*	*.92*	1.114		.265
	TV	120	^{s} *.04*	*1.08*	132	^{s} *-.16*	*1.00*	1.247		.212
non-VET	DG	13	.14	.80	14	^{s} *.72*	*1.13*	-1.694	†	.090
	DS	20	-.04	.99	18	^{s} *.25*	*1.03*	-.396		.692
	MC	54	^{s} *.08*	*.95*	49	^{s} *.10*	*1.06*	-.033		.974
	SC	55	^{s} *.13*	*.91*	56	^{s} *.12*	*.87*	.39		.697
	TR	20	^{s} *.21*	*.83*	17	.17	.96	.402		.688
	TV	27	^{s} *.07*	*.92*	27	^{s} *.38*	*.74*	-.976		.329
VET	DG	23	^{s} *-.30*	*.88*	22	.03	.97	-1.37		.171
	DS	30	^{s} *.22*	*1.01*	30	^{s} *-.33*	*.88*	2.358	*	.018
	MC	81	^{s} *.23*	*.96*	84	^{s} *-.21*	*.81*	2.972	**	.003
	SC	86	^{s} *.09*	*.96*	94	^{s} *-.23*	*1.10*	1.772	†	.076
	TR	31	^{s} *.09*	*.95*	29	^{s} *-.10*	*1.16*	.835		.403
	TV	48	^{s} *.01*	*.95*	42	^{s} *.08*	*.99*	-.382		.702

Note. † = p < .10. * = p < .05. ** = p < .01.
Z pos. = adv. of hands-on.
[a] = mainly R1
^{s} = Shapiro-Wilk p < .05, no normal distribution.
DG = Draw/Graph, DS = Draw/Sketch to explain, MC = Multiple-choice, SC = Single-choice,
TR = Text/state reason or equation, TV = Text/state or calculate a value

Table 5.9: Comparison of experimental conditions and VET in the first phase: test performance in individual *tasks* (R1. R2), grouped for learning objectives.

		hands-on			simulated			Mann-Whitney		
VET	LO	Tasks	M	SD	Tasks	M	SD	Z	p	
n/a[a]	BB	353	^{s} *.02*	*1.02*	380	^{s} *-.06*	*1.02*	1.23		.221
	BP	126	^{s} *.07*	*.98*	126	^{s} *-.07*	*1.03*	1.21		.226
	BSD	84	^{s} *-.02*	*1.00*	80	^{s} *-.04*	*1.02*	.45		.650
	ES	250	^{s} *.06*	*1.04*	256	^{s} *-.14*	*1.01*	1.90	†	.057
non-VET	BB	80	^{s} *.03*	*.90*	82	^{s} *.24*	*1.03*	-1.30		.193
	BP	30	^{s} *.05*	*1.00*	24	-.04	.83	.79		.431
	BSD	20	^{s} *.02*	*.89*	20	^{s} *.49*	*1.11*	-.85		.394
	ES	59	^{s} *.26*	*.88*	55	^{s} *.20*	*.82*	.09		.930
VET	BB	137	^{s} *.01*	*.99*	133	^{s} *-.06*	*.93*	.32		.751
	BP	50	^{s} *.21*	*.89*	40	^{s} *-.29*	*1.18*	2.21	*	.027
	BSD	24	^{s} *.19*	*.82*	36	^{s} *-.25*	*1.02*	1.65	†	.098
	ES	88	^{s} *.15*	*.98*	92	^{s} *-.21*	*.97*	2.57	**	.010
n/a[a]	all	813	^{s} *.04*	*1.02*	842	^{s} *-.09*	*1.02*	2.42	*	.015
non-VET	all	189	^{s} *.10*	*.91*	181	^{s} *.22*	*.96*	-.76		.449
VET	all	299	^{s} *.10*	*.96*	301	^{s} *-.16*	*.99*	3.08	**	.002

Note. † = p < .10. * = p < .05. ** = p < .01.
Z pos. = adv. of hands-on.
[a] = mainly R1
^{s} = Shapiro-Wilk p < .05, no normal distribution.
BB = Battery Behaviour, BP = Battery Parameters,
BSD = Battery System Design, ES = Experimental Setup

5.2.2 Second research phase (R6–R9): hands-on experiments vs. hidden simulations

In total, 139 students participated in the second phase.

5.2.2.1 Results of individual runs

During the experimental run with B. Eng. students in 2018 (R6), an effect indicating better learning with hidden simulations was discovered for one content areas. The other content areas showed no significant differences between the modes. Overall, students' test performances differed with a statistically significant weak effect towards hidden simulations ($t(109) = -1.74$, $p = .085$, Cohen's $d = -.33$).

During the international summer school of 2018 (R7), neither content area showed a significant effect, and both tended in different directions. The overall difference was not statistically significant, both modes performed identically ($t(36) = .09$, $p = .930$, Cohen's $d = .03$).

In the 2018 master's study program's preliminary laboratory (R8), the result of all content areas did not show a statistically significant difference either. The overall run showed equivalent learning with both modes ($t(64) = -.13$, $p = .896$, Cohen's $d = -.03$).

In the German B. Eng. study run R9 in 2019, none of the four content areas showed any significant effect. The test results after hidden simulations were not normally distributed. A U-Test was performed and showed no significant difference ($p = .876$). Altogether, the difference between modes was not statistically significant.

5.2.2.2 Analysis grouped by test

Table 5.10 shows the comparison of students' test performances after hands-on experiments and after hidden simulations.

Table 5.10: Comparison of experimental conditions in the second phase: test performance

	hands-on			hidden simulations								Mann-Whitney		
Study run	N	M	SD	N	M	SD	df	t		p	d	Z		p
R6 B. Eng.	55	-.16	.98	56	.16	.98	109	-1.74	†	.085	-.33	-1.65	†	.098
R7 SS	19	.02	.92	19	-.02	1.08	36	.09		.930	.03	.03		.977
R8 M. Sc.	33	-.02	1.00	33	.02	1.00	64	-.13		.896	-.03	-.13		.898
R9 B. Eng.	48	.02	1.08	50	^{s}*-.02*	*.90*						.16		.876
R6,9 Ger.	103	-.08	1.02	106	.07	.94	207	-1.09		.276	-.15	-1.04		.298
R7,8 int.	52	-.01	.96	52	.01	1.02	102	-.05		.959	-.01	.00		.999
All	155	-.05	1.00	158	.05	.96	311	-.92		.357	-.10	-.96		.336

Note. † = $p < .10$. d = Cohen's d (pos. = adv. of hands-on).
The last row shows the results weighted per test (i.e. students from double-crossover-experiments, taking four tests instead of two, count twice).
^{s} = Shapiro-Wilk $p < .05$, no normal distribution

When runs were evaluated as a group, neither the German runs (R6 and R8, t(207) = -1.09, p = .276, Cohen's d = -.15) nor the international runs (R7 and R8, t(102) = -.05, p = .959, Cohen's d = -.01) showed a significant effect.

Combining all five runs, 139 students and 313 returned knowledge tests of the second phase of the study showed no inclination of one mode outperforming the other (R6-R9, t(311) = -.92, p = .357, Cohen's d = -.10).

VET

As with previous runs, outcomes were analysed separately depending on whether students had finished a VET before their studies or not (see Table 5.11). No significant differences in students' test performances were found.

Table 5.11: Comparison of experimental conditions and VET in the second phase: test performance (R6, R9)

	hands-on			hidden simulations							Mann-Whitney	
VET	Tests	M	SD	Tests	M	SD	df	t	p	d	Z	p
No	44	-.28	1.07	46	-.05	.97	88	-1.07	.287	-.23	-.93	.351
Yes	59	.08	.97	60	.17	.91	117	-.52	.605	-.10	-.59	.553
All	103	-.08	1.02	106	.07	.94	207	-1.09	.276	-.15	-1.04	.298

Note. d = Cohen's d (pos. = adv. of hands-on).

Table 5.12 shows the comparison of non-VET and VET students based on the individual test performances of R6 and R9. Overall, former VET participants statistically significantly exceeded the scores of participants without VET, indicating a small effect (R6, R9; t(207) = -2.10, p = .037, Cohen's d = .29).

Table 5.12: Comparison of VET and experimental conditions in the second phase: test performance (R6, R9)

	non-VET			VET								Mann-Whitney		
Mode	Tests	M	SD	Tests	M	SD	df	t		p	d	Z		p
hands-on	44	-.28	1.07	59	.08	.97	101	-1.78	†	.078	-.35	-1.70	†	.089
hidden sim.	46	-.05	.97	60	.17	.91	104	-1.19		.239	-.23	-1.14		.255
Both	90	-.16	1.02	119	.12	.94	207	-2.10	*	.037	-.29	-2.01	*	.045

Note. † = p < .10. * = p < .05. d = Cohen's d (pos. = adv. of non-VET).

5.2.2.3 Analysis grouped by participant

The above results are based on individual knowledge test results. In the following, the analysed data were grouped by the participant. Every participant contributes to the statistics equally. Generating participant-related statistics allows the correlation of other data sources with the data collected in this part of the research. In contrast to the analysis by test result, a missing test result in a particular mode was extrapolated from the student's results achieved in other tests using the same teaching method if available.

Individual modes

Hands-on
The VET and non-VET students' average test performances after *hands-on experiments* were compared using a independent sample t-test. While evaluation weighted by test (Table 5.12) was statistically significant, no statistically significant differences between students with and without former VET education were found weighted by students (R6, R9; Non-VET: N = 24, M = -.29, SD = .92; VET: N = 30, M = .06, SD = .73; t(52) = -1.56, p = .126).

Hidden simulations
The VET and non-VET students' average test performances after *hidden simulations* were compared using an independent sample t-test. No statistically significant differences between students with and without former VET education were found (R6, R9; Non-VET: N = 24, M = -.09, SD = .74; VET: N = 30, M = .16, SD = .70; t(52) = -1.29, p = .203).

Overall test performance
Comparing the average test performance of all individual students between VET and non-VET in the second phase, no significant difference could be found (R6, R9; Non-VET: N = 24, M = -.18, SD = .76; VET: N = 30, M = .12, SD = .57; t(52) = -1.635, p = .108), while the evaluation weighted by test showed significance (Table 5.12).

Hands-on vs. hidden simulations
A paired-samples t-test was conducted to compare average *test performances* of participants after hands-on experiments with those after the hidden simulation condition. For the German (R6, R9; hands-on: N = 54, M = -.09, SD = .83; hidden simulation: N = 54, M = .06, SD = .72; t(15) = -1.39, p = .169) as well as for the international runs (R7, R8; hands-on: N = 55, M = -.00, SD = .96; hidden simulation: N = 55, M = .02, SD = 1.02; t(54) = -.185, p = .854) student learning after hands-on and hidden simulation laboratories did not differ significantly. VET and non-VET students of the German runs were analysed separately. No significant difference in performance after both conditions were found for former VET-participants (R6, R9, VET; hands-on: N = 30, M = .06, SD = .73; hidden simulation: N = 30, M = .17, SD = .70; t(29) = -.671, p = .508), or participants without VET (R2, non-VET; hands-on: N = 24, M = -.29, SD = .93; hidden simulation: N = 24, M = -.09, SD = .74; t(23) = -1.438, p = .164).

Superior learning mode

Approximately every second participant (52%) in the second phase of the research showed better learning after the hands-on condition, the others after the hidden simulations condition.

To validate better learning with hands-on experiments in the second phase (superior learning mode > 0), a one-sample t-test for difference from the neutral value

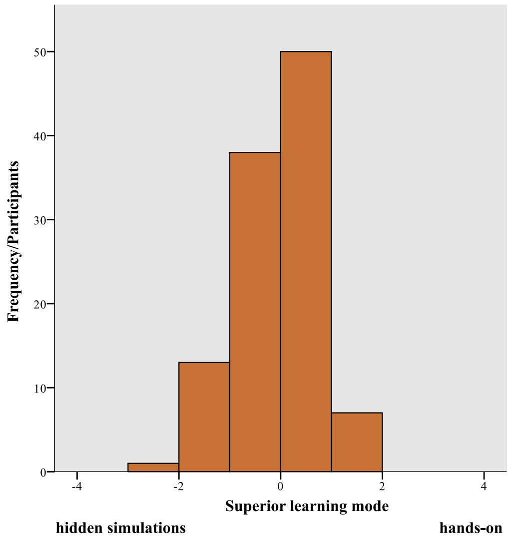

Figure 5.5: Second research phase: Histogram of the distribution of the *superior learning mode* of all participants. No significant trend towards a learning mode was found.

zero was performed. No significant difference was found (R6 up to R9: N = 109, M = -.08, SD = .80; t(108) = -1.09, p = .278). Participants showed no significant trend towards better performance after either of the compared modes.

Figure 5.5 shows a histogram plot of the distribution of the superior learning mode values of all participants of the second phase.

Superior learning mode – VET

Also in the second research phase, a t-test was calculated to compare the superior learning mode of individual students based on their former VET status. No significant difference between former VET-participants and students without VET was found (Non-VET: N = 24, M = -.20, SD = .68; VET: N = 30, M = -.10, SD = .85; t(52) = -.45, p = .655, Cohen's d = -.12), which confirms the analysis regarding the test performances in Table 5.11 and Table 5.12.

A paired-samples t-test was conducted to compare average test performances of participants after hands-on experiments with test performances after the hidden simulation condition. In both groups, no significant differences were found (VET R6, R9; hands-on: N = 30, M = .06, SD = .73; hidden simulation: N = 30, M = .17, SD = .70; t(29) = -.67, p = .508, Cohen's d = -.15 / non-VET R6, R9; hands-on: N = 24, M = -.29, SD = .93; hidden simulation: N = 24, M = -.09, SD = .74; t(23) = -1.44, p = .164, Cohen's d = -.24).

5.2.3 Bias in groups

Subgroups with outstanding students would bias the overall outcomes. Thus, the participants' background was examined. All participants of the German runs came from the same study program "Electric mobility" at UAS Ingolstadt. Objective background data was available for these students (see DS-F).

A clear individual participant data correlation was not possible as participants stayed anonymous in DS-A and DS-F, and code words were not available for the data provided by the university. Nevertheless, a rough comparison of the German study runs (R1, R2, R6, R9) was possible – based on the date of matriculation and the estimation every student proceeds the semesters in normal speed and order, data of DS-F could be used for the following analysis.

Similar data was not available for the two runs in the master's study program's preliminary laboratory (R5, R8), as it was conducted during the first semester of studies. The IEEE run in Chemnitz (R4) was not connected to any specific study program.

Equivalence of VET and non-VET participants

In order to test if either subgroup consisted of outstanding students and therefore if the cohorts / collected data were biased, average grades of VET and non-VET students were compared using the participants' average grades of the three semesters before the laboratory.

No significant differences were found between VET and non-VET students of R2 (R2; grades of semesters 1 to 3; non-VET: $N = 17$, $M = 3.08$, $SD = .88$; VET: $N = 12$, $M = 3.00$, $SD = .44$; $t(27) = .29$, $p = .773$, Cohen's $d = .11$), as both subgroups generally performed similarly in their studies. This indicates that the noticeable difference between hands-on experiments and simulation was not caused by a bias in student competencies.

The same analysis was performed for both German runs, comparing the hidden simulation condition with hands-on experiments (R6, R9; grades of semesters 1 to 3; non-VET: $N = 24$, $M = 3.05$, $SD = .63$; VET: $N = 28$, $M = 2.44$, $SD = .79$; $t(50) = 3.02$, $p = .004$, Cohen's $d = .84$; $Z = 3.11$, $p = .002$). Here, a significant bias towards better students in the VET group was detected (For Bavarian UAS, higher values indicate bad grades/low performance), which could somewhat explain the biased results of Table 5.12 towards better overall learning of VET participants in the second phase. Higher performance by VET students is not standard (see section 5.7), usually non-VET students earn equal (or slightly better) grades.

Equivalence of participants of the first and second phase

The average marks of the first three semesters of the relevant cohorts were compared between both research phases to evaluate if the second phase students could be seen as equivalent to the students of the first phase. No significant differences for non-VET students were identified ($p > .9$), while for former VET participants, the marks of the second phase were statistically significantly better than in the first phase (grades of semesters 1 to 3; R1, R2: $N = 34$, $M = 2.91$, $SD = .52$; R6, R9: $N = 28$, $M = 2.44$, $SD = .79$; $t(60) = 2.82$, $p = .006$; $Z = 3.14$, $p = .002$). The data was analysed further, and it was found that the participants of study run R6 contributed the most to the above difference – R6 had a lot of high-performing VET-students, while the data-sets attributed to R9 showed no statistically significant differences to R1 and R2.

Neither VET nor non-VET participants showed a statistically significant difference between the first phase and second phase regarding the following factors: age at completing the VET (if applicable), the years between VET and beginning of studies (if applicable), age at matriculation, credits earned during the first three semesters.

5.2.4 Summary of results of DS-A

Summary of section 5.2 "Laboratory learning using different learning modes (DS-A)"

In several study runs, test results related to knowledge acquisition as a result of conducting laboratory exercises in different modes were collected.

1. A counterbalanced within-subject research methodology was applied; it focused on the comparison of the laboratory modes hands-on and simulation.
2. A case study was performed, teaching battery basics and measurement methods for battery cells and energy storage systems.
   (a) Accompanying lectures, experimental instructions, teachers, learning objectives, tests, and many other variables were controlled for both groups.
   (b) Nearly identical experimental procedures were used in both modes.
3. The study was conducted in two consecutive phases.
4. Study runs of both phases were carried out in local access domain.
5. Nine study runs with German and international participants at two different universities were conducted.
6. Written anonymous knowledge tests were applied to validate student learning for individual content areas/modes.
7. In the first phase, the objective was to compare student learning through hands-on experiments with that from simulations. In this phase students were aware of the mode they use.
   (a) The analysis was performed weighted by test results and weighted by participants as well as comparing the results of both modes for individual participants (superior learning mode of an individual student). Using the three methods, the outcomes showed identical significance and trends.

(b) Correlating the average performance of the participants with their superior learning mode, no general trend was found.

(c) Presence/Absence of a former German Vocational Education and Training (VET) influenced student learning.

i. Comparing the modes

A. Overall, students with VET learned statistically significantly weaker with simulations than with hands-on experiments (Tests: d = .59, p = .021; Participants: d = .74, p = .008),

B. while students without former VET education showed no significant difference between both modes.

C. VET and non-VET participants' trend (superior learning mode) shows a significant difference (Mann-Whitney U-test, Z = -2.20, p = .027).

ii. Analysing the two modes separately,

A. after the hands-on condition, students with VET before studies and those without performed similarly, while

B. participants, who had completed a VET program prior to university enrolment, statistically significantly underperformed compared to their peers after simulated experiments (Tests: d = .72, p = .009; Participants: d = 1.01, p = .024).

iii. Analysing the general performance of the participants compared to their superior learning mode for VET and non-VET separately

A. students without former VET education are among the better performing participants; they show no clear trend towards a specific mode.

B. students with completed VET range from bad performers to good performers, and perform better after learning in the hands-on condition.

iv. Participants of the German runs in the first phase had no general bias (VET vs. non-VET) in study marks until the third semester, which supports the validity of the findings.

(d) The difference regarding the success of the learning modes derives from a broad basis of test items. The analysis resulted in no hints regarding an unintended influence on the research outcomes by selecting particular test questions.

8. Phase two was conceptualised to verify the quality of simulations used in phase one and at the same time to give insight into possible subjective influences of the laboratory mode itself. In this second phase, students

were not aware of the mode they used, as simulated results were presented as hands-on experiments.

(a) The second phase was conducted to:
    i. Validate the results of the first phase,
    ii. test the theory that differences in the first phase were caused by students' perception, and
    iii. eliminate weaknesses of the simulation model used in the first mode causing disadvantages when using simulations.
(b) The analysis was performed weighted by test results and weighted by participant as well as comparing the results of both modes for individual participants (superior learning mode). All methods showed identical trends and similar statistical significance (p).
(c) Generally, no significant differences between the test results of both modes were found. Students acquired nearly identical knowledge under both conditions; the outcomes tended insignificantly towards "hidden simulations".
(d) Presence/Absence of a former German Vocational Education and Training (VET) had little influence on student learning in the second research phase.
    i. Now (only perceiving hands-on experiments) former VET students performed slightly better than students without VET (Tests: $d = -.29$, $p = .037$; Participants: $p = .108$). This outcome needs to be treated carefully as the VET participants of the second phase also had a bias towards better study marks in the first three semesters.
    ii. Regarding the *superior learning mode* of VET and non-VET participants, the second phase showed no statistically significant difference.

9. Since students using the simulation model from the first phase in the hidden simulations performed similarly well as students using real hands-on data, the quality of the simulations could be confirmed. The difference between both modes in the first phase was perhaps due to the perception of the learning mode. The observed effect in the first phase does not appear to be based on bad experimental results caused by weaknesses in the simulation.

10. The results suggest that as long as the employed simulations are perceived as "real" hands-on experiments, no differences in learning between both modes exist.

5.3 Amount of Practical Experience (DS-B)

This investigation was done using the methodology outlined in section 4.3. A summary of the findings is presented on page 120.

5.3.1 Factor analysis

The Amount of Practical Experience of 263 participants (R1, R2, R5, R6, R8, R9) was collected using a questionnaire with 17 items. In proportion to the number of items, the sample size was adequate [128]. A *Principal Component Analysis (PCA)* was performed to extract the most important independent factors. Maximum correlations between items ranged between .277 and .613, showing that all items measured the same underlying construct while at the same time no item could be entirely replaced by another (multicollinearity). Anti-Image Correlation factors were > .7 for all items, indicating that all of them were adequate samples [129]. The Kaiser-Meyer-Olkin measure of sampling adequacy was .785, representing a proper factor analysis [130]. Bartlett's test of sphericity ($\chi^2(136) = 866$, $p < .0001$) indicated that correlations between items were sufficiently large for performing a PCA. Only factors with eigenvalues ≥ 1 were considered (Kaiser-Guttman-Criteria) [131, 132].

Examination of Kaiser's criteria and the scree-plot shown in Figure 5.6 provided empirical justification for obtaining four factors with eigenvalues exceeding 1, which accounted for 55.31% of total variance. Among the factor solutions, the varimax-rotated (orthogonal) four-factor solution generated the most interpretable solution, and most items loaded highly on only one of the four factors (see Table 5.13). Factors with loadings < .4 were excluded from the analysis. The loadings of the used items were squared and plotted (Figure 5.7). The plot presents a clear separation of the four dimensions.

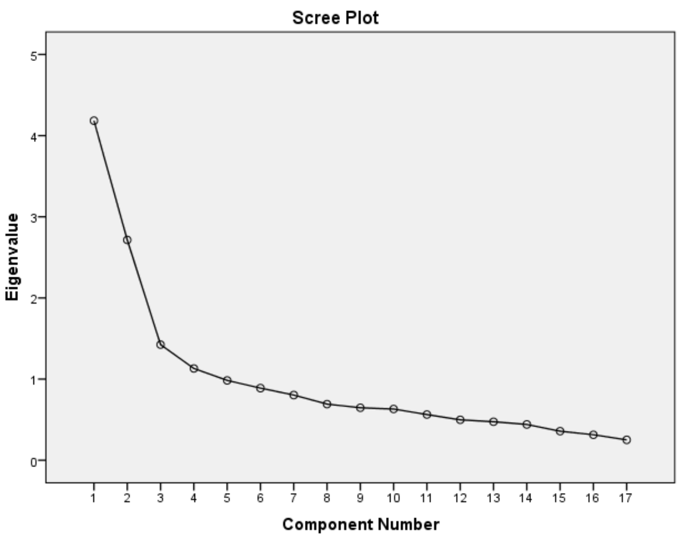

Figure 5.6: APE-Dimensions: Scree Plot

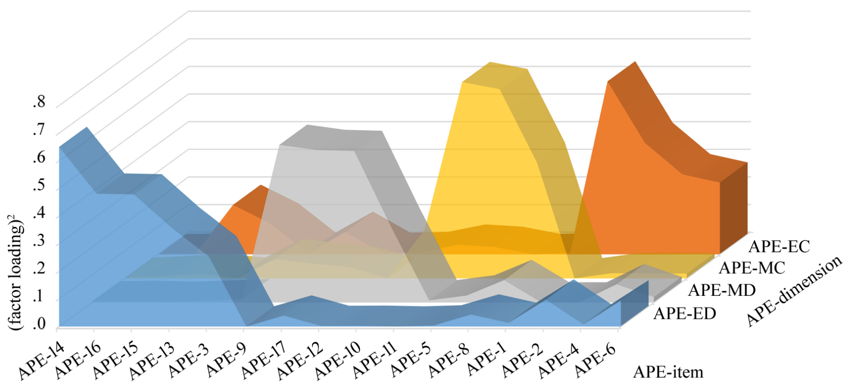

Figure 5.7: APE-Dimensions: Squared factor loadings

5.3.1.1 Dimensions

The following four dimensions were found. The underlying items (for the full list, see Table 5.13) showed clear connections in content, which made it easy to interpret and name the dimensions:

APE-ED (Electronics Do)

Practical experience in procedures which were mostly taught during studies in electronics and information technology. In contrast to the next component, these experiences are less creative in the sense of engineering or creating products. [APE-14, APE-16, APE-15, APE-13, APE-3]

APE-EC (Electronics Create)

Measures experience in designing electronic products. The items included focus on the engineering of electronic products, including the programming of firmware. [APE-1, APE-2, (APE-4), APE-6, APE-15]

APE-MC (Mechanics Car)

Practical experience in working with cars, e.g., changing motor oil, tires, or finding errors in the electrical system. [APE-9, APE-10, APE-12, APE-17]

APE-MD (Mechanics Do)

Experience in mechanical work, like cutting threads, sawing, and producing metal parts. [APE-5, APE-7, APE-8, APE-11]

APE-E, APE-M

Considering the meaning of the dimensions it was natural to group them into electronic and mechanic dimensions.

Table 5.13: APE: Rotated factor matrix showing factor loadings of the items used in the Principal Component Analysis to determine the four dimensions of practical experience. Factors were rotated orthogonally to maximise the explained variance. The rotation converged in 7 iterations, items in brackets were excluded from the factors.

Item	Content	APE-E		APE-M	
		APE-ED	APE-EC	APE-MC	APE-MD
APE-14	Simulate the frequency response of a simple RC element in SPICE.	.806			
APE-16	Explain to others what a C compiler does in principle.	.694			
APE-15	Read and explain a circuit diagram containing an operational amplifier.	.692	.423		
APE-13	Designed an analogue low pass-filter of 2nd order for a given requirement.	.599			
APE-3	Soldered electronic parts with a soldering iron.	.503			
APE-9	Searched an error in the electric system of a car, found it, and fixed it.			.755	
APE-17	Charge a lead-acid battery of a common car with a laboratory power supply.			.743	
APE-12	Changed a motor's oil.			.740	
APE-10	Changed a car's tires.			.519	(.421)
APE-7	Assembled a model kit.	(.420)		(.470)	
APE-11	Produced a metal part from a technical drawing.				.843
APE-5	Made a thread on a bore hole.				.828
APE-8	Sawed off a pipe.				.648
APE-1	Realised a function using a self-made circuit diagram.		.791		
APE-2	Transferred a circuit diagram to a PCB.		.635		
APE-4	Configured and assembled a desktop-PC.		(.538)		
APE-6	Programmed a micro controller.		.510		

5.3.2 Reliability

Cronbach's α was calculated to assess the internal consistency of the dimensions. The results and the specific number of items are presented in Table 5.14. With Cronbach's α = .797, the overall APE scale had a high reliability [129]. Internal consistencies were satisfying, with Cronbach's alphas for positive affects > .7 for all tested dimensions, except for APE-EC, which was weak (.594). After exchanging APE-4 with APE-15, the value increased to an acceptable value (.686) [133, p. 153].

Table 5.14: APE-dimensions: reliability

	N Items	Cronbach's α
APE	17	.797
APE-E	9	.752
APE-ED	5	.797
APE-EC	4	.686
APE-M	7	.786
APE-MC	4	.744
APE-MD	3	.745

5.3.3 Test for normal distribution

A Shapiro-Wilk test was performed to test the APE main dimension for normal distribution. The test confirmed normally distributed data for all groups except the German VET students in the second study phase. Here, the Shapiro-Wilk test showed a significant departure from normality ($W(31) = .97$, $p = .018$). After the histogram was analysed, a clear outlier was detected. Statistics without this outlier confirmed normal distribution of the underlying construct (see Table 5.15).

Table 5.15: APE: Test for normal distribution. Shapiro-Wilk statistics.

			W	df		p
hands-on vs. simulation	R2, German	non-VET	.92	11		.306
First Phase		VET	.92	16		.187
	R5, International		.98	40		.714
hands-on vs. hidden sim.	R6, R9, German	non-VET	.97	24		.770
Second Phase		VET	.92	31	*	.018
		VET[1]	.97	30		.437
	R8, International		.98	17		.935

Note. * = p < .05.
In the second research phase, an outlier caused significance of non-normal distribution (VET).
After removal of the outlier, normal distribution was achieved (VET[1]).

5.3.4 Mean Amount of Practical Experience (APE) / difficulties

Mean and median values are presented in Table 5.16. It is apparent that the average value of the overall dimension APE was positive (M = .35, Mdn = .38). In the study, participants more often stated "yes" or "rather yes" than "no" and "rather no". The following subsections compare the groups for the apparent differences between VET and non-VET, and German and international.

Table 5.16: APE: Difficulties of dimensions

	German UAS				International		all	
	VET		non-VET					
Students	47		35		55		181	
Item	Mean	Mdn	Mean	Mdn	Mean	Mdn	Mean	Mdn
APE	.69	.74	.23	.32	.05	.03	.35	.38
APE-E	.29	.39	.09	.17	.00	-.06	.13	.17
APE-ED	.50	.50	.40	.50	-.07	-.10	.27	.30
APE-EC	.07	.00	-.24	-.25	.06	.00	-.02	.00
APE-M	1.14	1.38	.40	.50	.10	.13	.60	.75
APE-MC	1.00	1.50	.26	.50	.07	.00	.46	.75
APE-MD	1.39	1.50	.43	.50	.11	.17	.73	1.17

Difficulties of individual items
To aid other researchers wishing to use the same or similar items in their research, the difficulties of all single items are reported in in the appendix in Table D.1, separated for German VET, non-VET and for international participants.

5.3.5 APE of German and international participants

For the international participants, the average APE value was neutral (M = .05, Mdn = .03), while the German participants overall had higher values (M = .50, Mdn = .50, not presented in table).

A Mann-Whitney U-test was performed to validate the general difference between Germans and internationals regarding the APE. The U-test showed a statistically significant difference between both groups ($Z = -5.225$, $p < .001$). German participants more often stated "yes" or "rather yes" than international participants. Further statistical tests comparing these two groups were not performed as the German subgroups (VET, non-VET) differed too much (see next section and Table 5.16).

5.3.6 APE of former VET participants and non-VET students

The Amount of Practical Experience was evaluated based on whether participants had previously completed a VET education in the German runs. APE was collected for R1, R2, R6, and R9. VET information was collected starting from R2.

To evaluate the relationship between the completion of a VET before studies and the self-reported Amount of Practical Experience, an independent-samples t-tests was computed. The result is presented in the first line of Table 5.17. Students who completed a VET reported statistically significantly higher Amounts of Practical Experience than their peers who had not. The effect size suggests a strong connection between VET and practical experience (t(85) = -4.32, p = .001, Cohen's d = -.93).

An independent samples t-test was computed to determine whether the Amount of Practical Experience in each of the six subdimensions depended on the participant's VET status. A Mann-Whitney U-test covered groups which were not normally distributed. The results are shown in Table 5.17. In general, VET participants reported statistically significantly higher amounts of practical experience than their non-VET peers.

The effect was stronger with experience in the mechanical domain APE-M than in the electrical domain APE-E.

The difference in mechanical experience can be explained easily, as 70% of the VET students in the Bachelor runs held VET degrees in mechanical or mechatronical professions (Table 5.18), which may have caused the significant difference compared to students without VET. With the exception of one subject ("theoretical physics"), mechanical topics are not taught in the electrical study programs. This made it unlikely for the students to gain more experience in these topics during their studies, which explains the clear differences.

The results of the Mann-Whitney test were caused by a difference in medians, as well as by a strong difference in kurtosis. Many former VET-participants stated the maximum value 1.5.

Table 5.17: APE-dimensions: t-test and U-test regarding the completion of a VET education before studies

	non-VET			VET								Mann-Whitney		
	N	M	SD	N	M	SD	df	t		p	d	Z		p
APE	39	.22	.58	48	.68	.41	85	-4.32	**	.001	-.93	-4.08	**	.000
APE-E	39	.03	.65	48	.27	.55	85	-1.87	†	.066	-.40	-1.83	†	.067
APE-ED	39	.31	.66	48	.47	.62	85	-1.18		.243	-.25	-1.13		.257
APE-EC	39	-.25	.80	48	.05	.70	85	-1.88	†	.064	-.40	-1.84	†	.066
APE-M	39	.43	.74	48	^{s}*1.14*	*.47*						-5.13	**	.000
APE-MC	39	.26	.87	48	^{s}*1.00*	*.75*						-4.27	**	.000
APE-MD	39	^{s}*.52*	*.97*	48	^{s}*1.39*	*.25*						-5.33	**	.000

Note. † = p < .10. * = p < .05. ** = p < .01. d = Cohen's d (pos. = more experience of non-VET).
^{s} = Shapiro-Wilk p < .05, no normal distribution

In the electrical domain, the differences were not as clear. Here, only the APE-EC "electronics create" dimension showed a trend towards former VET students having more practical experience, while the APE-ED "electronics do" dimension showed no significant differences. This was surprising, as more than 80% of former VET-participants had completed an apprenticeship dealing at least in part with electronic components (Table 5.18).

The smaller effect of APE-E compared to APE-M seems to be caused by experience gained during their University studies. Most of the topics mentioned in the items of the APE-ED dimension are covered in subjects of the first semesters of studies. This allows students without VET to catch up, leading to smaller differences between VET and non-VET.

Considering the fact that the students were enrolled in a study program focusing on electronics, means for the electronics categories are low – for both groups of students (Table 5.17). The APE-EC "electronics create" items focus (in contrast to APE-ED) on operations necessary for implementing features in real products. The more pronounced difference in APE-EC (vs. APE-ED) suggests that former VET participants are more likely to try out their gained theoretical knowledge in a practical way.

As most did a VET related to mechanics, the difference in experience was likely due to differences in behaviour *during* their studies.

VET vs. non-VET of individual items

The table with the results of a non-parametric Mann-Whitney U-Test for individual items is presented on page 249 in the appendix. More than 50% of the items show significant trends towards higher experience of former VET-participants, while no single trend was found pointing to higher experience of non-VET students. Looking to individual items did not lead to any more profound insights.

APE of VET and non-VET in both phases

The APE dimensions were compared using a Mann-Whitney U-test to check the equivalence of the participants of both phases separately for VET and non-VET participants. All APE dimensions showed no statistically significant differences between the study phases ($p > 0.1$), except for the subdimension APE-ED regarding participants without VET. The APE-ED value was smaller in the first phase for this group (Phase 1: .10, Phase 2: .53, $Z = -1.986$, $p = .047$). This specific analysis has to be treated carefully as Phase 1 data based only on study run R2 (missing VET information with R1).

5.3.7 VET professions of the participants of the German study runs

In R6 and R9, the VET students were asked for the field of their VET profession. The stated professions were coded in four categories, the results are presented in Table 5.18. Most professions could be categorised under electronics or mechatronics, with a small share falling under purely mechanical professions.

The stated professions were also analysed in a second way: 77% (N = 23) of professions were related to the subject of the study program (B. Eng. Electrical Engineering and Electric Mobility), while 23% (N = 7) showed no clear connection to vehicles or electronics. It can be inferred that the ratios were similar in R1 and R2, as these cohorts were enrolled in the same study program at THI.

Table 5.18: Coded VET-professions and APE sub dimensions, data of R6 & R9

VET-Profession	N	%	M APE-M	M APE-ED	M APE-EC
other	1	3.3	(1.5)	(.90)	(.75)
mechanic	4	13.3	.81	.30	-.50
mechatronics	17	56.7	1.35	.62	.26
electronic	8	26.7	.72	.65	.28
Total VET R6, R9	31		1.12	.57	-.18
comparison (all runs, incl. non-VET)	181		.60	.27	-.02

5.3.8 APE vs. test performances

A Spearman rank-order correlation coefficient was computed to assess the relationship between the APE dimensions and the students' overall test performances (TP) after the laboratory experiments in the different modes. No generally valid significant correlations between test performances after hands-on and after simulated experiments and the APE dimension were found. (Five individual items showed statistically significant correlations, but were distributed over all APE-dimensions. These statistics allowed no further insights and are presented on page 251 in the appendix.)

Nevertheless, looking to Figure 5.8 made clear that correlations existed. It was possible to identify regions in the plot for VET and non-VET separately. A least-square fit for a linear correlation between APE and TP was calculated. The results are presented in Table 5.19 and Table 5.20. The correlation for former VET-participants was statistically significant, while the correlation for non-VET and all participants showed no statistical significance. Table 5.21 presents the results of a t-test for difference of the slopes of correlations of non-VET and former VET students. The difference between the correlations of both compared groups was statistically significant ($t = 2.53$, $p < .014$).

This result was quite interesting. Figure 5.8 presents the summary of the correlation APE vs. test performance of German and International runs from both research phases. In the following, the different research phases and participant groups were evaluated separately.

5.3.8.1 Analysis of German runs

First research phase (hands-on vs. simulations)

Figure 5.9 presents the scatter plot for all data of German participants (R1, R2). The plot shows the students' APE versus the students' average test performance in the tests. While former VET-participants showed better test results, if they had higher APE, no statistically significant effect was found for non-VET (R2, Table 5.19 to Table 5.21).

In separate comparisons of the results after hands-on and simulated experiments former VET participants had statistically significant positive correlations between

Table 5.19: Amount of Practical Experience vs. Test performances, German Runs Correlations, Students *without* former VET education

		German non-VET									
APE vs.	shown in	N	B	SE	β	F	df1	df2		p	R^2
TP all	Figure 5.8	34	.21	.21	.17	.97	1	32		.332	.03
TP Ph1 (R2)	Figure 5.9	10	-.33	.40	-.28	.68	1	8		.433	.08
TP Ph1 sim.	Figure 5.10	9	.15	.39	.14	.15	1	7		.715	.02
TP Ph1 hands-on	Figure 5.11	10	-.46	.43	-.35	1.15	1	8		.316	.13
TP Ph2 (R6, R9)	Figure 5.12	24	.49	.24	.40	4.08	1	22	†	.056	.16

Note. † = p < .10. * = p < .05. ** = p < .01.
Positive B-values describe advantages in test performance of students with high APE.

Table 5.20: Amount of Practical Experience vs. Test performances, German Runs Correlations, Students *with* former VET education

		German VET									
APE vs.	shown in	N	B	SE	β	F	df1	df2		p	R^2
TP all	Figure 5.8	46	.91	.17	.64	29.84	1	44	**	.001	.40
TP Ph1 (R2)	Figure 5.9	16	1.16	.30	.71	14.52	1	14	**	.002	.51
TP Ph1 sim.	Figure 5.10	16	1.56	.39	.73	15.58	1	14	**	.001	.53
TP Ph1 hands-on	Figure 5.11	16	.78	.38	.49	4.34	1	14	†	.056	.24
TP Ph2 (R6, R9)	Figure 5.12	30	.78	,20	.59	15.07	1	28	**	.001	.35

Note. † = p < .10. * = p < .05. ** = p < .01.
Positive B-values describe advantages in test performance of students with high APE.

Table 5.21: Amount of Practical Experience vs. Test performances, German Runs Correlations, Checks for significant differences of effects between VET and non-VET

APE		t-test for different effects					
vs.	shown in	N	B	SE	t		p
TP all	Figure 5.8	80	.70	.28	2.53	*	.014
TP Ph1 (R1, R2)	Figure 5.9	26	1.48	.49	3.05	**	.006
TP Ph1 sim.	Figure 5.10	25	1.41	.55	2.56	*	.018
TP Ph1 hands-on	Figure 5.11	26	1.24	.56	2.22	*	.037
TP Ph2 (R6, R9)	Figure 5.12	54	.29	.32	.90		.371

Note. † = p < .10. * = p < .05. ** = p < .01.
Positive coefficients correspond with higher positive slope of former VET participants

APE and TP after both modes. The correlation for students without former VET was statistically insignificant after both simulations (Figure 5.10) and hands-on experiments (Figure 5.11).

The difference of VET and non-VET correlations was statistically significant for both modes (see Table 5.21).

It was checked if the significant slopes of the former VET participants in Figure 5.10 and Figure 5.11 differed statistically significantly between modes. Table 5.22 presents the results of the relevant t-test. The difference between the correlations was not statistically significant.

Former VET participants with higher APE were more successful in tests than former VET participants with lower APE – the measured effect was stronger after simulations, but did not differ in a statistically significant way between the compared modes in the first research phase.

Table 5.22: Amount of Practical Experience vs. Test performances, German Runs Correlations, VET and non-VET participants, Checks for significant differences of effects between test performance after hands-on experiments and simulated lab.

APE		t-test for different effects				
vs.	shown in	N	B	SE	t	p
TP Ph1 VET	Figure 5.10/Figure 5.11	32	-.77	.54	-1.42	.166
TP Ph1 non-VET	Figure 5.10/Figure 5.11	19	-.61	.59	-1.03	.318

Note: Positive coefficient corresponds with stronger slope (dependency) of hands-on test performance compared to performance after simulations

Further analysis separated by learning objectives

To find out if specific learning objectives (see page 339) in connection with APE have different influence on the test performance of a student, non-parametric correlations

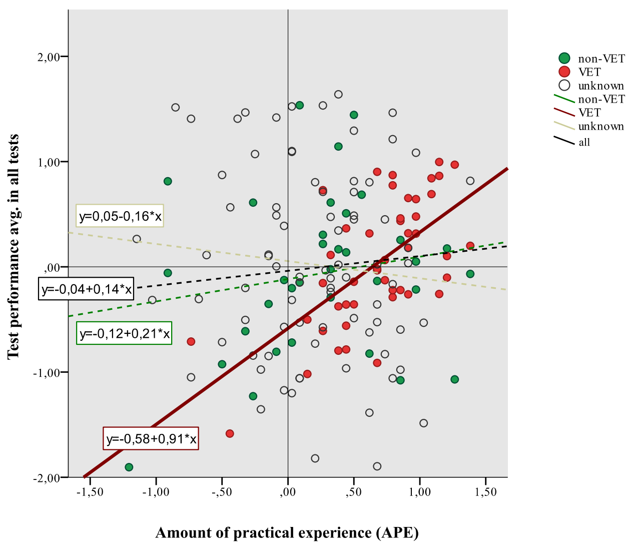

Figure 5.8: All available runs, R1, R2, R5, R6, R8, R9: Correlations, Amount of Practical Experience vs. Student's average test performance

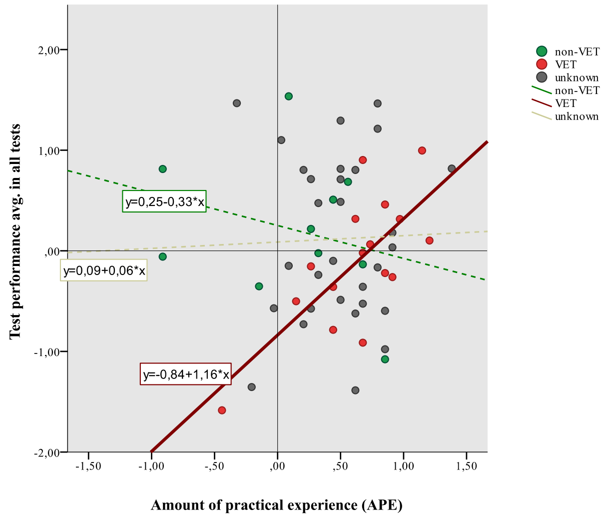

Figure 5.9: First Phase, R1 and R2, German students: Correlations, Amount of Practical Experience vs. Student's average test performance

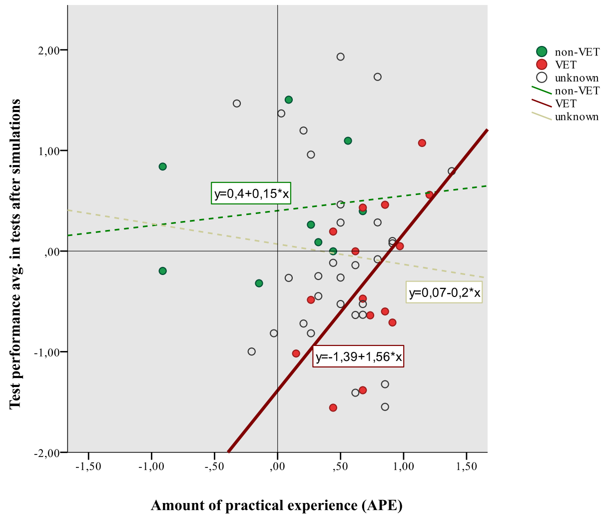

Figure 5.10: First Phase, R1 and R2, German students: Correlations, Amount of Practical Experience vs. Student's test performance after *simulations*

were calculated. The results are shown in Table 5.23.

Former VET participants showed positive correlations between APE and test performance for the learning objectives "Battery Behaviour" and "Experimental Setup" ($p < .1$). VET participants tend to perform well in these categories, if they achieved a high APE value. For the other categories no statistically significant trend was detected.

Non-VET students showed the opposite trend for "Experimental Setup" (hands-on: $\rho = -.259$, $p = .048$): If a non-VET participant had a high APE, he tended to perform bad in tasks related "Experimental Setup" after hands-on experiments.

Second research phase (hands-on experiments vs. hidden simulations)

Figure 5.12 shows data from the German runs (R6, R9) in the second phase, compar-

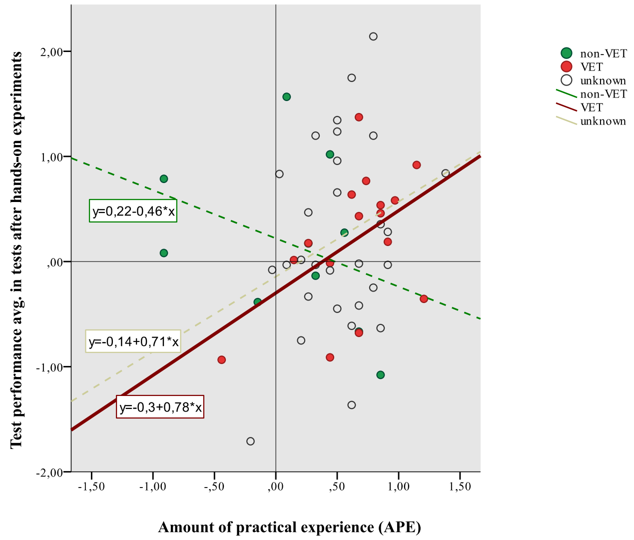

Figure 5.11: First Phase, R1 and R2, German students: Correlations, Amount of Practical Experience vs. Student's test performance after *hands-on experiments*

Table 5.23: APE: Spearman rank-order correlation between the test performance and APE, grouped for VET and learning objective. R1, R2, first study phase.

		no VET info		non-VET		VET	
		ho	sim	ho	sim	ho	sim
Battery Behaviour	ρ	.109	.026	.035	.067	.153	.191
	p	.104	.673	.758	.549	† .075	* .028
	N	225	270	80	82	137	133
Battery Parameters	ρ	-.132	-.010	-.303	-.117	-.126	.236
	p	.215	.931	.104	.588	.381	.143
	N	90	80	30	24	50	40
Battery System Design	ρ	-.095	-.143	.000	-.194	.133	.191
	p	.504	.260	1.00	.412	.537	.265
	N	52	64	20	20	24	36
Experimental Setup	ρ	.083	-.064	-.259	-.005	.229	.324
	p	.289	.389	* .048	.969	* .032	** .002
	N	164	182	59	55	88	92

Note. † = p < .10. * = p < .05. ** = p < .01.
Positive ρ values describe advantages of a high APE in the tests.

ing hands-on experiments with hidden simulations. This time, positive and statistically significant correlations between APE and TP were found for both VET and non-VET (see Table 5.19, Table 5.20). The more experience the participants possessed in terms of APE, the better they performed in the knowledge tests. The difference between the VET and non-VET slope was not statistically significant (see the last line in Table 5.21).

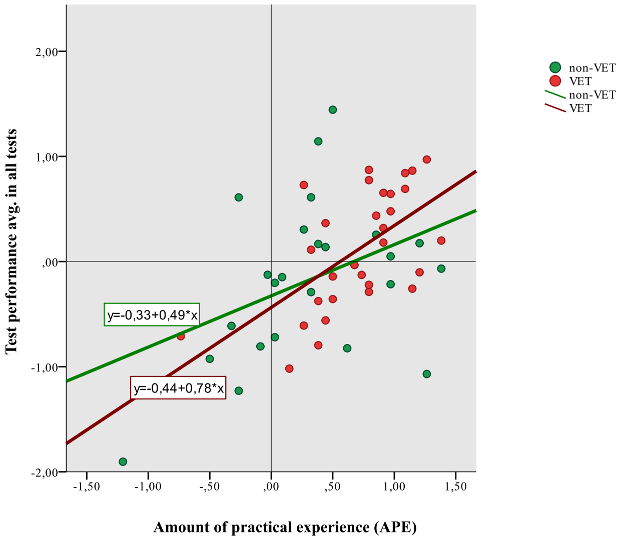

Figure 5.12: Second Phase, R6 and R9, German students: Correlations, Amount of Practical Experience vs. Student's average test performance

5.3.8.2 Analysis of the international runs

The international run in both research phases was also plotted to investigate the correlation between APE and TP. The results of the hands-on vs. simulations phase are presented in Figure 5.13. The results of the second research phase are shown in Figure 5.15. Table 5.24 shows the calculated results of correlations for the international runs. In the following, all data is reported and considered without the outlier that is visible in Figure 5.13.

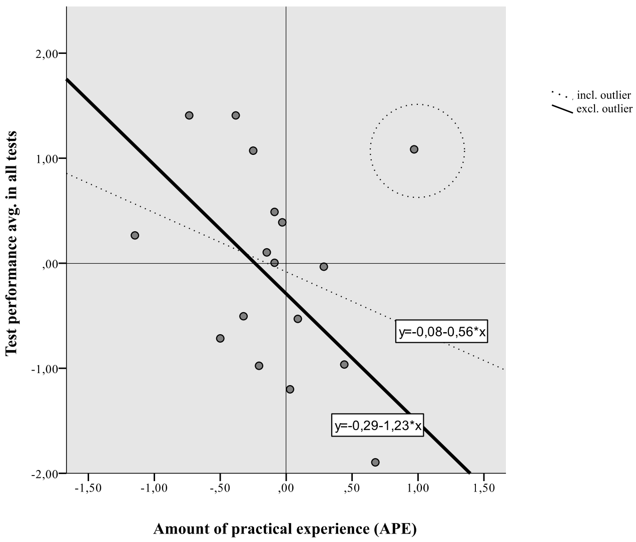

Figure 5.13: First Phase, R5, International students: Correlations, Amount of Practical Experience vs. Student's average test performance. An outlier would influence the correlation.

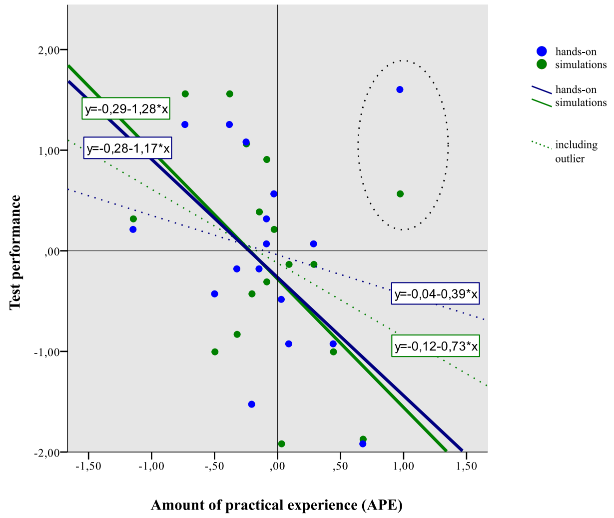

Figure 5.14: First Phase, R5, International students: Correlations, Amount of Practical Experience vs. Student's average test performance, for both modes separately

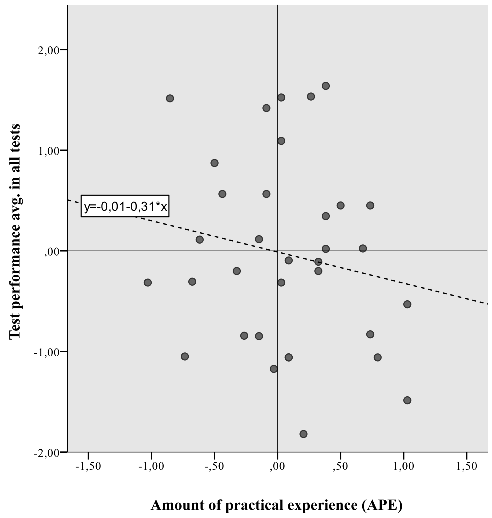

Figure 5.15: Second Phase, R8, International students: Correlations, Amount of Practical Experience vs. Student's average test performance

Comparing German and international participants

In contrast with the German students, the international students showed the opposite trend, which was statistically significant. The higher the stated APE of an international student, the lower the respective participant performed in the knowledge tests after experimenting ($N = 16$, $B = -1.23$, $SE = .47$, $F(1,14) = 6.73$, $p = .021$). This trend did not differ between content areas taught with hands-on experiments and content areas taught with simulations (see Figure 5.14).

The combined results displayed in Figure 5.9 through Figure 5.15 clearly show that the probability of a German UAS participant with little practical experience and good test performance is low (upper left quadrant nearly empty). In contrast, for international participants, this quadrant contained most results.

Table 5.24: Amount of Practical Experience vs. Test performances, International runs correlations.

		International									
APE vs.	shown in	N	B	SE	β	F	df1	df2		p	R^2
TP Ph1 R5	Figure 5.13	16	-1.23	.47	-.57	6.73	1	14	*	.021	.276
TP Ph1 R5 sim.	Figure 5.14	16	-1.28	.55	-.52	5.30	1	14	*	.037	.223
TP Ph1 R5 hands-on	Figure 5.14	16	-1.17	.47	-.56	6.33	1	14	*	.025	.262
TP Ph2 R8	Figure 5.15	34	-.31	.31	-.17	.99	1	32		.327	.000

Note. † = p < .10. * = p < .05. ** = p < .01.
Positive B-values describe advantages in test performance of students with high APE.
R5 was calculated without the outlier.

5.3.8.3 APE vs. test performance in individual learning modes

To find out if the trends were based on the TP of one of the two compared modes, a Spearman rank-order correlation coefficient was computed to assess the relationship between the individual APE items and dimensions and the test performance after simulations and hands-on experiments separately. No significant correlation between test performances after hands-on and after simulated experiments and the APE dimension were found. As the results of German-VET, German-non-VET and internationals were mixed for this analysis, these results have to be seen critically. The following subsection investigates the superior learning modes of single students with their APE.

5.3.9 APE vs. superior learning mode

A Spearman rank order correlation analysis was performed to find correlations between the APE dimensions and the students' superior learning mode (performance after hands-on relative to performance after simulated experiments, as defined in subsection 4.2.7). The results are presented in Table 5.25. A positive, but weak, correlation between the two items was found only for the APE-ED dimension (Spearman's $\rho(71) = .22$, $p = .061$). Students who stated a lot of experience in the "electronics do" category learned better with hands-on experiments than with simulated laboratories. The dimension APE-ED describes practical experience in procedures which were taught during the study program, it was the only dimension where no significant differences regarding the completion of a VET program were found. Also, for the overall APE a correlation was detected, which had similar intensity and showed somewhat statistical relevance ($p = .099$). As the results of German-VET, German-non-VET and internationals were mixed for this analysis, these results have to be seen critically.

The table with the analysis for individual items is presented on page 251 in the appendix. It did not lead to any more profound insights.

Table 5.25: APE: Spearman rank-order correlation between superior learning mode and all APE dimensions – all runs first study phase.

[a]	superior learning mode			
	ρ		p	N
APE	.20	†	.099	71
APE-E	.17		.164	71
APE-ED	.22	†	.061	71
APE-EC	.08		.493	71
APE-M	.13		.262	71
APE-MC	.05		.657	71
APE-MD	.19		.122	71

Note. † = p < .10.
Positive "superior learning mode"-values describe advantages of hands-on.
[a] As the results of German-VET, German-non-VET and internationals were mixed for this analysis, these results have to be seen critically.

5.3.9.1 Analysis of German runs (incl. VET vs. non-VET)

Figure 5.16 presents an interesting aspect found in the first phase for the participants of the German runs, when comparing the APE to the superior learning mode of the individual students. The former VET participants can be separated clearly from the students without a VET degree.

1. As discussed above, the figure shows that the former VET participants generally have a higher APE (pos. offset on the x-axis).

2. Former VET-participants tend to have better learning results after hands-on than after simulations (pos. offset on the y-axis, relative to non-VET).

3. The opposite direction of slopes in the graphs comparing APE vs. TP for simulations (Figure 5.10) and hands-on (Figure 5.11), which were compared in Table 5.22, lead to very similar slopes of VET and non-VET regarding the superior learning mode: German participants with low APE tend to achieve better results after hands-on experiments, while German participants with higher APE performed better in tests after simulations.

A least-square fit for a linear correlation between APE and TP was calculated. The results are presented in Table 5.26. Even where the negative correlations for VET and non-VET are clearly visible in Figure 5.16, neither correlation was statistically significant. This was due to the nonlinear type of correlation. Linear fitting cannot accurately describe the curved correlations visible in the figure.

Thus, a non-parametric Spearman rank order correlation was calculated. The results are presented in Table 5.27. It showed somewhat statistically significant trends for both VET and non-VET (VET: ρ = -.42; non-VET: ρ = -.59; $p \approx .10$).

From Figure 5.16 and the calculated results it can be drawn, that – considering German participants with VET and without VET degree separately – the higher the APE the better the performance after simulations when compared with hands-on experiments. When former VET participants reported experience with all items (in

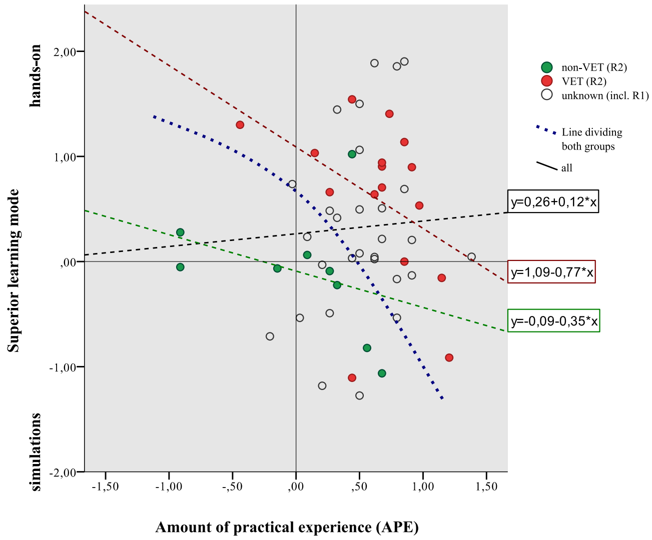

Figure 5.16: First Phase, R1 and R2, German students: Correlations, Amount of Practical Experience vs. Student's superior learning mode

average "yes" or "rather yes"), they were performing equally in both modes. The same goes for non-VET students, which reported approximately 50% (or less) of the items. Besides former VET participants' shift towards hands-on, the intensity of the trend did not differ between the VET and non-VET groups.

This leads to the hint that overall data regarding the first run (R1), when combining VET and non-VET students, might lead to wrong conclusions regarding the trend for this analysis. These two groups need *separate* analysis.

Table 5.26: Amount of Practical Experience vs. Student's superior learning mode, *linear* correlations

APE vs.	shown in	N	B	SE	β	F	df1	df2	p	R^2
SLM Ph1 R2 VET	Figure 5.16	16	-.77	.46	-.41	2.81	1	14	.116	.108
SLM Ph1 R2 non-VET	Figure 5.16	9	-.35	.36	-.34	.92	1	7	.369	-.010
SLM Ph1 R1&R2 all [a]	Figure 5.16	54	.12	.25	.07	.24	1	52	.629	-.015
SLM Ph1 Intl R5	Figure 5.17	17	.34	.34	.25	1.03	1	15	.327	.002

Note. † = $p < .10$. * = $p < .05$. ** = $p < .01$.
Positive B-values describe advantages in test performance of students with high APE.
R^2 = adjusted R^2.
[a] = As of the offset for VET vs. non-VET this value can be seen critically.

Superior learning mode, grouped for VET and non-VET

A Spearman rank-order correlation coefficient was computed to assess the relationship between the APE dimensions and the students' superior learning mode, grouped for VET and non-VET students. The results are presented in Table 5.28. All statistically significant correlations of the dimensions suggested that students who had *not* previously had much practical experience performed better in tests after hands-on experiments compared to simulated experiments (SLM).

Table 5.27: APE: Spearman rank-order correlation Amount of Practical Experience vs. Student's superior learning mode – all individual runs first study phase

APE		superior learning mode			
	shown in	ρ		p	N
Ph1 R2 VET	Figure 5.16	-.42		.103	16
Ph1 R2 nonVET	Figure 5.16	-.59	†	.097	9
Ph1 R1&R2 all [a]	Figure 5.16	.10		.477	54
Ph1 Intl R5	Figure 5.17	.25		.337	17

Note. † = $p < .10$.
Positive "superior learning mode"-values describe advantages of hands-on.
[a] = As of the offset for VET vs. non-VET this value can be seen critically.

For these statistics on individual APE items see section D.3 in the appendix.

Table 5.28: APE: Spearman rank-order correlation between the superior learning mode and all the APE dimensions, grouped for VET. R2, first study phase.

APE vs. SLM	VET N = 16			non-VET N = 9		
	ρ		p	ρ		p
APE	-.42		.103	-.59	†	.097
APE-E	-.23		.392	-.31		.417
APE-EC	-.14		.617	-.35		.351
APE-ED	-.18		.501	.08		.847
APE-M	-.66	**	.005	-.67	*	.049
APE-MC	-.62	*	.011	-.80	*	.010
APE-MD	-.34		.205	-.17		.660

Note. † = $p < .10$. * = $p < .05$. ** = $p < .01$.
Positive "superior learning mode"-values describe advantages of hands-on.
[a] = All participants of the VET/non-VET group answered identically.

5.3.9.2 Analysis of international runs

Figure 5.17 presents the gained data regarding APE and superior learning mode for the international run of the first phase. The regression statistics are presented in Table 5.26. International participants with higher APE performed somewhat better in the hands-on condition, but the identified trend was not statistically significant.

Comparing German and international participants

The positive connection between APE and SLM in internationals ran contrary to the behaviour of the German participants. The difference in slope was checked for statistical significance, the results are presented in Table 5.29. A significant difference was found: While international students with higher APE tended to perform better in hands-on experiments, German former VET participants with higher APE tended towards better results with simulations. For German non-VET, no statistically significant difference to internationals was identified.

Comparing Figure 5.16 to Figure 5.17, it becomes clear that the probability of

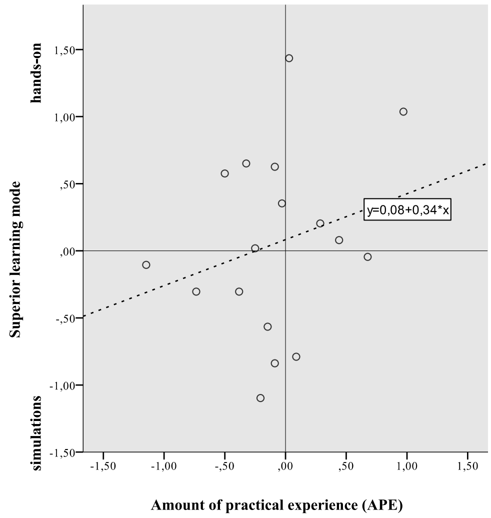

Figure 5.17: First Phase, R5, International students: Correlations, Amount of Practical Experience vs. Student's superior learning mode

a German UAS participant with little practical experience and simulations as superior performing learning mode is low (lower left quadrant nearly empty), while for international participants this quadrant contained most results.

Table 5.29: Amount of Practical Experience vs. superior learning mode, Checks for significant differences of effects of international and German participants

APE		t-test for different effects					
vs.	shown in	N	B	SE	t		p
SLM Ph1 R2 VET vs. R5 int.	Figure 5.16/Figure 5.17	33	-1.12	.57	-1.95	†	.060
SLM Ph1 R2 non-VET vs. R5 int.	Figure 5.16/Figure 5.17	26	-.69	.49	-1.39		.176

Note. † = p < .10. Positive coefficient corresponds with stronger slope (pos. dependency) of hands-on test performance compared to performance after simulations

5.3.10 Summary of results of DS-B

Summary of section 5.3 “Amount of Practical Experience (DS-B)”

Personal interest can be seen as a precursor for learning success since it has a profound effect on cognitive functioning, including but not limited to focus and attention as well as knowledge gain [119]. A widely renowned concept that describes interest on a general level is Holland´s RIASEC-typology (refer section 5.4, DS-C). As this concept might be too broad to capture students’ interests in a science-based topic, a questionnaire was created. It comprised of 17 questions about prior experiences in practical tasks with higher specificity towards technical practical experiences.

Amount of Practical Experience (APE)

- Based on the collected data, the overall APE scale had a high reliability. Furthermore two electronic and two mechanic sub-scales were identified.
- In general, participants who completed a VET before university reported statistically significantly higher amounts of practical experience than their peers who had not. None of the employed items showed statistically significantly higher experience of the non-VET group.
- The APE difference (VET vs. non-VET) was much more significant regarding mechanical topics than electrical topics. This might be due to the electrical experience gained during UAS studies, which accounts for both VET and non-VET, while mechanical experience seemed to be gained mainly from the VET-programme.
- In the field of electronics, former VET participants showed more experience in the APE-EC category compared to APE-ED, where no significant differences were found.

Amount of Practical Experience (APE) vs. test performance

- Among all subgroups, the identified correlations between APE and average test performance cancel each other out. To be able to explain the correlation between the dimensions, these subgroups (Germans, Internationals, VET, non-VET) have to be analysed separately.
- German participants with higher APE performed better compared to Germans with lower APE.
  - This is due to former VET-participants, who performed statistically significantly better when a high APE was reported (mainly based on tasks on “Battery Behaviour” and “Experimental Setup”).

- The effect for non-VET-participants was overall insignificant.

- In the first phase international participants showed the opposite trend: Internationals with lower APE generally performed better compared to internationals with high APE.

- Analysed separately for both modes, former VET participants showed statistically significant positive correlations between APE and TP in both modes (Participants with a lot of practical experiences performed well in the tests), while the correlation for students without former VET was statistically insignificant. The trends for simulations and the trend for hands-on experiments did not differ statistically significantly – neither for VET nor for non-VET.

- VET compared to non-VET behaviour differed statistically significantly in both modes.

Amount of Practical Experience (APE) vs. superior learning mode

- German participants with higher APE tended towards the superior learning mode "simulations", while Germans with lower APE tended towards "hands-on" experiments. It is important to precede carefully, as a big offset (equivalent to 50% of experienced APE-items) regarding the better performing mode between VET and non-VET students was found. Nevertheless, the mentioned trend (the slope, not general) was identical in both groups.

- The international participants showed the opposite trend (very weak and not statistically significant), participants with lower APE tended to learn somewhat better with simulated laboratories than with hands-on experiments.

Second research phase

- Overall, the APE of participants of the first and second phase could be considered as equal, except for the sub-dimension APE-ED of participants without VET.

- In the second research phase, when hiding the simulations, no differences between VET and non-VET were found regarding the effects of APE on the learning outcomes.

- Like in the first phase, a higher APE was connected to generally better average test performances among the German students.

- The international participants showed the opposite trend again, but in the single run recording that information, the trend was not statistically significant.

5.4 Personality/RIASEC (DS-C)

This investigation was done using the methodology outlined in section 4.4. Thirty-five students responded. A summary of the findings is presented on page 129.

5.4.1 Study participants in general

The Explorix test manual [126] states approximate RIASEC values regarding different categories: Average participant, age, educational level, and job. In the following, these values are compared to the data collected in the present study.

Average Participant

The item difficulty for all items [126, p. 77ff] allows for comparison to the average participant of Explorix (Table 5.30). This was also possible for the average participants of AIST (Table 5.31) [122, p. 53ff].

Table 5.30: RIASEC: Comparison to Explorix mean participant values

	R	I	A	S	E	C
N	35	35	35	35	35	35
Mean	.81	.68	.36	.50	.67	.49
Std. Deviation	.15	.16	.21	.18	.22	.19
Item difficulty [126, p. 77ff]	.48	.56	.53	.55	.57	.47
Deviation	.33	.13	-.18	-.06	.09	.03

Table 5.31: RIASEC: Comparison to AIST mean participant values

	R	I	A	S	E	C
N	29	29	29	29	29	29
Mean	.73	.71	.41	.50	.59	.50
Std. Deviation	.14	.15	.17	.18	.17	.15
Item difficulty [122, p. 53ff]	.41	.45	.45	.47	.49	.38
Deviation	.32	.27	-.04	.03	.11	.12

As shown in Table 5.30 and Table 5.31, the participating B. Eng. electric mobility students had unexpectedly high R and I values considering item difficulties, while the rest of categories were similar to the published difficulties.

The collected data regarding the A dimension differed only from the published Explorix difficulties, while the study participants were similar regarding the published AIST difficulties. As no differences between VET vs. non-VET regarding the A dimension were found (see below), the focus was only on R and I.

In detail, for the R dimension and Explorix, Capabilities (M = .85, SD = .18 vs. item difficulty M = .55 [126, p. 78]) and Activities (M = .78, SD = .16 vs. item difficulty M = .41 [126, p. 77]) were higher. In the I dimension, Capabilities (N = 35, M = .80, SD = .19) were rated clearly higher than the average test set participant (item difficulty M = .60 [126, p. 78]). In I activities, non-VET (M = .62, SD = .15) scored higher than average (item difficulty M = .46 [126, p. 77]), while the VET values (M = .47, SD = .21) are similar to the published item difficulty [126].

Scaling of values of the test manual

The full Explorix is based on four different types of questionnaire: activities, capabilities, professions, and self-assessment. For the research, only two of four components were used: activities and capabilities. As item difficulties differ significantly [126, p. 36] between the four questionnaires, correction factors were calculated and further applied (ref. Table 5.32) to be able to compare the statistics of the test manual with the participants of the research. For both the R and I category, a factor of 1.23 was determined between the two used (activities, capabilities) and all four questionnaires.

Table 5.32: Explorix, corrected average item difficulties based on [126]

		Exp-R	Exp-I	
Age	19-22 y	.43	.49	[126, p. 75]
	23-29 y	.57	.66	
Edu. Level	Uni / UAS	.52	.71	[126, p. 76]
Job	Student	.42	.55	[126, p. 76]
Participants		.81	.68	
VET		.87	.62	
non-VET		.76	.73	

Age

The participants (age M = 22.5 y) have higher scores (N = 35, M = .81, SD = .15) for “Realistic” than the average participant of Explorix in the same age. For “Investigative” non-VET (N = 18, M = .73, SD = .13) have higher scores than the average participant, and the VET (N = 16, M = .62, SD = .18) are similar to the average.

Educational level: University / UAS

The participants (N = 35, M = .81, SD = .15) have higher scores for “Realistic” than the average participant of Explorix at the same educational level. For “Investigative” non-VET (N = 18, M = .73, SD = .13) have average scores, and the VET (N = 16, M = .62, SD = .18) are below the average.

Main job = Student

The participants (N = 35, M = .81, SD = .15) had higher scores for “Realistic” than the average higher-education level student. The participants (N = 35, M = .68, SD = .16) also had higher scores for “Investigative” than the average higher-education level student.

Correlations between RIASEC Dimensions

Holland found the most intense relationship between neighbouring dimensions [126, p. 14] in the circular RIASEC arrangement. Table 5.33 supports this finding (except E): A Spearman rank-order correlation coefficient was computed to assess the relationship between the RIASEC dimensions. Trends were reported in bold font. The strongest correlation was found between neighbouring dimensions.

Table 5.33: RIASEC: Spearman rank-order correlations between dimensions

		R	I	A	S	E
I	ρ	.27				
	p	.163				
	N	29				
A	ρ	-.05	.22			
	p	.812	.243			
	N	29	29			
S	ρ	.00	.13	† **.36**		
	p	.984	.497	.054		
	N	29	29	29		
E	ρ	† **.33**	† **.34**	* **.46**	.12	
	p	.078	.072	.013	.540	
	N	29	29	29	29	
C	ρ	* **.41**	.27	.19	.00	** **.49**
	p	.029	.157	.333	.992	.007
	N	29	29	29	29	29

Note. p < .10 marked bold.
† = p < .10. * = p < .05. ** = p < .01.

5.4.2 Study participants with and without Vocational Education and Training

An independent samples t-test was conducted to compare the RIASEC values of VET and non-VET students. The results are presented in Table 5.34. The table presents all subcategories of both test sets as well as the combined values. A Shapiro-Wilk test was conducted for all compared groups. If a subgroup was not normally distributed, the descriptives were marked with ^{s}. Additionally, non-parametric tests were conducted to allow for comparison of not normally distributed subgroups. In case of significance of the U-test, further descriptives (Kurtosis, Skewness, Medians) were analysed to clarify that the U-test result was based mainly on different medians.

Looking at the data presented in the table, it becomes evident that the AIST and the Explorix test set gained similar results, which acknowledges these methods.

When comparing the results, trends between VET and non-VET were found. The R and I category differed most, while the other categories were showing no differences between the groups.

In sum, VET-graduates scored – compared to students without former VET – higher values in the "Realistic" category ($t(26) = -1.94$, $p = .064$, Cohen's $d = -.73$, trend), while their scores in the "Investigative" category were lower (t(24) = 2.00, p = .057, Cohen's d = .76, trend).

Table 5.34: RIASEC: Explorix, AIST and Sum compared for VET / non-VET using an independent samples t-test as well as a non-parametric Mann-Whitney U-test

	non-VET			VET								Mann-Whitney		
	N	M	SD	N	M	SD	df	t		p	d	Z		p
Exp_Inte_R	18	.74	.19	16	.84	.10	32	-1.87	†	.072	-.64	-1.742	†	.088
Exp_Inte_I	18	.62	.15	16	.47	.21	32	2.42	*	.022	.83	2.167	*	.033
Exp_Inte_A	18	.35	.20	16	.36	.24	32	-.06		.955	-.02	-.139		.905
Exp_Inte_S	18	.43	.25	16	.37	.16	32	.88		.387	.30	.770		.463
Exp_Inte_E	18	.69	.22	16	.66	.23	32	.42		.680	.14	-.489		.646
Exp_Inte_C	18	.35	.20	16	^{s}*.30*	*.31*						1.149		.266
Exp_Capa_R	18	^{s}*.79*	*.21*	16	^{s}*.90*	*.12*						-1.671		.109
Exp_Capa_I	18	^{s}*.84*	*.16*	16	.76	.22	32	1.17		.252	.40	1.039		.313
Exp_Capa_A	18	^{s}*.34*	*.28*	16	.38	.23	32	-.36		.723	-.12	-.645		.528
Exp_Capa_S	18	.58	.24	16	.62	.19	32	-.58		.564	-.20	-.628		.551
Exp_Capa_E	18	.59	.27	16	^{s}*.71*	*.25*						-1.521		.135
Exp_Capa_C	18	.65	.22	16	^{s}*.72*	*.24*						-1.117		.281
Exp_R	18	.76	.18	16	^{s}*.87*	*.09*						-2.004	*	.046
Exp_I	18	^{s}*.73*	*.13*	16	.62	.18						2.088	*	.039
Exp_A	18	.35	.21	16	.37	.22	32	-.24		.811	-.08	-.242		.825
Exp_S	18	.51	.22	16	.49	.15	32	.17		.870	.06	.052		.959
Exp_E	18	.64	.23	16	.68	.20	32	-.57		.570	-.20	-.414		.695
Exp_C	18	.50	.18	16	.51	.21	32	-.13		.899	-.04	-.243		.825
Exp_sum	18	.58	.12	16	.59	.09	32	-.23		.817	-.08	-.293		.772
AIST_R	14	.69	.18	14	.77	.09	26	-1.53		.139	-.58	-1.223		.227
AIST_I	14	.74	.17	14	^{s}*.68*	*.13*						1.336		.194
AIST_A	14	.41	.17	14	^{s}*.41*	*.17*						.438		.667
AIST_S	14	.53	.22	14	.48	.14	26	.72		.476	.27	.253		.804
AIST_E	14	.58	.20	14	.60	.14	26	-.27		.793	-.10	-.276		.804
AIST_C	14	.51	.17	14	.49	.13	26	.38		.704	.15	.462		.667
AIST_sum	14	.58	.13	14	.57	.08	26	.14		.893	.05	.230		.839
R	14	.73	.17	14	.82	.07	26	-1.94	†	.064	-.73	-1.471		.150
I	14	.75	.12	14	.65	.14	26	2.00	†	.057	.76	1.839	†	.069
A	14	.38	.17	14	.38	.16	26	.14		.892	.05	.138		.910
S	14	.50	.21	14	.48	.13	26	.31		.761	.12	.115		.910
E	14	.61	.22	14	.63	.14	26	-.32		.754	-.12	.000		1.000
C	14	.52	.15	14	.50	.13	26	.32		.752	.12	.299		.769

Note. † = p < .10. * = p < .05. ** = p < .01. d = Cohen's d, Z (pos. = higher values of non-VET).
^{s} = Shapiro-Wilk p < .05, no normal distribution.
One dataset was ignored for this analysis as of missing information about VET.

In the Explorix test, questions were analysed in categories based on their nature, questions for interests "interested in doing those/like the activities" showed somewhat higher differences between VET and non-VET than stated capabilities "able to do". The German questions including English translations for the two relevant categories are presented in the appendix, pages 253ff.

Correlations between RIASEC dimensions separated for VET/non-VET

A Spearman rank-order correlation coefficient was computed to assess the relationship between the RIASEC dimensions, depending on whether a participant did a

VET before enrolling or not. The results are presented in Table 5.35. Trends were reported in bold font. Looking to the table, the general finding of strong neighbouring correlations can also be supported for the subgroups.

Table 5.35: RIASEC: Spearman rank-order correlation between dimensions and average test performance. top right: non-VET. bottom left: VET

		R	I	A	S	E	C	TP
R	ρ		* **.64**	.16	.15	* **.62**	.37	-.17
	p		.013	.578	.605	.018	.197	.612
	N		14	14	14	14	14	11
I	ρ	.06		.26	-.02	.42	.40	.44
	p	.846		.378	.940	.139	.158	.180
	N	14		14	14	14	14	11
A	ρ	-.36	.12		.43	† **.48**	† **.52**	.52
	p	.202	.686		.126	.083	.055	.102
	N	14	14		14	14	14	11
S	ρ	-.30	† **.50**	.23		.16	.13	-.03
	p	.291	.072	.436		.584	.659	.937
	N	14	14	14		14	14	11
E	ρ	.07	.14	.21	-.11		* **.64**	-.05
	p	.811	.642	.464	.714		.015	.894
	N	14	14	14	14		14	11
C	ρ	* **.57**	.15	-.18	-.20	.41		.21
	p	.033	.615	.533	.483	.144		.537
	N	14	14	14	14	14		11
TP	ρ	.26	.11	-.08	.01	-.19	.30	
	p	.383	.720	.803	.986	.541	.316	
	N	13	13	13	13	13	13	

Note. $p < .10$ marked bold.
† = $p < .10$. * = $p < .05$.

Generally, correlations between the dimensions for non-VET were stronger. The correlations were compared ($\Delta\rho$, p) between VET and non-VET. To check if the correlations differed statistically significantly between VET and non-VET students, Equation I.13 was used. The results are presented in Table 5.36. For two correlations, a statistically supported difference was identified (A-C: $p = .073$; R-I: $p = .097$). In both cases, the correlation for former VET participants was absent ($p > .5$), while for students without VET statistically supported correlations were measured ($\rho > .5$, $p < .1$).

Analysing the scatter diagram shown in Figure 5.18, no meaningful difference of the A-C correlation between VET and non-VET could be found. Nevertheless, the R-I correlation (see Figure 5.19) showed interesting behaviour: while former VET participants had a high R-value, independent of their I-value, non-VET's R-value depended significantly on the I-value.

Non-VET Students show high “Investigative” values if their “Realistic” values are high, while students with a VET degree show no dependency and have high “Realistic” values generally.

Table 5.36: RIASEC: Spearman rank-order correlation between dimensions: Statistical significance of the difference of the Spearman rank-order correlations between former VET participants and students without VET

			VET		non-VET				
			ρ	N	ρ	N		z	p
A	-	**C**	-.18	14	.52	14	†	-1.80	.073
R	-	**I**	.06	14	.64	14	†	-1.66	.097
R	-	E	.07	14	.62	14		-1.54	.123
I	-	S	.50	14	-.02	14		1.32	.186
R	-	A	-.36	14	.16	14		-1.28	.202
R	-	S	-.30	14	.15	14		-1.09	.275
S	-	C	-.20	14	.13	14		-.79	.429
E	-	C	.41	14	.64	14		-.73	.463
A	-	E	.21	14	.48	14		-.72	.473
I	-	E	.14	14	.42	14		-.72	.474
I	-	C	.15	14	.40	14		-.64	.522
S	-	E	-.11	14	.16	14		-.63	.527
R	-	C	.57	14	.37	14		.62	.539
A	-	S	.23	14	.43	14		-.54	.593
I	-	A	.12	14	.26	14		-.33	.740

Note. † = p < .10.

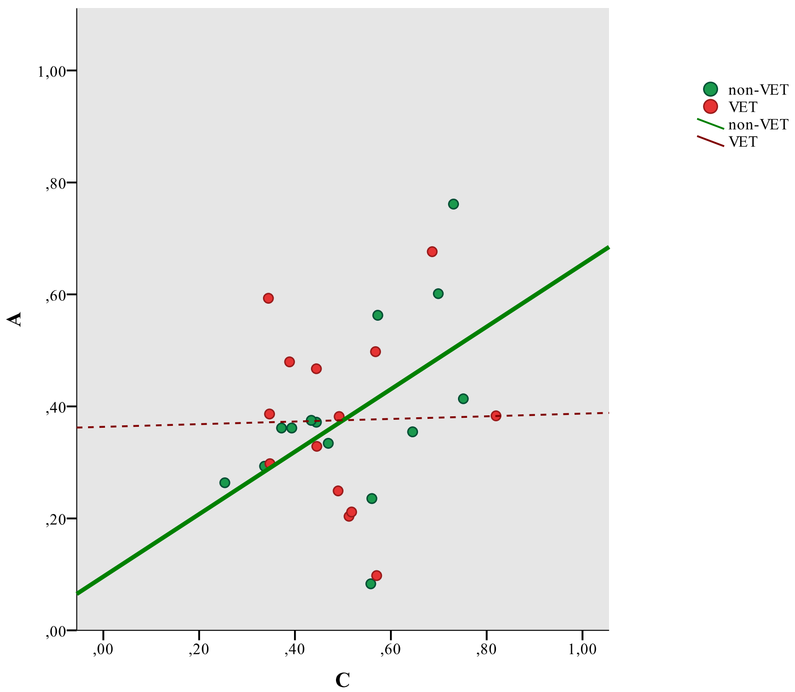

Figure 5.18: Correlation, dimension A vs. C, and grouped for the completion of a VET education. The lines represent a least-square error fit for a linear correlation.

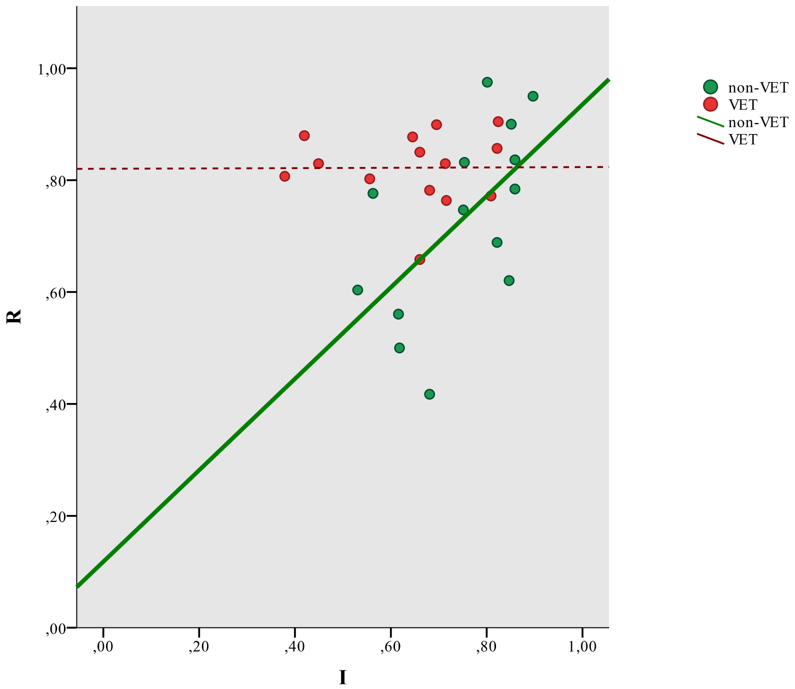

Figure 5.19: Correlation, dimension R vs. I, and grouped for the completion of a VET education. The lines represent a least-square error fit for a linear correlation.

5.4.3 RIASEC model vs. APE

Non-parametric Spearman rank-order correlation tests were calculated to further investigate the meaning of APE against the known RIASEC dimensions. The results are presented in Table 5.37.

The table makes it clear that APE is correlated to the “Realistic” dimension of the Holland categories. The most direct correlation can be seen with the APE-M category, especially APE-MC. Participants who had high scores for APE-M and APE generally had also high scores in the “Realistic” category of the test sets. All other categories showed no correlation to APE.

It would be interesting to compare the “Investigative” category with the superior learning mode (hands-on vs. simulations). Unfortunately, the RIASEC data were only collected from German runs in the second research phase, where simulations were hidden and no differences in student learning were discovered.

5.4.4 RIASEC model vs. test performances (DS-A)

The second research phase allowed for a comparison of the RIASEC model with students’ performances (in sum, after hands-on, after hidden simulations). Non-parametric Spearman rank-order correlation tests were calculated to investigate if RIASEC-values were a precursor for the performance in the knowledge tests. Only one trend was found: RIASEC-C correlated somewhat with the test performance after experimenting with hidden simulations (R6, R9: $\rho = .35$, $p = .096$, $N = 24$). It means students with high value in the “Conventional” category had better results af-

Table 5.37: RIASEC: Spearman rank-order correlation between the RIASEC and APE dimensions

		R	I	A	S	E	C
APE	ρ	* **.42**	.18	.04	.02	.15	.11
	p	.025	.355	.845	.906	.458	.595
APE-E	ρ	.20	.24	-.04	-.08	.03	.07
	p	.297	.211	.854	.696	.890	.717
APE-ED	ρ	.07	.21	.03	-.12	.17	.21
	p	.712	.280	.895	.559	.383	.275
APE-EC	ρ	.13	.09	-.15	-.05	-.12	-.06
	p	.495	.643	.458	.787	.538	.745
APE-M	ρ	* **.40**	.06	.09	.11	.23	.10
	p	.034	.766	.638	.584	.235	.622
APE-MC	ρ	** **.48**	.16	.18	.23	.27	.14
	p	.009	.403	.370	.229	.164	.463
APE-MD	ρ	.13	-.23	-.18	-.20	.12	.09
	p	.494	.232	.362	.298	.541	.658

Note. N = 28. p < .10 marked bold.
* = p < .05. ** = p < .01.
Please see section 4.3 and section 5.3 for details on APE.
Data is based on R9 and partially on R6.

ter hidden simulations than their peers with lower "Conventional" value. Concerning the small number of data sets, the high p-value ($\approx .1$) and the overall inconspicuous "Conventional" category, this trend can be seen as random.

The analysis was repeated separately for non-VET and VET participants (Table 5.35). No significant correlations between the test performances of the second phase and RIASEC were detected for these groups.

In sum, no significant correlations of the RIASEC values with student performances were found.

5.4.5 Summary of results of DS-C

Summary of section 5.4 "Personality/RIASEC (DS-C)"

- AIST and Explorix tests generated similar outcomes, which supports the RIASEC system and the quality of the employed test sets.
- The collected data acknowledged Holland's findings that neighbouring dimensions in the circular RIASEC arrangement are related. This result is valid overall as well as for VET and non-VET groups separately.
- Compared to the published difficulties of the Explorix test manual
  – The electric mobility students had higher "Realistic" and "Investigative" scores than those given by the average item difficulties pub-

lished in the test manual.

- A look at the detailed scales (compared to scales of well-educated young students, derived from the test manual), shows the "Realistic" value of the electric mobility students to be above average. The "Investigative" value was only above average for students without a VET degree, while it was average for former VET-participants.
- Regarding the other four categories (ASEC), the electric mobility students had very similar values as those given by the published difficulties.

- Comparing the VET and non-VET participants
  - VET-graduates tend to have higher values regarding the "Realistic" category than non-VET (trend, $t(26) = -1.94$, $p = .064$, Cohen's $d = -.73$),
  - while their results regarding the "Investigative" category tend to be lower (trend, $t(24) = 2.00$, $p = .057$, Cohen's $d = .76$).
  - In the other categories (ASEC) no significant differences between both groups were detected.
  - Participants without a completed VET education tended to have a high "Realistic" values if they had high "Investigative" values (significant correlation).
  - For former VET participants, this effect was not found, the VET-participants showed generally high "Realistic" values, independent of their "Investigative" score.
- APE seems to be correlated with the "Realistic" category of the RIASEC dimensions. No correlation with the other categories was found.
- No significant correlations between the RIASEC values and student performances were found.

5.5 Survey for participants' subjective opinion after conducting the individual experiments (DS-D)

For DS-D, after performing the experiments, the participants gave feedback on the previously conducted experimental session. In that way *indirect* opinions on the compared learning modes were collected. For the methodology of DS-D see section 4.5.

The student feedback analysis is based on all results of the German Bachelor runs R1, R2, R6 and R9. Student feedback was not collected during the other runs due to time constraints and the absence of the default laboratory survey in these programs. In these runs the second experiment (C*) was conducted directly after the first experiment (B*), with no time for filling out the surveys for each experiment. Thus, the results of this survey are based only on German bachelor students of UAS Ingolstadt (THI). A summary of the findings is presented on page 143.

5.5.1 Return rates

For DS-D, additional data could be collected from a default laboratory survey the university implemented in the new laboratory modules of the study program: THI aimed for iterative improvement of the recently installed laboratory experiments in the study program. Due to low participation in this survey during the first year of its implementation, the leader of the study program began (in 2017) to offer a small incentive for participation. The minimum passing score of 50% of points in the laboratory protocol was reduced to 45% for participation in the online survey of the respective experiment. As a result, the return rate increased, see Table 5.38. Both this survey and the incentive were default for *all* laboratory classes and *independent* of the conducted research.

A Mann-Whitney U-test was performed to check for influences of the differing return rates on the outcomes. The answers regarding the base items (a) to (d) were compared between R1 (low return rate) and R2 (high return rate), separated for the learning modes. It showed no statistically significant differences ($p < .05$).

5.5.2 Phase 1, comparing hands-on and simulated experiments (students were aware of the used mode)

71% of the participants answered that they gained new insights/comprehension while experimenting (a).

41% of the participants stated problems (b) when asked: "At which point in the experiment did you have the biggest problem proceeding with the experiment?".

The median difficulty (c) was rated feasible. The answers separated for both modes are presented in Table 5.39.

When asked if the participants rate the experiment relevant outside the university (d), more than 30% somewhat or entirely agreed, while 14% somewhat or entirely disagreed. The answers, separated by mode, are presented in Table 5.40.

Table 5.38: Survey for participants after experimenting: return rates

Year, Run	Experiment	Filled	Students	Percentage	Percentage
	A	23		58%	
	B	23		58%	
2016, R1	C1	10	40	25%	35%
	C2	9		23%	
	D	5		13%	
	A	23		77%	
	B	24		80%	
2017, R2	C1	20	30	67%	70%
	C2	22		73%	
	D	16		53%	
	A	28		97%	
	B	23		79%	
2018, R6	C1	14	29	48%	66%
	C2	18		62%	
	D	13		45%	
	A	16		62%	
	B	16		62%	
2019, R9	C1	13	26	50%	51%
	C2	11		42%	
	D	10		38%	

Table 5.39: Subjective opinions after laboratory session: answers regarding (c), difficulty

	R1, R2				R6, R9			
(c)	hands-on		simulated		hands-on		hidden simulations	
difficulty	N	percent	N	percent	N	percent	N	percent
easy	25	29.1	20	21.7	12	21.8	13	22.4
feasible	58	67.4	68	73.9	41	74.5	43	74.1
difficult	3	3.5	4	4.3	2	3.6	2	3.4
Total	86	100.0	92	100.0	55	100.0	58	100.0

Table 5.40: Subjective opinions after laboratory session: answers regarding (d), relevance for later professional life/outside university

	R1, R2				R6, R9			
(d)	hands-on		simulated		hands-on		hidden simulations	
relevance	N	percent	N	percent	N	percent	N	percent
fully disagree	1	1.2	3	3.3	1	1.8	0	.0
somewhat disagree	6	7.0	15	16.3	4	7.3	2	3.4
neutral	51	59.3	48	52.2	25	45.5	30	51.7
somewhat agree	20	23.3	17	18.5	15	27.3	15	25.9
fully agree	8	9.3	9	9.8	10	18.2	11	19.0
Total	86	100.0	92	100.0	54	100.0	58	100.0

5.5.2.1 Statistical comparison of hands-on and simulations

A non-parametric Mann-Whitney U-test was performed to compare the answers regarding the items (a) through (d) after experimenting in the hands-on and simulated condition (first research phase). The results are presented in Table 5.41. For items (a) and (b), both binary choices, additionally fourfold tests were conducted. No statistically significant differences were found. Nevertheless, trends were visible:

- (a) In hands-on mode somewhat more students expressed that they had acquired new insights/comprehension (77% vs. 66%; $\widehat{\chi^2} = 2.370$, $p = .124$; $Z = 1.535$, $p = .125$).
- (d) In the feedback form on the experiments, students who conducted the experiments in the hands-on mode rated the execution of the experiments slightly more beneficial for their future professional life ($Z = 1.419$, $p = .156$). After hands-on experiments, 8% somewhat or entirely disagreed with the relevance, while after simulations, a higher share, 19% of the participants disagreed. After hands-on experiments, 33% stated the experimenting somewhat or entirely relevant, while after using simulations, the amount was slightly smaller, 28%, see Table 5.40.

Regarding mentioned problems (b) and difficulty of conduction (c), experiments in both modes were rated similar, which could be expected as experimental procedure and user interface were the same in both modes:

- (b) A similar amount of students in both modes mentioned problems while conducting the experiments (45% vs. 39%; $\widehat{\chi^2} = .705$, $p = .401$; $Z = .837$, $p = .402$).
- (c) Conducting of experiments in both modes was rated as similarly difficult ($Z = -1.116$, $p = .265$).

Table 5.41: R1, R2; Subjective opinions after laboratory session: comparison of participants' opinions after experimenting in different conditions

	hands-on		simulated		Fourfold		Mann-Whitney	
	N	M	N	M	$\widehat{\chi^2}$	p	Z	p
(a) new insights	86	77%	92	66%	2.370	.124	1.535	.125
(b) problems	86	45%	92	39%	.705	.401	.837	.402
(c) difficulty	86		92				-1.116	.265
(d) relevance	86		92				1.419	.156

5.5.2.2 Correlation between items

A Spearman rank-order correlation coefficient was computed to assess the relationship between the answers of all four items (a, b, c, d). The results are presented in Table 5.42.

Correlation between new insights (a) and relevance (d)

A correlation between the relevance for professional life (d) and the presence of new insights while experimenting (a) was detected. It shows that students who claimed that they gained new insights also tended to believe that the execution of the experiment will help them in their future professional life (Spearman's $\rho = .35$, $p < .001$).

That effect was cross-checked by an Spearman rank-order correlation (Table 5.43), which acknowledged that highly significant correlation for both modes separately. If students perceive no new insights while conducting an experiment, they do not judge the execution as beneficial for their life outside university.

Table 5.42: R1, R2; Subjective opinions after laboratory session: Spearman rank-order correlation between items

		(a) new insights	(b) problems mentioned	(c) difficulty
(b) problems	ρ	-.06		
	p	.403		
	N	178		
(c) difficulty	ρ	.07	-.08	
	p	.377	.308	
	N	178	178	
(d) relevance	ρ	** **.35**	-.12	-.09
	p	.001	.118	.253
	N	178	178	178

Note. $p < .10$ marked bold. ** = $p < .01$.

5.5.2.3 Correlations between items identified for individual modes

To find individual correlations, existing in one of the modes only, the rank-order analysis was repeated for the hands-on and the simulations condition separately. The results are presented in Table 5.43. Trends are reported in bold font. Besides the general correlation between (a) and (d) noted above, two additional trends were found for the *hands-on* condition :

- First, if a *hands-on* experiment was deemed difficult (c), problems (b) were mentioned less often (Spearman's $\rho = -.20$, $p = .065$, $N = 86$). A more in-depth look at the answers from the students reveals no correlation between these two items, a common mentioned problem was the lack of available measurement devices.

- Secondly, students who (c) deemed a *hands-on* experiment difficult, tended to not (d) consider it beneficial for their future professional life (Spearman's $\rho = -.22$, $p = .046$, $N = 86$).

Table 5.43: R1, R2; Subjective opinions after laboratory session: Spearman rank-order correlation between items. top right: after hands-on. bottom left: after simulations

hands-on / sim.		(a) new insights	(b) problems mentioned	(c) difficulty	(d) relevance
(a) new insights	ρ		.00	.07	* **.27**
	p		.972	.527	.013
	N		86	86	86
(b) problems	ρ	-.14		† **-.20**	-.07
	p	.199		.065	.520
	N	92		86	86
(c) difficulty	ρ	.09	.06		* **-.22**
	p	.416	.590		.046
	N	92	92		86
(d) relevance	ρ	** **.40**	-.17	.04	
	p	.001	.107	.729	
	N	92	92	92	

Note. $p < .10$ marked bold. † = $p < .10$. * = $p < .05$. ** = $p < .01$.

5.5.2.4 The difference of correlations between modes

In the last section, correlations between the items for both modes were examined individually. Dependent on the learning mode, the correlations had different strength. This section investigates if this difference was statistically significant in order to gain insight into students' relation to the learning modes.

To test whether the correlations differed statistically significantly between both modes, Equation I.13 was used. The results are presented in Table 5.44.

For the correlations, which were detected in hands-on but not in simulations, a somewhat statistically supported difference between the learning modes was found (b-c: $p = .088$; c-d: $p = .092$). In both cases, the correlation after simulated experiments was absent ($\rho \approx .05$), while a weak correlation was measured after hands-on experiments ($\rho \approx .2$, $p < .1$).

No other pair showed significant differences between both modes, including the correlation between the relevance for professional life (d) and the stated new insights (a) while experimenting (a-d: $p = .342$).

5.5.2.5 Moodle feedback vs. test performance and influence of Vocational Education and Training

The survey used in R1 and R2 did not allow for establishing correlations between individual learning success and students' feedback, as the feedback form did not ask for the self-created code-word that was used in the knowledge tests. After an ethics amendment for the following runs, the questionnaire was updated to include the information.

Table 5.44: R1, R2; Subjective opinions after laboratory session: Statistical significance of the difference of the Spearman rank-order correlations between modes

			hands-on		simulated				
			ρ	N	ρ	N		z	p
(b)	-	**(c)**	-.20	86	.06	92	†	-1.704	.088
(c)	-	**(d)**	-.22	86	.04	92	†	-1.684	.092
(a)	-	(d)	.27	86	.40	92		-.950	.342
(a)	-	(b)	.00	86	-.14	92		.917	.359
(b)	-	(d)	-.07	86	-.17	92		.660	.509
(a)	-	(c)	.07	86	.09	92		-.111	.912

Note. † = $p < .10$.

5.5.3 Phase 2, comparing hands-on experiments with hidden simulations (students were not aware of the used mode)

In the second phase of the research, during runs R6 and R9, 89% of the participants answered that they gained new insights/comprehension while experimenting (a), which is a higher rate than in the first phase.

37% of the participants stated problems (b) when asked “At which point in the experiment did you have the biggest problem proceeding with the experiment?”, which is approximately the same rate as in the first phase.

The median difficulty (c) was rated similar number to the first phase. The answers, separated by mode, are presented in Table 5.39.

When asked if the participants rate the experiment relevant (d) outside the university, more than 44% somewhat or entirely agreed, while 6% somewhat or entirely disagreed. Compared to the first phase, this is an improved response. The answers, separated by mode, are presented in Table 5.40.

5.5.3.1 Statistical comparison of hands-on and hidden simulations

A non-parametric Mann-Whitney U-test was conducted to compare the answers to the items (a) through (d) after experiments in the hands-on and simulated condition (second research phase). For items (a) and (b), both binary choices, additionally four-fold tests were conducted. The results are presented in Table 5.45. No statistically significant differences were found. Nevertheless, one trend was visible:

- (b) In the hidden simulations mode, slightly more problems during the experiment were mentioned (31% vs. 45%; $\widehat{\chi^2} = 2.320$, $p = .128$; $Z = -1.516$, $p = .129$).

The other points were rated similar in both modes:

- (a) In both modes, the students expressed that they had acquired similar amounts of new insights/comprehension (87% vs. 93%; $\widehat{\chi^2} = 1.092$, $p = .296$; $Z = -1.040$, $p = .298$).

- (c) The engagement with the experiments in both modes was stated equally difficult (Z = .087, p = .931).
- (d) Students gave the execution of the experiments nearly identical ratings regarding the benefit for their future professional life (Z = -.301 p = .764).

Table 5.45: R6, R9; Subjective opinions after laboratory session: comparison of participants' opinions after experimenting in different conditions

	hands-on		simulated		Fourfold		Mann-Whitney	
	N	M	N	M	$\widehat{\chi^2}$	p	Z	p
(a) new insights	55	87%	58	93%	1.092	.296	-1.040	.298
(b) problems	55	31%	58	45%	2.320	.128	-1.516	.129
(c) difficulty	55		58				.087	.931
(d) relevance	55		58				-.301	.764

To investigate the reasons for the slight increase in problems encountered during hidden simulations, the answers were analysed in detail: 60% of problems concerned the use of the GUI to control the battery tester, 11% concerned hardware problems (like cooling speed or defective battery cells), and 29% were problems regarding calculations, claiming the laboratory evaluation is too time-consuming, or about bad teamwork and inter-teamwork (for example data exchange of weird data). As the GUI and evaluation were identical and hardware problems were only possible when students used real hands-on equipment, the difference regarding (b) between modes can be disregarded. It was not statistically significant (p » .05) anyhow.

This being said, no relevant differences between both modes were found.

5.5.3.2 Correlation between items

A Spearman rank-order correlation coefficient was computed to assess the relationship between the answers of all four items (a, b, c, d). The results are presented in Table 5.46.

Correlation between new insights (a) and relevance (d)

Like in the first phase, a correlation between the relevance for professional life (d) and the presence of new insights while experimenting (a) was detected.

It confirms that students who claimed that they gained new insights also tend to believe that the execution of the experiment will help them in their future professional life (Spearman's ρ = .20, p = .031).

5.5.3.3 Correlations between items analysed for individual modes

To find differences between the modes, the Spearman rank-order analysis was repeated for the hands-on and the hidden simulations (perceived as hands-on) condition separately. The results are presented in Table 5.47. Trends are reported in bold font.

Table 5.46: R6, R9; Subjective opinions after laboratory session: Spearman rank-order correlation between items

		(a) new insights	(b) problems men.	(c) difficulty	(d) relevance
(b) problems	ρ	-.12			
	p	.192			
	N	121			
(c) difficulty	ρ	-.05	.04		
	p	.571	.653		
	N	121	121		
(d) relevance	ρ	* **.20**	.09	-.07	
	p	.031	.345	.439	
	N	121	121	121	
test performance	ρ	* **.20**	.07	-.13	** **.36**
	p	.035	.447	.174	.001
	N	112	112	112	112

Note. p < .10 marked bold. * = p < .05. ** = p < .01.

For the hands-on condition, an additional significant correlation was found: If, after *hands-on* experiments, participants claimed to gained new insights (a), problems (b) were mentioned less often (ρ = -.33, p = .013, N = 55). Once again, this correlation can be disregarded. As mentioned above, a more detailed look at students' answers makes it clear that the stated problems were caused by the use of the GUI and the complexity of test sequences, which were identical in both modes.

Table 5.47: R6, R9; Subjective opinions after laboratory session: Spearman rank-order correlation between items. top right: after hands-on. bottom left: after hidden simulations

hands-on / hidden sim.		(a) new insights	(b) problems men	(c) difficulty	(d) relevance	test perf.
(a) new insights	ρ		* **-.33**	-.14	.18	.21
	p		.013	.323	.190	.137
	N		55	55	55	54
(b) problems	ρ	-.03		-.01	.15	.14
	p	.833		.916	.286	.297
	N	58		55	55	54
(c) difficulty	ρ	.18	.13		-.04	-.15
	p	.174	.330		.763	.268
	N	58	58		55	54
(d) relevance	ρ	.18	-.01	-.02		** **.35**
	p	.167	.960	.859		.010
	N	58	58	58		54
test performance	ρ	.17	-.02	-.11	** **.37**	
	p	.206	.872	.426	.004	
	N	58	58	58	58	

Note. p < .10 marked bold. * = p < .05. ** = p < .01.

5.5.3.4 The difference of correlations between modes

Again, differences between the strengths of the correlations of both modes were analysed for the results of the second phase (Table 5.48). One relevant pair was found (a-c: $p = .098 < .1$), which would mean that in hands-on mode experiments rated as more difficult (c) lead to fewer perceived new insights (a), while in the hidden simulations less difficult rated experiments lead to comparatively more insights. As the p-value reaches nearly .1 and both base correlations (a-c, for hands-on and hidden simulations) show no significance (Table 5.47), that outcome can be considered a statistical effect and disregarded.

Table 5.48: R6, R9; Subjective opinions after laboratory session: Statistical significance of the difference of the Spearman rank-order correlations between modes

			hands-on		hidden simulations				
			ρ	N	ρ	N		z	p
(a)	-	**(c)**	-.14	55	.18	58	†	-1.654	.098
(a)	-	(b)	-.33	55	-.03	58		-1.617	.106
(b)	-	test perf.	.14	54	-.02	58		.861	.389
(b)	-	(d)	.15	55	-.01	58		.799	.424
(b)	-	(c)	-.01	55	.13	58		-.754	.451
(c)	-	test perf.	-.15	54	-.11	58		-.246	.806
(a)	-	test perf.	.21	54	.17	58		.195	.846
(d)	-	test perf.	.35	54	.37	58		-.133	.894
(c)	-	(d)	-.04	55	-.02	58		-.092	.927
(a)	-	(d)	.18	55	.18	58		-.025	.980

Note. † = $p < .10$.

5.5.3.5 Test performance

The survey used in R1 and R2 (first phase) did not allow for establishing correlations between individual learning success and student's feedback, as the feedback form did not include the self-created code-word used in the knowledge tests. After an ethics amendment for the following runs, the questionnaire was updated and included the code-word (R6, R9).

In the second phase, students who considered a specific experiment relevant for life outside university performed better in the regarding knowledge test ((d)-test performance: $\rho = .36$, $p < .001$, $N = 112$, last line of Table 5.47). This dependency is shown in Figure 5.20.

Students who stated new insights during experimentation (a), also performed better in the tests ((a)-test performance: $\rho = .20$, $p = .035$, $N = 112$).

Data in Table 5.48 shows that both correlations ((a)-tp, (d)-tp) were nearly identical for both modes; no significant difference in the effects between the two modes (hidden simulations, hands-on) could be found.

Additionally, a Spearman rank order correlation calculation was conducted to

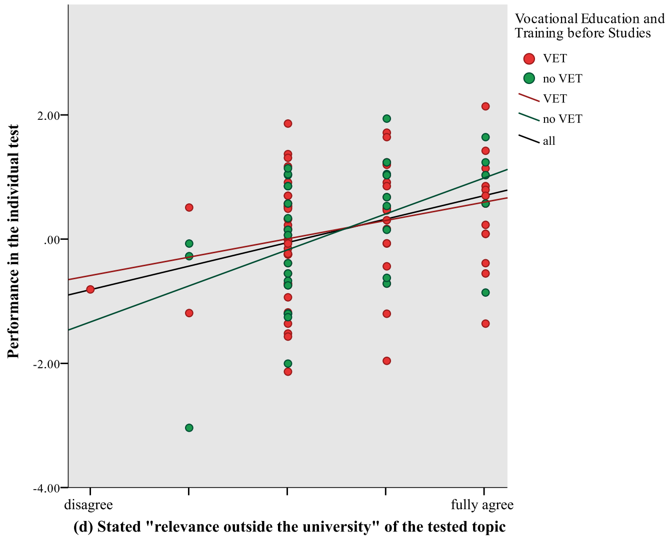

Figure 5.20: Phase 2, R6, R9: Test performance in a test vs. perceived relevance of the respective experiment. Students who perceived an experiment to be less relevant performed lower in the respective test. The lines represent a linear fitted trend line, assuming a linear distribution from fully disagree to fully agree.

investigate the relation between the superior learning mode of a student (hidden simulations vs. hands-on) and the items (a) to (d). It did not show a single statistically significant result ($p < .05$), which supports the equivalent perception of both modes in research phase two.

5.5.3.6 Vocational Education and Training (VET)

Beginning with run R6, the use of a code-word allowed also for comparison regarding a completed VET program before studies.

A non-parametric Mann-Whitney U-test was performed for both modes together, and both modes separately; the result is shown in Table 5.49. No significant differences in answers were found.

5.5.3.7 Amount of Practical Experience (DS-B)

In the second research phase, keywords were available to match the subjective moodle feedback to specific participants' APE. To be able to correlate a student's subjective responses with their characteristics, their average response (available responses on the experiments) to items (a) to (d) was calculated. A Spearman rank order correlation analysis was performed to find correlations between the items and students' *Amount of Practical Experience* (DS-B). The results are presented in Table 5.50. Two major and statistically significant correlations were detected:

Students who had a high APE-E more often reported having gained new insights (APE-E vs. (a); Spearman's $\rho = .345$, $p = .025$, $N = 42$). This trend was explicitly caused by a high APE-ED (APE-ED vs. (a); Spearman's $\rho = .43$, $p = .005$, $N = 42$).

Students who had a high APE more often considered the experiments relevant

Table 5.49: Moodle-Feedback, R6 and R9: Answers compared for VET / non-VET

		non-VET		VET		Mann-Whitney	
		N	M	N	M	Z	p
(a) new insights	both	38	89%	81	89%	.095	.924
	hands-on	17	88%	38	87%	.142	.887
	hid. sim.	17	94%	41	93%	.195	.846
(b) problems	both	38	39%	81	37%	.254	.799
	hands-on	17	29%	38	32%	-.159	.873
	hid. sim.	17	53%	41	41%	.793	.428
(c) difficulty	both	38		81		1.070	.285
	hands-on	17		38		.132	.895
	hid. sim.	17		41		.840	.401
(d) relevance	both	38		81		-.876	.381
	hands-on	17		38		.214	.831
	hid. sim.	17		41		-1.344	.179

Note. † = p < .10. * = p < .05. ** = p < .01.
As some feedback could not assigned to a mode (e.g. wrong code-word), "both" covers more feedback as the modes in sum.

outside of a university setting (APE vs. (d); Spearman's $\rho = .41$, $p = .008$, $N = 42$). This trend was caused specifically by a high APE-E (APE-E vs. (d); Spearman's $\rho = .43$, $p = .005$, $N = 42$), and APE-ED (APE-ED vs. (a); Spearman's $\rho = .43$, $p = .005$, $N = 42$).

Table 5.50: R6, R9; Subjective opinions after laboratory session: Spearman rank-order correlation with APE

		APE	APE-ED	APE-EC	APE-M	APE-MC	APE-MD
(a) new insights	ρ	.294	.429	.249	.264	.263	.036
	p	.062	.005	.112	.091	.092	.821
		*	**				
(b) problems	ρ	-.168	-.123	-.206	-.228	-.247	-.217
	p	.293	.437	.191	.146	.115	.168
(c) difficulty	ρ	-.272	-.297	-.230	-.283	-.256	-.274
	p	.085	.056	.143	.069	.101	.080
(d) relevance	ρ	.408	.428	.309	.245	.308	-.036
	p	.008	.005	.046	.118	.047	.822
		**	**	*		*	

Note. N = 42; Note. * = p < .05. ** = p < .01. APE-dimensions are described on page 102.

5.5.3.8 RIASEC (DS-C)

The same investigation was conducted to correlate the *RIASEC* dimensions with the moodle responses (a) to (d). The results are presented in Table 5.51. One statistically significant correlation was detected:

Student who had a high *Investigative*-value were more likely to judge the experiments as relevant outside of a university setting (I vs. (d); Spearman's $\rho = .48$, $p = .043$, $N = 18$).

Table 5.51: R6, R9; Subjective opinions after laboratory session: Spearman rank-order correlation with RIASEC

		R	I	A	S	E	C
(a) new insights	ρ	-.145	.380	.025	.128	-.091	-.034
	p	.565	.120	.923	.612	.719	.892
(b) problems	ρ	.027	-.067	.281	.210	.005	.025
	p	.914	.793	.259	.404	.983	.921
(c) difficulty	ρ	-.330	-.219	.219	.015	-.088	-.289
	p	.181	.383	.383	.954	.727	.245
(d) relevance	ρ	.027	.481	.187	.407	.086	.167
	p	.915	.043 *	.457	.093	.734	.507

Note. N = 18; Note. * = p < .05.

5.5.4 Comparing responses in both research phases

To detect whether students perceived the hands-on mode in the second research phase as the hands-on mode of the first phase, a non-parametric Mann-Whitney U-test was performed. Differences would give indication that the participants or the experiments had changed over time.

71% of the participants in the first research phase vs. 89% in the second research phase answered that they gained new insights/comprehension (a) while experimenting. No statistically significant differences between hands-on experiments in the first and the second phase were found ($Z = -1.544$, $p = .123$).

41% of the participants in the first research phase vs. 37% in the second research phase stated problems (b). A slight, and statistically not significant trend of fewer problems in the second phase was detected ($Z = 1.703$, $p = .089$), which could be caused by the improvement of experiments during the runs.

Regarding stated difficulty (c), data in Table 5.39 shows a very similar distribution of answers in both phases after hands-on experiments. This was confirmed by a U-test. No statistically significant differences were found between hands-on experiments in the first and the second phase ($Z = 0.891$, $p = .373$).

Table 5.40 shows a very similar data distribution in both phases with regards to relevance (d). No statistically significant differences were found between hands-on experiments in the first and the second phase ($Z = -1.414$, $p = .157$).

Due to the missing information about VET during R1, and the missing code-word in the moodle questionnaire for R2, unfortunately a comparison between the phases with regards to VET info was not possible.

5.5.5 Summary of results of DS-D

Summary of section 5.5 "Subjective opinions after conducting the experiments (DS-D)"

Each week, after experimenting in one of both modes, the students were asked to participate in an online survey. It was located on the THI intranet. The feedback was not mandatory, but the students were used to the existence of the form from their study program's other laboratories. It was anonymous, but as the subgroup of the run was recorded, the answers could be analysed according to the respective learning modes.

In the first phase, when comparing the hands-on experience with simulated experiments, the following was found:

- The difficulty of the experiments conduction was, on average, considered feasible, and no significant differences between the two modes were detected.
- A similar number of problems was reported after the experiments in both modes.
- A high share of participants perceived a gain of new insights through the laboratory experiments. After hands-on experimenting, a higher share of students claimed to have acquired new insights/comprehension than after simulations (77% vs. 66%), but the difference was not statistically significant ($p = .125$).
- When asked if the experiments are relevant outside the university, in the first phase, more than 30% of the participants agreed, while only 14% disagreed.

  After hands-on experiments, the execution of the experiment was rated slightly more beneficial for future professional life on average, but that difference was not statistically significant either ($p = .156$). Additionally, the share of participants who perceived no benefit for life outside the university was much higher after simulations (19.6%) than after hands-on experiments (8.2%).
- The more *hands-on* experiments were perceived as difficult in the first phase, the less the execution of the experiment was considered relevant (very weak correlation, statistically significant). The difference in strength compared to the same correlation regarding simulated experiments (where that effect was absent) was also statistically significant. As that effect for the hands-on mode should have been acknowledged in the second phase,

and it was not, the reader needs to be careful – it might have been based on random statistical effects in the first phase.

- When the execution of an experiment was rated to give new insights, it was also understood as helpful for later professional life (weak correlation, statistically significant). That weak trend was independent of the mode, and observed in both research phases. In the first phase, the correlation was more intensive for simulated experiments, but the difference between hands-on and simulations was not statistically significant.

- In the second phase, the comparison with knowledge test results was possible and showed a slight correlation ($\rho = .36$) between perceived relevance and test performance (Figure 5.20), which was statistically significant ($p < .001$). Students who rated experiments as relevant tended to perform better in the respective test. The difference between VET and non-VET students was not significant, it was a general trend.

- Students who stated that an experiment delivered new insights tended to produce better test results in the correlated test ($\rho = .20$, $p = .035$). No differences were found between VET and non-VET students.

Starting with the second phase,

- when comparing two modes which were both perceived as hands-on experiments, no differences was found in the student response. None of the employed items showed significant differences between the modes. Both modes led to the same opinions about the experiment and were perceived as identical.

- the overall distribution and the percentage of students who gave a low rating for relevance was similar to the first phase (9.1%) for real hands-on. With hidden simulation, it was only 3.4%, which is a big contrast to overt simulations. This provides another indicator that hidden simulations were effectively perceived as hands-on experiments.

- the share of students claiming to have gained new insights increased, but did not differ between modes (87% vs. 93%, $p = .298$).

- the items were also evaluated with regards to the completion of a German VET before enrolment, and no differences were identified. That is not surprising, as generally no differences between modes of the second phase were identified. Thus, this outcome does not allow for the assumption that these differences would not have been found when comparing perceived simulations to perceived hands-on experiments in the first phase, while recording VET info in the Moodle feedback.

- the experiments were more likely to be considered relevant by participants who had a high Amount of Practical Experience (specifically caused by the Electronics-Do dimension) and by participants who had a high Investigative RIASEC value.
- students who had a high APE-E (specifically caused by the Electronics-Do dimension) stated more often to have gained new insights.

5.6 Survey for participants' and other persons' subjective opinions on the learning modes (DS-E)

To collect a broader view of opinion on the learning modes, a second online questionnaire was developed. This survey included students and employees of other universities in Germany and foreign countries, which did not conduct the laboratory experiments. For the methodology of DS-E see section 4.6. A summary of the findings is presented on page 162.

5.6.1 Participants

Several universities were contacted and invited to join the study. The system collected 285 responses. 55% of the participants completed their questionnaires (filled the form to the end, eventually skipped some items). Data of the partially completed questionnaires was considered in the analysis as long the data for the targeted analysis necessary was fully available. The data sets of the partially completed questionnaires were considered as long all items for the specifically targeted analysis were available.

Many of the participants were undergraduates (34%). 33% had finished bachelors, while 36% had finished masters or the German diploma. Only 6% had a doctoral degree.

On average, participants had 4 years of professional experience. Naturally, the standard deviation was big (4.8 years). A high share of participants were students (45% Bachelors, 34% Masters, 7% PhD).

A significant amount of participants came from electrical engineering (48%), and mechanical engineering (23%). Only 1.6% of participants were not from the engineering field.

The statistics for the participating universities are presented in Table 5.52. The highest share of data (70%) was returning from German Universities of Applied Sciences and German "traditional/full" universities. Even when more universities were asked to participate, the return rate of Universities of Applied Sciences in Germany was higher. As a result, only 22.5% of the German answers came from traditional universities. Considering the non-German schools as traditional universities, 42% of data returned from traditional universities.

Among German students 30% of participants had completed a VET education, while 70% had not.

25% of the data originated from participants of the main study. Nevertheless, only 8% of the respondents provided code-words. This meant that the data could not be linked to the main research.

5.6.2 General outcomes

5.6.2.1 Paired single choice questions

A Wilcoxon signed ranks test was computed to compare the pairs of statements on both modes. For all pairs of questions, a highly statistically significant and strong trend towards favouring the hands-on mode was found (all $p < .001$, $Z > 6$). The detailed results are presented in Table 5.53.

In sum, the results suggest that hands-on experiments are favoured by the participants of the questionnaire. Specifically, participants of this part of the study

- expected to learn better with hands-on experiments compared to simulations (Pair 1, $d = .96$).
- believed in a more authentic laboratory experience with hands-on experiments (Pair 2, $d = 1.17$).
- thought the outcomes of hands-on experiments are more accurate (Pair 3, $d = .75$).
- expressed the opinion that others (the majority of students) also believe to benefit from hands-on experiments most (Pair 4, $d = .93$).
- claimed they would visit hands-on laboratories more often, compared to simulated experimental lessons, if visiting was optional (Pair 5, $d = 1.05$).

5.6.2.2 Single choice questions

Participants agreed in average both to Q1 "I always simulate technical problems by myself in order to understand them better and/or to check my assumptions." ($M = \mathit{4.65}^{s}$, $Mdn = 5$) as well as Q2 "Students use simulations in their studies far too often, instead of trying things out." ($M = \mathit{4.46}^{s}$, $Mdn = 5$) slightly (see Table 5.54). A one-sample Wilcoxon signed rank test was performed to check for a statistically significant deviation to the neutral value (4). Both trends were statistically significant (Q1: $Z = 7.490$, $p < .001$; Q2: $Z = 5.715$, $p = .001$)).

A Spearman's rank correlation coefficient was computed to assess the relationship between both items (Q1, Q2); and a statistically significant, but weak, correlation between both variables was found ($\rho = .27$, $p < .001$). The correlation was rated irrelevant by analysing a scatter plot, which showed a widespread field with nearly all combinations of both items.

When directly asked which mode the participants would recommend (Q3), more than 70% of the participants expressed a preference for the hands-on mode, even when allowing for the possibility to choose a remote laboratory in simulations (Table 5.55). Four percent of the participants expressed none of the default options and stated their preference in the free text. All of them asked for a combination of simulations and hands-on experiments.

Table 5.52: Survey for participants and other persons, participating universities

Country	University	Frequency	Percent
Australia	RMIT Melbourne	11	4.8
Denmark	Aarhus University	10	4.4
	University of Southern Denmark	4	1.8
Germany	Aalen University (trad. Uni.)	5	2.2
	Bayreuth (trad. Uni.)	7	3.1
	Chemnitz (trad. Uni.)	23	10.1
	Stuttgart (trad. Uni.)	1	.4
	Hamburg (UAS)	19	8.3
	Ingolstadt (UAS)	100	43.9
	München (UAS)	2	.9
	Ostfalia (UAS)	1	.4
	Stuttgart (DHBW)	1	.4
	Westsächsische Hochschule Zwickau (UAS)	1	.4
Italy	Politecnico di Milano	16	7.0
Spain	Universidade da Coruña/Ferrol	7	3.1
Tunesia	ENSTAB / Université de Carthage	12	5.3
	Other	8	3.4
	Total	228	100.0

Table 5.53: Subjective opinions: Wilcoxon signed ranks test, comparison of the opinions on both modes

	Hands-on			Simulations			Wilcoxon		
	N	M	SD	N	M	SD	Z		p
Pair 1	162	^{s}*5.93*	*1.12*	162	^{s}*4.70*	*1.42*	7.365	**	.000
Pair 2	161	^{s}*5.89*	*1.12*	161	^{s}*4.30*	*1.57*	8.493	**	.000
Pair 3	161	^{s}*5.73*	*1.13*	161	^{s}*4.81*	*1.30*	6.518	**	.000
Pair 4	161	^{s}*5.56*	*1.24*	161	^{s}*4.32*	*1.42*	7.564	**	.000
Pair 5	163	^{s}*5.99*	*1.28*	163	^{s}*4.45*	*1.63*	7.626	**	.000

Note. ** = p < .01. Z pos. = better opinions regarding hands-on.
All questions (see page 67) were coded in a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree (4 = neutral).
^{s} = Shapiro-Wilk p < .05, no normal distribution.

Table 5.54: Subjective opinions: descriptive statistics of Q1, Q2 and results of a one-sample Wilcoxon signed rank test, compared to the neutral value (4)

	Q1	Q2
N	167	167
Mean	^{s}*4.65*	^{s}*4.46*
Std. deviation	*1.66*	*1.61*
Median	5	5
Range	6	6
Minimum	1	1
Maximum	7	7
Z	7.490	5.715
p	.000	.001

Note. ** = p < .01. Z pos. = agreement.
All questions (see page 67) were coded in a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree (4 = neutral).
^{s} = Shapiro-Wilk p < .05, no normal distribution.

A non-parametric Mann-Whitney U-test was used to determine if the number of answers in the category "Hands-on, in small working teams, at a specific time in the laboratory" compared to the amount of the rest of the options differed between participants with (80% preferred hands-on in laboratories) and without VET education (72% preferred hands-on in laboratories). No significant difference was found ($Z = -0.614$, $p = .539$).

5.6.2.3 Free text questions

Direct questions were employed to determine the main aspect/opinion of the participants on both modes. The participants were asked for their general opinion on both modes individually.

Simulations

Regarding simulations, the participants of the questionnaire submitted negative and positive comments. The answers were evaluated and grouped:

- Negative comments (36×)
  - Doubts on reality (25×)
  - Faster execution of often pre-set experiments may hinder understanding (3×)
  - Doubts about learning all necessary aspects for working in real laboratories (3×)
  - General negative opinions (2×)
  - General doubts, as special knowledge to use simulations is necessary (2×)
  - Claims that simulations cannot replace real practice (1×)
- Neutral comments (20×)
  - Asking for combination with hands-on (11×)
    - Combination of both is the best variant (no chronological order recommended) (5×)
    - Recommendation to use hands-on after theoretical knowledge was gained from simulations (4×)
    - Recommendation to use simulations to deepen knowledge after hands-on experiments (2×)
  - Simulations are appreciated, but should not fully replace hands-on experiments (2×)
  - No final trend or opinion in the answer, or discussing both advantages and disadvantages (7×)
- Positive comments (49×)

    - General positive opinions (19×)
    - Simulations deliver a fast overview/basic concept for preparation (9×)
    - More efficient - time / energy / cost (7×)
    - Less dangerous, or a good replacement for a similar reason (5×)
    - Allows learning by trial and error (4×)
    - Allows a focus on the essential result (3×)
    - Good for deepening the understanding (2×)

Hands-on Experiments

Regarding hands-on experiments, the participants of the questionnaire submitted negative and positive comments on the learning mode. The answers were evaluated and grouped for the opinion.

- Negative comments (10×)
    - Less efficient in terms of time/energy/cost/effort (10×)
- Neutral comments (3×)
    - No trend in the answer (2×)
    - Conducting dangerous experiments needs combination with simulations (1×)
- Positive comments (90×)
    - General positive opinions (36×)
    - Hands-on experimentation is authentic, includes nature and real life (18×)
    - Leads to deeper understanding (13×)
    - Practical learning instead of theory (8×)
    - Working/learning in a multi-sensorial way (5×)
    - Good for teaching measurement deviations etc. (3×)
    - Boost the confidence of students (3×)
    - Introduction to scientific work (2×)
    - More interesting (2×)

5.6.3 The trend between participant characteristics

As the general trend tended towards advantages of hands-on laboratories, further statistical tests were conducted to compare the strength of opinion towards hands-on laboratories between several participant characteristics.

Paired single choice questions

To allow for the comparison, the trend of each paired question was determined. To do so, the differences between the coded response of hands-on and the coded response of simulations were calculated. The resulting value ranged between -6 and 6 and described the opinion of the participant in contrast between both modes. A positive number describes a participant with positive opinion towards the hands-on mode (as in, the participant agreed to the statements favouring the hands-on mode more those favouring simulations), while a negative number would describe a participant favouring the simulated laboratory mode.

Single choice questions

The answers regarding Q1 and Q2 were further evaluated using Mann-Whitney U-tests/independent-samples t-tests; the data were grouped for the respective participant characteristics.

5.6.3.1 Status student/graduate

Students as well as persons with a finished higher education (in the following referred to as *graduates*) filled out the questionnaire. A Mann-Whitney U-test was conducted to compare both groups; and the results are shown in Table 5.56.

Paired single choice questions

A statistically significant difference was found in the trends between both modes of both groups regarding question pairs 2, 3, 4, and 5. Both students and graduates endorsed hands-on stronger than simulations ($M > 0$), but graduates showed a stronger trend. Only pair 1 showed no statistical supported difference ($p > .1$) between the trends of both groups. These results suggest that graduates see more advantages of hands-on laboratories (when compared to simulations) than persons who are still in university.

Specifically, the results suggest a stronger endorsement of hands-on laboratories by graduates than students regarding:

- the authentic laboratory experience (pair 2)
- accurate learning of the real behaviour of a specimen (pair 3) and
- the superiority of the laboratory mode to study the real behaviour of a specimen (pair 4).

Nevertheless, the trend of the other pairs also went towards a stronger opinion of graduates than of students.

To investigate if the trends were caused by deviating opinions on hands-on or simulations, Mann-Whitney U-tests were computed for all five items, both for the answers regarding hands-on and simulations. The results are also presented in Table 5.56. The table shows that the differences in the trends between the groups were mainly (except pair 3) caused by a difference in the opinions regarding simulations, as effect sizes here were generally bigger.

Single choice questions

Also, the answers to Q1 and Q2 were compared regarding differences between students and graduates. The results are presented in the two last lines of Table 5.56.

The answers regarding Q1 ("I simulate technical problems by myself in order to understand them better") and Q2 ("Students use simulations in their studies too often") showed no statistically significant trends.

5.6.3.2 Country

In the questionnaire, participants from different countries took part. A Mann-Whitney U-test was conducted to compare German participants to internationals; the results are shown in Table 5.57. No statistically significant differences were identified.

5.6.3.3 Academic degree / postgraduate vs. undergraduate

In the questionnaire, participants with and without academic degree contributed data.

Paired single choice questions

A Mann-Whitney U-test was conducted to compare the preferences of both groups; the results are shown in Table 5.58. There was no significant difference in the answers to the question pairs detected. These results suggest that both groups have similarly strong opinions to the laboratory modes.

Single choice questions

Nevertheless, the Mann-Whitney U-test regarding Q1 and Q2 resulted in differences between the compared groups.

The test showed a significant difference in the scores for the frequency of usage of simulations of undergraduates (Q1: N = 51, M = 4.12^s) and postgraduates (Q1: N = 112, M = 4.85^s); Z = -2.485, p = .013. The results suggest a much more frequent usage of simulations among postgraduates than among undergraduates.

A significant difference was also found between both groups in the opinion that students use simulations too often (instead of trying things out). The difference was statistically significant (Z = -2.203, p = .028). Specifically, undergraduates had neutral opinion (N = 51, M = 4.04^s) while postgraduates slightly agreed to that statement (N = 112, M = 4.63^s).

5.6.3.4 Practical experience

For the found differences between students and participants after studies, a Spearman's rank correlation coefficient was computed to assess the relationship between all items and the self-stated years of practical experience of the participants. No

Table 5.55: Subjective opinions: Frequencies of answers on question Q3 “Which learning mode for laboratories would you recommend for university students?”

Preferred laboratory mode	Frequency	Percent
Hands-on, in small working teams, at a specific time in the laboratory	120	73.2
Simulations, in small working teams, at a specific time, at the university	28	17.1
Simulations, alone, when I want, online (remote, from home or any place)	9	5.5
Free text response, asking for a combination of hands-on and simulations,		
- without any specification about time and locality	4	2.4
- in small working teams, at a specific time in the lab/university	3	1.8
Total	164	100.0

Table 5.56: Subjective opinions: non-parametric Mann-Whitney U-test, students' opinions compared to the opinions of participants who had already completed their studies

	Student			After Studies			Mann-Whitney		
	N	M	SD	N	M	SD	Z		p
Pair 1	115	^{s}*1.23*	*1.61*	26	1.77	2.42	-1.743	†	.081
hands-on	116	^{s}*5.92*	*1.06*	27	^{s}*5.93*	*1.24*	.353		.724
sim.	115	^{s}*4.69*	*1.34*	26	4.15	1.67	1.752	†	.080
Pair 2	115	^{s}*1.51*	*1.72*	26	2.62	1.77	-2.834	**	.005
hands-on	117	^{s}*5.86*	*1.01*	27	^{s}*6.15*	*.99*	-1.619		.106
sim.	115	^{s}*4.36*	*1.49*	26	^{s}*3.58*	*1.39*	2.446	*	.014
Pair 3	114	^{s}*.75*	*1.48*	26	^{s}*1.81*	*1.86*	-2.436	*	.015
hands-on	115	^{s}*5.57*	*1.09*	27	^{s}*6.11*	*.93*	-2.388	*	.017
sim.	116	^{s}*4.82*	*1.19*	26	^{s}*4.35*	*1.60*	1.338		.181
Pair 4	114	^{s}*1.07*	*1.74*	26	2.15	1.95	-2.473	*	.013
hands-on	115	^{s}*5.50*	*1.22*	27	^{s}*5.74*	*1.02*	-.769		.442
sim.	115	^{s}*4.43*	*1.37*	26	3.62	1.50	2.420	*	.016
Pair 5	115	^{s}*1.44*	*1.95*	27	^{s}*2.26*	*2.75*	-2.118	*	.034
hands-on	116	^{s}*5.96*	*1.19*	27	^{s}*5.89*	*1.55*	.393		.694
sim.	115	^{s}*4.51*	*1.59*	27	3.63	1.69	2.451	*	.014
Q1	119	^{s}*4.44*	*1.65*	27	^{s}*4.89*	*1.60*	-1.373		.170
Q2	119	^{s}*4.33*	*1.56*	27	4.30	1.88	.020		.984

Note. † = $p < .10$. * = $p < .05$. ** = $p < .01$. Z pos. = more agreement of students;
Z pos. for pairs = students' endorsement of hands-on is stronger.
All questions (see page 67) were coded in a 7-point Likert scale
ranging from 1 = strongly disagree to 7 = strongly agree (4 = neutral).
^{s} = Shapiro-Wilk $p < .05$, no normal distribution.

Table 5.57: Subjective opinions: Mann-Whitney U-test, comparison of German participants' opinions to the opinions of participants from other countries

	German			other			Mann-Whitney	
	N	M	SD	N	M	SD	Z	p
Pair 1	109	^{s}*1.06*	*1.74*	47	1.60	1.83	-1.488	.137
hands-on	110	^{s}*5.89*	*1.17*	47	^{s}*6.13*	*.99*	-1.255	.209
sim.	109	^{s}*4.83*	*1.37*	47	^{s}*4.53*	*1.50*	1.271	.204
Pair 2	108	^{s}*1.56*	*1.86*	47	^{s}*1.77*	*1.68*	-.759	.448
hands-on	109	^{s}*5.89*	*1.12*	47	^{s}*5.94*	*1.13*	-.385	.700
sim.	109	^{s}*4.36*	*1.63*	47	^{s}*4.17*	*1.42*	.985	.325
Pair 3	108	^{s}*.91*	*1.76*	47	^{s}*1.00*	*1.35*	-.152	.879
hands-on	109	^{s}*5.76*	*1.12*	47	^{s}*5.62*	*1.11*	.891	.373
sim.	109	^{s}*4.87*	*1.35*	48	^{s}*4.60*	*1.16*	1.519	.129
Pair 4	108	^{s}*1.19*	*1.87*	47	^{s}*1.26*	*1.39*	-.193	.847
hands-on	109	^{s}*5.50*	*1.24*	47	^{s}*5.68*	*1.11*	-.692	.489
sim.	109	^{s}*4.31*	*1.46*	47	^{s}*4.43*	*1.31*	-.351	.726
Pair 5	109	^{s}*1.57*	*2.15*	47	^{s}*1.55*	*1.98*	.206	.837
hands-on	110	^{s}*6.07*	*1.25*	47	^{s}*5.87*	*1.36*	1.054	.292
sim.	109	^{s}*4.50*	*1.64*	47	^{s}*4.32*	*1.63*	.604	.546
Q1	113	^{s}*4.60*	*1.69*	46	^{s}*4.65*	*1.52*	-.133	.894
Q2	113	^{s}*4.26*	*1.66*	46	^{s}*4.72*	*1.41*	-1.589	.112

Note. Z pos. = more agreement of Germans;
Z pos. for pairs = Germans' endorsement of hands-on is stronger.
All questions (see page 67) were coded in a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree (4 = neutral).
^{s} = Shapiro-Wilk p < .05, no normal distribution.

Table 5.58: Subjective opinions: Mann-Whitney U-test, comparison of undergraduate students to graduated participants of the survey

	Undergraduate			Postgraduate			Mann-Whitney		
	N	M	SD	N	M	SD	Z		p
Pair 1	49	^{s}*1.02*	*1.68*	109	^{s}*1.34*	*1.83*	-1.077		.282
Pair 2	49	^{s}*1.55*	*1.79*	109	^{s}*1.65*	*1.80*	-.325		.745
Pair 3	48	.75	1.63	109	^{s}*1.03*	*1.66*	-1.009		.313
Pair 4	49	^{s}*.94*	*1.69*	108	^{s}*1.35*	*1.77*	-1.109		.268
Pair 5	49	^{s}*1.39*	*2.07*	110	^{s}*1.57*	*2.12*	-.490		.624
Q1	51	^{s}*4.12*	*1.77*	112	^{s}*4.85*	*1.56*	-2.485	*	.013
Q2	51	^{s}*4.04*	*1.51*	112	^{s}*4.63*	*1.61*	-2.203	*	.028

Note. * = p < .05.
Z pos. = more agreement of undergraduates;
Z pos. for pairs = undergraduates' endorsement of hands-on is stronger.
All questions (see page 67) were coded in a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree (4 = neutral).
^{s} = Shapiro-Wilk p < .05, no normal distribution.

significant and strong correlation between the variables was found (all Spearman's $\rho < .17$, $p > .07$), see Table 5.59.

5.6.3.5 Bachelor- / Master-students

A Mann-Whitney U-test was calculated to compare the opinions of participating bachelor and master students. While the paired questions showed no difference, a statistically significant difference ($Z = -1.836$, $p = .066$) regarding Q2 "Students use simulations in their studies far too often, instead of trying things out." was found (Table 5.60). Bachelor students had, on average, a neutral opinion ($N = 59$, $M = 4.07^s$) while master students agreed more to the statement ($N = 58$, $M = 4.60^s$).

Semester in Bachelor studies

As differences were found between bachelor and master students, a Spearman's rank correlation coefficient was computed to assess the relationship between all items and the self-stated semester of studies of the participating bachelor students. The results are presented in Table 5.59. No correlation between the variables was found (all Spearman's $\rho < .20$, all $p > .16$).

5.6.3.6 Vocational training (VET)

An independent samples t-test as well as a Mann-Whitney U-test was computed to assess the influence of a completed VET on the opinions. As the specific type of VET is only offered in Germany, only participants of German universities were compared for this analysis (Table 5.61).

One statistically significant difference was detected: Participants with VET agreed to "hands-on laboratory experiments offer authentic laboratory experience" (Part of pair 2, only opinion on hands-on: $N = 22$, $M = 5.55^s$) statistically significantly less than participants without VET-degree ($N = 67$, $M = 6.03^s$); $Z = 2.061$, $p = .039$. This effect must be understood relative to the outcomes for the same statement regarding simulations, where the VET also agreed less. The paired question finally delivered no trend between the compared learning modes. VET participants rated laboratory experiments (independent from the mode) as less authentic than persons without a former VET education.

According to the data presented in Table 5.61, no other significant trends between persons who finished a VET and persons without a VET degree were found.

5.6.3.7 Type of university

An independent samples t-test as well as a Mann-Whitney U-test was conducted to compare the opinions between students of German Universities of Applied Sciences and students of German traditional universities. No significant differences were found (Table 5.62), except for Q1 and Q2. UAS students ($N = 85$, $M = 4.47^s$) stated to use simulations less often than students of German traditional universities ($N = 28$,

Table 5.59: Subjective opinions: Spearman's rank correlation, no strong and statistically significant correlation was found.

		Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Q1	Q2
Years of	ρ	-.162	.159	-.002	.041	-.031	.011	-.021
practical	p	.073	.079	.981	.651	.733	.903	.810
experience	N	124	123	123	124	124	127	127
Semester of	ρ	-.095	-.003	.020	.192	-.015	-.156	.135
bachelor	p	.492	.984	.886	.161	.914	.237	.309
studies	N	55	55	54	55	55	59	59

Table 5.60: Subjective opinions: Mann-Whitney U-test, comparison of bachelor and master students

	Bachelor			Master			Mann-Whitney		
	N	M	SD	N	M	SD	Z		p
Pair 1	55	^{s} *.98*	*1.73*	58	^{s}*1.48*	*1.49*	-1.603		.109
Pair 2	55	^{s}*1.49*	*1.74*	58	^{s}*1.53*	*1.75*	-.179		.858
Pair 3	54	^{s} *.78*	*1.56*	58	^{s} *.74*	*1.41*	.027		.979
Pair 4	55	^{s}*1.00*	*1.61*	57	^{s}*1.16*	*1.88*	-.107		.915
Pair 5	55	1.45	2.06	58	^{s}*1.48*	*1.87*	-.003		.998
Q1	59	^{s}*4.29*	*1.79*	58	^{s}*4.62*	*1.51*	-.900		.368
Q2	59	^{s}*4.07*	*1.55*	58	^{s}*4.60*	*1.54*	-1.836	†	.066

Note. † = p < .10.
Z pos. = more agreement of bachelor students;
Z pos. for pairs = bachelor students' trend towards hands-on is stronger.
All questions (see page 67) were coded in a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree (4 = neutral).
^{s} = Shapiro-Wilk p < .05, no normal distribution.

M = 5.00); Z = -2.076, p = .038. While UAS students had in average neutral opinions towards the statement “Students use simulations in their studies far too often, instead of trying things out” (N = 85, M = *4.05*^{s}), German traditional university students tend to agree to the statement more (Q2: N = 28, M = 4.89); Z = -2.358, p = .018.

5.6.4 Results in the knowledge tests (DS-A)

The questionnaire for subjective opinions was published during the second phase and independent from the study runs. All active THI STEM students were invited to participate. Thus, students of different study programs participated.

Nevertheless, some of the data sets could be correlated by stated keywords regarding the main study: The questionnaire asked whether respondents had participated in the main study. If they had, they were asked to provide their code word (optional).

Twenty-three partly incorrect code words were provided. Fifteen of these data sets could be correlated with the data of the other data sources (R8: N = 11; R3: N = 3; R2: N = 1). Most of this data was derived from an international run of the second phase. The first phase’s low response rate can be explained by the fact that students of the first study phase had mostly left university by the time the questionnaire was published.

5.6.4.1 Paired single choice questions

A Spearman’s rank correlation coefficient was computed to assess the relationship between students’ average test performance as well as students’ superior learning mode (hands-on vs. hidden simulations) and the responses to the paired questions (see Table 5.63). No significant effect was found for the two modes in the second phase, which were perceived as identical. In test performance, statistically significant effects were identified. The following participant groups tended to have *bad test results*:

- Participants who stated to learn well with simulations (ρ = -.56, p = .028).
- Participants who stated that the outcomes of hands-on laboratory experiments allow learners to familiarise with the actual behaviour of batteries (ρ = -.55, p = .034).
- Participants who stated that most of the students believe that hands-on experiments provide them with best opportunities to study the real-world behaviour of batteries (ρ = -.60, p = .017).

Analysis of the pairs delivered no statistically significant correlations.

These outcomes need to be treated with care, as some of the statistically significant items pointed towards hands-on, some of them pointed towards simulation, and none of the pairs delivered statistically significant correlations. Only a single data set was derived from the German runs.

Table 5.61: Subjective opinions: independent samples t-test and Mann-Whitney U-test, comparison of participants from German universities with/without VET degree

	Non-VET			VET							Mann-Whitney		
	N	M	SD	N	M	SD	df	t	p	d	Z		p
Pair 1	67	^{s}*1.25*	*1.75*	21	^{s}*1.14*	*2.03*					.251		.802
hands-on	67	^{s}*5.94*	*1.09*	22	^{s}*5.64*	*1.33*					.898		.369
sim.	67	^{s}*4.69*	*1.35*	21	^{s}*4.48*	*1.54*					.489		.625
Pair 2	67	^{s}*1.70*	*1.83*	21	1.71	2.05					-.030		.976
hands-on	67	^{s}*6.03*	*.94*	22	^{s}*5.55*	*1.01*					2.061	*	.039
sim.	67	^{s}*4.33*	*1.57*	21	3.86	1.56					1.216		.224
Pair 3	66	1.00	1.81	21	.95	1.53	85	.11	.914	.03	.253		.800
hands-on	66	^{s}*5.77*	*1.09*	22	5.59	.85					1.069		.285
sim.	67	^{s}*4.79*	*1.45*	21	4.67	1.11					.517		.605
Pair 4	67	1.13	2.00	20	^{s}*1.45*	*1.93*					-.400		.689
hands-on	67	^{s}*5.49*	*1.21*	21	^{s}*5.29*	*1.15*					.859		.390
sim.	67	^{s}*4.36*	*1.51*	21	^{s}*3.86*	*1.28*					1.233		.217
Pair 5	67	1.70	2.30	21	1.62	1.69	86	.15	.880	.04	.288		.774
hands-on	67	^{s}*6.00*	*1.30*	22	^{s}*6.05*	*.79*					-.744		.457
sim.	67	^{s}*4.30*	*1.69*	21	^{s}*4.43*	*1.54*					-.195		.846
Q1	68	^{s}*4.63*	*1.58*	24	4.08	1.84					1.409		.159
Q2	68	^{s}*4.06*	*1.62*	24	4.04	1.88					.036		.971

Note. * = p < .05. d = Cohen's d
Z pos. = more agreement of non-VET; Z pos. for pairs = non-VET trend towards hands-on is stronger.
All questions (see page 67) were coded in a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree (4 = neutral).
^{s} = Shapiro-Wilk p < .05, no normal distribution.

Table 5.62: Subjective opinions: independent samples t-test and Mann-Whitney U-test, comparison of German UAS and German traditional university students

	UAS			trad. Uni.							Mann-Whitney		
	N	M	SD	N	M	SD	df	t	p	d	Z		p
Pair 1	80	^{s}*1.14*	*1.83*	29	.86	1.48					.942		.346
Pair 2	80	^{s}*1.61*	*1.80*	28	1.39	2.04					.432		.666
Pair 3	79	^{s} *.96*	*1.71*	29	.76	1.94					.636		.525
Pair 4	79	^{s}*1.33*	*1.97*	29	.83	1.54					1.051		.293
Pair 5	80	1.50	2.17	29	1.76	2.13	107	-.55	.582	-.12	-.511		.610
Q1	85	^{s}*4.47*	*1.62*	28	5.00	1.87					-2.076	*	.038
Q2	85	^{s}*4.05*	*1.68*	28	4.89	1.42					-2.358	*	.018

Note. * = p < .05. d = Cohen's d
Z pos. = more agreement of UAS students;
Z pos. for pairs = UAS students' trend towards hands-on is stronger.
All questions (see page 67) were coded in a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree (4 = neutral).
^{s} = Shapiro-Wilk p < .05, no normal distribution.

Table 5.63: Subjective opinions: Spearman's rank correlation with students' test performance and students' superior learning mode

	TP N = 15 ρ		TP N = 15 p	SLM 2nd phase N = 14 ρ	SLM 2nd phase N = 14 p
Pair 1	.25		.360	-.01	.975
hands-on	-.20		.472	.32	.264
sim.	-.56	*	.028	.15	.609
Pair 2	-.01		.969	-.15	.599
hands-on	-.33		.228	.01	.971
sim.	-.14		.621	.23	.421
Pair 3	-.23		.407	.23	.438
hands-on	-.55	*	.034	.15	.618
sim.	.07		.791	-.19	.527
Pair 4	-.21		.453	-.13	.652
hands-on	-.60	*	.017	-.18	.530
sim.	.00		.987	.14	.627
Pair 5	-.15		.590	-.08	.798
hands-on	-.25		.364	-.04	.896
sim.	.03		.926	.11	.707
Q1	-.22		.436	.11	.698
Q2	-.49	†	.062	-.12	.685

Note. † = $p < .10$. * = $p < .05$.
(ρ pos. = students who agreed with the statement performed better in the tests;
ρ pos. for pairs = students who more strongly agreed with the statement when related to hands-on experiments than when related to simulations performed better in the tests).

5.6.4.2 Single choice questions

The same analysis was performed for the single choice questions. A trend was detected: Participants who stated that students use simulations far too often, instead of trying things out, tended to perform worse in the tests ($\rho = -.49$, $p = .062$). As with the paired questions, no significant trend was detected for the superior learning mode.

5.6.5 Amount of Practical Experience

Sixteen data sets could be correlated (R2: N = 1; R6: N = 3; R8: N = 12). Most of the data derived from an international run.

5.6.5.1 Paired single choice questions

A Spearman's rank correlation coefficient was computed to assess the relationship between students' APE and the responses to the paired questions (see Table 5.64).

- Participants with a high APE-M tended to state that they learned well with hands-on experiments ($\rho = .511$, $p = .043$).
- Participants with a high APE rated the authentic laboratory experience of hands-on experiments higher than the authenticity of simulations ($\rho = .512$, $p = .042$).
  - This effect was caused especially by the higher rating of the hands-on mode ($\rho = .535$, $p = .033$); opinions on simulations showed no statistically significant correlation to the APE.
  - APE-M showed high correlation ($\rho = .581$, $p = .018$), while APE-E showed no statistically significant correlation.
- Participants with a high APE rated the hands-on laboratory experiments better than simulations in allowing learners to familiarise with the actual behaviour of batteries ($\rho = .621$, $p = .010$).
  - This effect was mainly caused by the lower rating of simulations ($\rho = -.636$, $p = .008$); opinions on hands-on experiments showed no statistically significant correlation to the APE.
  - Again, APE-M showed high correlation ($\rho = .596$, $p = .015$), while APE-E showed no statistically significant correlation.

5.6.5.2 Single choice questions

The same analysis was performed for the single choice questions. No statistically significant trend was detected.

Table 5.64: Subjective opinions: Spearman's rank correlation with APE

		APE	APE-E	APE-ED	APE-EC	APE-M	APE-MC	APE-MD
Pair 1	ρ	.148	-.183	-.035	-.217	.412	.185	.423
	p	.584	.498	.899	.419	.113	.492	.103
hands-on	ρ	.188	-.143	-.038	-.111	.511	.433	.422
	p	.487	.598	.889	.682	* .043	† .094	.103
sim.	ρ	-.061	.119	.017	.177	-.183	-.003	-.232
	p	.823	.660	.951	.513	.498	.991	.387
Pair 2	ρ	.512	.137	.266	.046	.581	.444	.413
	p	* .042	.614	.320	.865	* .018	† .085	.112
hands-on	ρ	.535	.185	.295	-.002	.523	.511	.194
	p	* .033	.494	.268	.993	* .038	* .043	.472
sim.	ρ	-.202	.108	-.055	.104	-.352	-.181	-.405
	p	.453	.691	.841	.701	.181	.501	.120
Pair 3	ρ	.621	.237	.382	.084	.596	.613	.256
	p	** .010	.378	.144	.756	* .015	* .012	.338
hands-on	ρ	.145	.026	.172	-.093	.118	.146	-.060
	p	.591	.923	.523	.731	.664	.589	.826
sim.	ρ	-.636	-.212	-.305	-.089	-.673	-.697	-.348
	p	** .008	.431	.250	.742	** .004	** .003	.187
Pair 4	ρ	.034	-.353	-.110	-.437	.230	.221	.096
	p	.901	.180	.685	† .091	.392	.411	.722
hands-on	ρ	.052	-.149	.048	-.254	.131	.040	.104
	p	.847	.583	.860	.342	.629	.884	.702
sim.	ρ	.013	.339	.167	.370	-.144	-.141	-.033
	p	.961	.200	.538	.158	.595	.602	.904
Pair 5	ρ	.317	-.030	.096	-.115	.346	.403	.053
	p	.232	.913	.723	.671	.189	.122	.846
hands-on	ρ	.326	.002	.097	-.087	.420	.427	.181
	p	.219	.995	.719	.749	.105	.099	.501
sim.	ρ	-.240	.072	-.041	.142	-.264	-.332	.016
	p	.371	.790	.879	.600	.323	.210	.952
Q1	ρ	-.003	.208	.040	.268	-.134	.005	-.278
	p	.991	.440	.882	.315	.621	.986	.296
Q2	ρ	-.099	-.106	-.075	-.292	-.006	-.176	.090
	p	.715	.696	.783	.272	.982	.513	.739

Note. N = 16. † = p < .10. * = p < .05. ** = p < .01.
(ρ pos. = students who agreed to the question had high APE;
ρ pos. for pairs = students who agreed more with the question regarding hands-on in comparison to simulations had high APE).

5.6.6 Summary of results of DS-E

Summary of section 5.6 "Subjective opinions on the learning modes, general (DS-E)"

In the online questionnaire, approximately 70% of the participants were from Germany; the rest of the data were collected from other countries.

Paired single choice questions
Statistically significant and strong effects were found. Participants

- expected to learn better with hands-on experiments compared to simulations.
- believed in a more authentic laboratory experience with hands-on experiments.
- stated that the outcomes of hands-on experiments more accurately allow learners to familiarise with the behaviour of batteries.
- believed that others/the majority also benefit more from hands-on experiments to study the real behaviour of rechargeable batteries.
- claimed they would visit hands-on laboratories more often, compared to simulated experimental lessons, if visiting was optional.

The strength of preference of hands-on experiments was analysed for different participant characteristics separately and it was found that

- the trend of *students* towards hands-on experiments was statistically significantly weaker than the trend of the *rest of participants.* Here, the differences were caused mainly by differing opinions on the simulated experiments, while the opinions on the hands-on mode did not differ significantly. This *could* be correlated with the lower age of this group, but the data does not allow to investigate that further.
- Participants with a high APE-M were more likely to state that they learned well with hands-on experiments (mainly based on international run R8).
- Participants with a high APE (mainly APE-M) rated the authentic laboratory experience of hands-on experiments higher than the authenticity of simulations (mainly based on international run R8).
- Participants with a high APE (mainly APE-M) rated the outcomes of hands-on laboratory experiments better than simulations to allow learners to familiarise with the actual behaviour of batteries (mainly based on international run R8).

No differences were found regarding

- German vs. international participants
- undergraduate and postgraduate students.
- the years of practical experience.
- students in Bachelor and Master programs
- the semester of study of Bachelor students.
- if a VET was finished or not.
- the type of school (trad. university or UAS).

Participants with former VET education showed a trend towards more doubts about the authenticity of laboratory experiments.

Single choice questions
On average, participants stated that they used simulations to understand technical problems better (Q1). They also somewhat agreed that students use simulations far too often instead of trying things out (Q2). Both trends were statistically significant.

The trend towards the opinion that students use simulations too much in their studies was slightly weaker in Germany than in other countries.

When asked directly for the recommended mode (Q3), more than 70% of the participants favoured hands-on experiments in the university, instead of simulations at the university or simulations at home.

Postgraduates claim to use simulations more often to understand technical problems, compared to undergraduate students (Q1).

The group of postgraduates and the group of master students share the opinion that students generally use simulations too often (Q2) while the undergraduates/bachelor students on average reported a neutral opinion on that point. German traditional university students also were more likely to express that students use simulations in their studies too often, instead of trying things out than German UAS students. Also, students from traditional universities state to use more often simulations compared the UAS students. As traditional university students were mainly masters, these items need to be treated carefully, as they may be linked.

Free-text responses
In two free-text questions, the participants were asked to state their general opinion on both modes separately. Both with simulations and with the hands-on mode, most participants emphasised advantages. Nevertheless, the share was highest for the hands-on experiments (87% vs. 46%). While the only negative

aspect mentioned for the hands-on mode was the higher effort required to perform the experiments, for simulated experiments the participants stated different negative aspects. The share of negative opinions on simulations was much higher (35%) compared to those on hands-on experiments (10%). The participants mainly criticised the learning mode by stating doubts about the realism of simulations. Also, neutral comments were stated more often (19% vs. 3%), which mainly requested a combination of simulations with hands-on experiments. The outcome of the evaluation of the free-text responses fits the image derived from the single-choice responses well.

5.7 Objective data – THI university database (DS-F)

This investigation was done using the methodology outlined in section 4.7. A summary of the findings is presented on page 169.

University records of 11,476 students were analysed. 39% of the students in the STEM field had finished a VET before enrolling at UAS Ingolstadt (THI), while only 30% of the students outside STEM (e.g. business administration) came from a VET background. This research focused on the data of STEM/engineering programs of the electrical and mechanical faculty; the following chapter describes the results of 7,930 student careers (including 860 students who were still enrolled when the data was requested). As nearly all of the compared subsets were not normally distributed non-parametric tests were employed to compare the data.

5.7.1 Secondary school education and work experience

The following paragraph analyses the former career of the enrolled VET-participants and the timespan between finishing the VET program and enrolling. Three groups could be identified:

First, students who did their VET after upper secondary school (12% of all students with VET degree, N = 367; M = 1.25 years, SD = 2.89 years) or specialised upper secondary school (16% of all students with VET degree, N = 484; M = 1.30 years, SD = 2.02 years) worked in their VET profession (or did different things, like a gap-year) for approximately one year between VET completion and enrolment (Figure 3.1, dashed line). These students had qualifications necessary for enrolment anyway, independent of their VET. This means that 28% of the STEM students with a VET education had decided in their previous life to pursue VET education *instead* of higher education and enrolled additionally at university after successfully completing the VET program, without more mandatory schooling in secondary education.

Secondly, former VET participants which had to add secondary school education for study allowance. This major part had previously graduated from senior vocational school (65% of all students with VET degree) and took approximately one year longer on average between finishing their VET and enrolment (N = 2008; M = 2.54 years, SD = 1.85 years) compared to the first group. Nevertheless, this does not equate to more industry experience, as the difference to the first type of student (+1 year) directly corresponds to the necessary additional time to complete the required upper secondary school education *after* the VET-program. Thus, the average working experience after completing the VET is similar among the first two groups.

Thirdly, a small portion (7% of all students with VET degree) received their university entrance qualification through further training qualifications within the VET system (e.g. attending a master craftsman course). This group averaged a much longer timespan between finishing their VET and enrolling (N = 211; M = 4.77 years, SD = 3.34 years) and thus had a lot more work experience than the other groups with VET degrees.

5.7.2 Age at enrolment

The aforementioned delays also explain well the differences regarding age at enrolment, which is approximately two years higher for VET graduates than for those who enrol directly after school (all STEM study program students arriving: Non-VET: N = 4817, M = 20.63 y, SD = 2.39 y; VET: N = 3113, M = 22.82 y, SD = 2.78 y). VET-Participants are two years older when enrolling as they invested two to three years into their VET program, plus approximately one year of working experience. On the other hand, they saved time for theoretical schooling (10 years before VET plus 1 year after, instead of 13 years).

These numbers do not change when including admissions from non-STEM subjects, so age does not seem to be correlated to the main subject of study (all arriving, including Business School: Non-VET: N = 7289, M = 20.65 y, SD = 2.42 y; VET: N = 4191, M = 22.95 y, SD = 2.93 y).

5.7.3 Influence of former VET on studies

To investigate the influence of former VET education on the students' studies a non-parametric Mann-Whitney-U-test was conducted. The result is shown in Table 5.65.

Table 5.65: Non-parametric Mann-Whitney-U-test on university (THI) data of successful STEM *graduates* grouped by VET

	non-VET				VET						
	N	M Rank	Mdn	M	N	M Rank	Mdn	M	Z		p
Matri age	2080	1367	20.04	20.36	1665	2505	21.88	22.39	-31.99	**	.000
Matri len (sem.)	2080	1941	8	7.93	1665	1787	8	7.74	4.55	**	.000
Degree age	2078	1423	23.99	24.30	1665	2436	25.74	26.24	-28.49	**	.000
Mark first 4 sem.	2078	1857	2.40	2.35	1665	1891	2.42	2.36	-.95		.340
Mark all 7 sem.	1404	1320	2.29	2.26	1301	1388	2.33	2.30	-2.25	*	.024
∑ ECTS sem. 1	1764	1639	19	16.60	1474	1596	18	16.20	1.32		.186
∑ ECTS sem. 2	2033	1832	46	41.58	1635	1837	47	41.70	-.14		.891
∑ ECTS sem. 3	1957	1734	70	65.60	1554	1784	72	67.16	-1.47		.141
∑ ECTS sem. 4	1958	1713	96	93.18	1594	1855	104	97.05	-4.11	**	.000
∑ ECTS sem. 5	1739	1484	119	115.12	1450	1728	127	123.28	-7.46	**	.000
∑ ECTS sem. 6	1950	1661	162	153.56	1586	1901	170	16.31	-6.97	**	.000
∑ ECTS sem. 7	1831	1585	195	181.62	1514	1780	196	188.49	-5.98	**	.000

Note. † = p < .10. * = p < .05. ** = p < .01. pos. Z = higher value of non-VET.
In German UAS, lower numbered grades correspond to better-performing students (scale 1-5).
In all significant categories, the statistical significance was caused by differences of mean/median.

5.7.3.1 Reason for un-enrolment

The end of a student's tenure at THI came for two reasons, which seemed to be correlated to the student's age at enrolment: successfully graduating or being un-enrolled early, usually for failing a test. In order to compare the different histories of student groups, this analysis was performed for VET (failed: N = 352, M = 23.39 y, SD = 2.84 y; graduated: N = 1,665, M = 22.39 y, SD = 2.44 y; Z = 7.17, p < .001) and non-VET (failed: N = 667, M = 21.21 y, SD = 2.60 y; graduated: N = 2,080, M = 20.36, SD = 1.80 y; Z = 7.56, $p < .001$) separately. When students are older at enrolment, they are more likely to fail. In both groups, the difference of median ages (fail/graduate) was approximately one year, and the effect size was very similar for both VET and non-VET.

5.7.3.2 Duration of successful studies

Former VET participants complete successful studies statistically significantly faster, but the difference is very small (Non-VET: N = 2,080, M = 7.93 sem, SD = 1.29 sem; VET: N=1,665, M = 7.74, SD = 1.21; Z = 4.55, $p < .001$) Thus, the age at successful completion of engineering studies deviates also by approximately 2 years (Non-VET: N = 2,078, M = 24.30 y, SD = 1.92 y; VET: N = 1,665, M = 26.24 y, SD = 2.53 y; Z = -28.49, $p < .001$) – like the age at enrolment.

5.7.3.3 The collection of ECTS points

Looking to the process of successfully collecting ECTS points (see [1]), former VET participants fall behind their colleagues in the first semester, but speed up over the semesters, leading to a significant advantage for VET-participants from semester four on-wards (see Table 5.65).

5.7.3.4 Reasons for un-enrolment from the university

The objective data from the university administration was analysed for the reasons of un-enrolment in detail. The semester dependent results, comparing VET and non-VET students, are shown in Figure 5.21. Differences between both analysed groups become noticeable in the second semester, when former VET-participants less often decide to give up their studies or change to another university.

In Bavaria, if a UAS-student fails the test for the same class three times, he/she is forced to un-enrol from the study program. Between the third and the fifth semester, former VET-participants less often trigger mandatory un-enrolment as of this reason. This could be due to a more unobstructed view on the study programs' topics in the first semesters. Mandatory un-enrolment in the third through fifth semester is usually triggered by finally failing subjects which are intended to be passed in the very first semesters. These subjects are mostly basic engineering subjects, like mathematics, physics, etc.

Starting from semester seven (the regular length of a bachelor program), a higher share of former VET-participants successfully finished their study program. Only 57% of students without a VET (who were enrolled for at least half a semester) received a bachelor's degree in the intended program, whereas 64% of former VET participants graduated successfully.

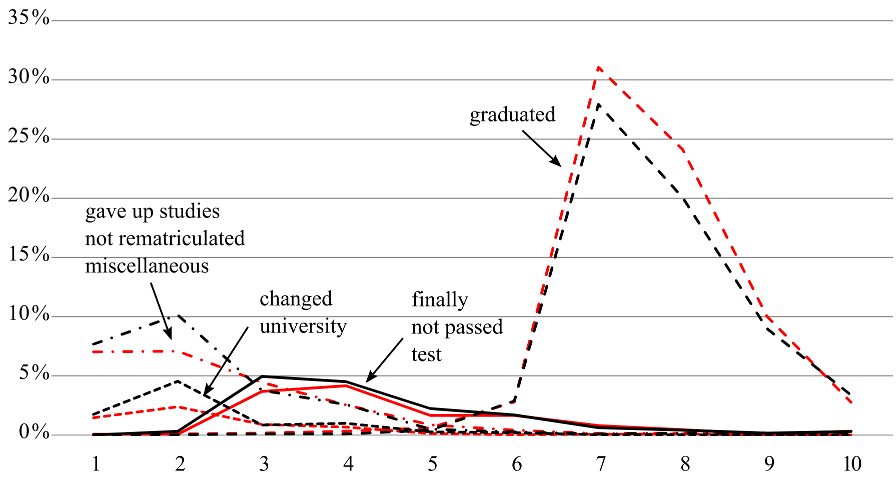

Figure 5.21: Analysis of all former (un-enrolled) THI STEM students' careers for reasons of un-enrolment grouped by semester. 100% = all former VET students (red), 100% = all non-VET (black), N=9,956 students. Enrolled students who never took up their studies at THI were not considered.

5.7.3.5 Grades

VET participants finished their studies marginally faster than their colleagues without VET. However, they performed statistically significant lower regarding study grades, although the effect was very small (All grades of the bachelor study programs, semester 1-7, only students who successfully reached the bachelor degree; non-VET: N = 1,404, M = 2.26, SD = .47; VET: N = 1,301, M = 2.30, SD = .46; Z = 2.25, p = .024). Grades in German tertiary education are assigned from 1 = high distinction to 4 = pass and 5 = fail.

The average grades of the first four semesters (see Table 5.65) did not differ significantly. For technical reasons, only passed exams were considered for the grade analyses, and counted in the semesters the tests were intended for. Thus, a student who failed a test twice and passed the third run with a very good mark contributed with this good mark to his/her final grade. The database did not allow for analysis of the older results.

Grades of cohorts taking part in the main study

The data were also filtered by year of enrolment and study program to examine possible biases in the cohorts taking part in the main research study. The outcomes are presented in the respective analysis, see subsection 5.2.3.

5.7.3.6 Share of former VET participants

Based on the university data, the share of former VET participants fluctuated between 40% to 60% in different cohorts, but was in sum equivalent in both phases (see Table 4.5). The self-stated information of the actual participants is more stable and reliable (between 50 and 60% VET in R2, R6, R9).

5.7.4 Summary of results of DS-F

Summary of section 5.7 "Objective data – THI university database (DS-F)"

To establish objective background information, the student database of THI university was accessed. The analysis was focused on the differences between VET and non-VET students to explain the trends found in the main research. Only STEM/engineering study programs were evaluated, data of the business school was excluded. Data of 7,930 students were analysed.

- 39% of the students in the STEM field had finished a VET before enrolling.
- 28% of the STEM bachelor students with a former VET education had qualifications necessary for enrolment, independent of their VET. They had decided in their previous life to pursue VET education *instead* of higher education and enrolled afterwards at university. 65% of the former VET participants had to visit additional secondary schools to enrol at university. A small portion (7%) of former VET students received their entrance qualification through further training qualifications within the VET system.
- The work experience of students who completed a regular VET program *and* gained a regular university entrance qualification through a secondary school education did not differ based on the order they did those things in.
- The work experience of students who got their university entrance qualification based on further training qualifications, like master craftsman courses, was significantly higher.
- Former VET-participants (23 years old) were approximately two years older than students without VET when enrolling at the university.
- The older students are at enrolment, the more likely they are to fail their studies. (Since VETs are naturally older at enrolment, this analysis was performed for VET and non-VET separately, and the effect size was similar for both groups.)
- On average, former VET-participants complete their bachelor studies somewhat faster (7.74 vs. 7.93 semesters). The difference was statistically significant but practically irrelevant. The difference was caused by former VET participants speeding up in later studies (4th semester onwards).
- VET graduates more seldom give up their studies in the first semesters. Non-VET leave their study program more often (e.g. to change to another one or another university). Later on, former VET participants are less

likely to trigger mandatory un-enrolment for failing a test. The basic lectures in the first semester are more often reasons to fail on for the non-VET than for the VET-participants.

- Finally, 57% of STEM students without a completed VET graduate at THI, while 64% of former VET-participants are successful.
- Regarding final grades, VET-participants perform somewhat weaker than the non-VET students. The difference was statistically significant but was practically irrelevantly small.

It is necessary to mention some limitations for generalisation: UAS are in an extraordinary situation compared to traditional universities in Germany. It might be the case that students with the best grades in standard secondary education choose traditional universities (e.g. studying human medicine, law, etc.), instead of studying at a UAS. Thus, the results might be biased and cannot generalised for full German universities. The presented data naturally compares only groups of students enrolled at UAS – and it is based on one university (THI) only.

Furthermore, it seems plausible that only former VET candidates who performed well regarding the VET schooling component tend to later go on to university, as they make a conscious decision to rejoin a formal school environment. Thus, the results include that inherent preselection.

Chapter 6

Analysis and discussion

The link between theoretical learning and laboratory experiments is particularly relevant at German Universities of Applied Sciences, which attribute great importance to practice-guided learning. Equipment for hands-on laboratory experiments as well as laboratory supervision of classes that require physical equipment can be costly, especially when dealing with potentially dangerous materials such as lithium-ion battery cells. Therefore, hands-on laboratories are often replaced by computer-based learning in the form of simulated experiments.

It is necessary to compare the effectiveness of simulated laboratory experiments to the effectiveness of hands-on laboratory experiments to avoid a deterioration in learning quality. Most scholars did not view different learning modes as directly competitive solutions for the same educational objectives, but instead tried to achieve different study goals, thus developing and optimising each mode independently. A summary of this work is presented in section 2.6.

While the positive effect of laboratories on student learning has been widely accredited, optimal modalities of laboratory experiments have not been established unanimously. This study seeks to add more generalisable insights to the existing body of knowledge by comparing the impact of two different modes of laboratory work on learning while minimising the influence of as many interfering factors as possible (listed in detail in the appendix, page 239). In both modes, hands-on and simulated laboratories, learning objectives and the experimental approach of laboratory exercises were identical.

Execution of the case study had four goals:

- First, the validation of whether the strict methodology is suitable to compare different laboratory modes.
- Second, the comparison of the effectiveness of teaching battery basics in computer-based laboratories with practical hands-on exercises in a case study with nine study runs.
- Third, to monitor whether the learning success remains the same if students are not aware that they used simulations.

- Fourth, the investigation if there is a relationship between student qualities / attributes / educational background and the more successful learning mode in the case study.

As it was found in the first study runs that the superior learning mode differed for students with and without completed Vocational Education and Training (VET) program, the study was expanded to investigate the reasons behind these results. Further methods were included to investigate the unique characteristics of these groups, which could explain the reasons for different behaviour. The data were analysed to compare students' background and general appraisal of the experimental modes, as well as their satisfaction with the modes.

6.1 Identified aspects

6.1.1 The case study did not reveal a universally superior learning mode

Nine study runs with German and international participants at two different universities were conducted. A counterbalanced within-subject research methodology was applied; it focused on the comparison of the laboratory modes hands-on and simulation in the local domain. A case study was performed on the teaching of battery basics and measurement methods for battery cells and energy storage systems.

Accompanying lectures, experimental instructions, teachers, learning objectives, tests, and many other variables were controlled in both groups. Identical experimental procedures were used in the compared modes.

Test results related to knowledge acquisition as a result of conducting laboratory exercises in different modes were collected.

The research was split into two phases: In the first phase hands-on experiments were compared to *overt* simulations. In the second phase the simulation condition was *hidden* and students thought they were conducting hands-on experiments. Their results were compared to those from the hands-on condition.

When the simulation condition was overt, a weak, but significant outcome of better knowledge acquisition with hands-on laboratory experiments was achieved. This is against the trend of the recent literature that reported on better or equal learning with nontraditional (virtual/simulated) laboratories [9–11]. The two study runs with German University of Applied Sciences (UAS) students showed statistically significant trends towards better learning with hands-on experiments, while students in the three international runs performed similarly in both modes (Table 5.3).

Results of the German study run in 2017 (R2) indicated a significant impact of the laboratory mode on the students' performance depending on whether a participant had completed a Vocational Education and Training (VET) program. In the *hands-on* condition, German participants who had completed such a VET program before enrolment at university performed similarly to their peers who enrolled immediately after standard secondary school education. In contrast, *simulated* experiments had less positive effects on former VET participants' learning compared to their peers without a VET degree ($t(46) = 2.74$, $p < .01$, Cohen's d= = .72).

This outcome fits the trend detected by Tsihouridis *et al.* [22]. Their literature review found a higher need for hands-on experimentation with a lower educational level of participants. In the present study, the educational level was the same (former VET as well non-VET participants were enrolled in the same bachelor program), but the history of the participants (subsection 3.2.1) needs to be taken into consideration.

The Moodle questionnaire data (DS-D) supports the claim that the experimental procedures in both modes can be considered equal. The problems mentioned ($Z = .837$, $p = .402$) as well as difficulties participants encountered ($Z = -1.116$, $p = .265$) accounted only for very slight and statistically not significant differences.

Participants with a VET degree performed statistically significantly better after hands-on experiments than after simulations ($t(58) = 2.38$, $p < .05$, Cohen's $d = .59$), while students who had not completed a VET program demonstrated no statistically significant differences.

The summarised results of the first phase (both for German runs separately (R1, R2), as well as all participating students (R1 – R5)) suggest that traditional hands-on experiments lead to better learning compared to simulations when teaching battery basics. One needs to be careful when interpreting these two statistics: The share of German students contributing to this result was high (54% in the first research phase). The share of VET participants within the German runs was also high (59%, R2), as in the German runs all students originated from a bachelor program at a University of Applied Sciences. Considering the somewhat weaker effect of the German 2016 study run (R1) – where the information about VET was not recorded – and also the strong dependency on the students with former VET, the found effect can be interpreted as being completely based on the VET subgroup of the German runs.

6.1.2 The participants of both phases were equivalent

In every phase, two runs were conducted in the same German study program "Electric Mobility" (R1, R2 vs. R6, R9) including the VET topic. The share of VET-participants did not differ substantially between cohorts. Additionally, an international run in the same study program "Renew. Energy Systems" was included in each phase. In both study programs no major revision of the curriculum occurred while the study was running. Only study run R4 in the first phase had no equivalent in the second phase – but neutrally contributed to the study.

The equivalence of groups was further evaluated by comparing data about the students provided by the university (DS-F). There was a focus on the German runs for these calculations, as they showed the major effects. Data had to be attributed to the study runs based on the date of matriculation. VET and non-VET participants did not differ statistically significantly between the first phase and second phase regarding age at matriculation and the process of collecting ECTS points in the first three semesters. For non-VET, no differences in grades could be identified between both phases. Student data sets attributed to R6 showed better marks for the former VET-participants compared to all other runs, but this effect could not be replicated in R9. As neither run showed statistically significant results with regards to the overall research question, this did not affect the outcomes of the study. Furthermore, VET participants of all cohorts were equivalent with regards to their age at completion of the VET as well as the time span between VET completion and enrolment.

No major differences regarding APE (VET, non-VET separately) or age were identified between phases (see also Table 4.3).

The moodle responses on the (perceived) hands-on experiments were analysed with no significant differences being found, which suggests that experimentation, perception of experiments and thus participants were comparable between both phases.

6.1.3 The second phase was helpful to verify the results

Results of the first phase of the case study suggested that traditional hands-on experiments led to better learning compared to simulations when teaching battery basics. This result was unexpected, since hands-on experiments were conducted very similarly to simulated laboratories.

When conducting the hands-on laboratory exercises, students only monitored currents and voltages that were displayed on the measurement equipment. Similar to the students conducting simulations, they had to trust the displayed currents and voltages.

Moreover, the experiments for three (B, C, D) out of four content areas did not involve any physical interaction with the equipment during hands-on experiments, as all measurement equipment was controlled by software.

On the one hand, students were aware of the modes utilised during the study runs. Perhaps the known absence of a real device/a real battery made the experimental results of simulations appear insubstantial to the students, causing motivation to suffer.

On the other hand, simulations were created and parameterised to closely imitate the actual properties of real battery cells used in the hands-on mode. Much effort was spent on designing simulations that realistically imitate battery behaviour (see appendix, section F.3). The model was reviewed, see section F.5 to ensure that all participants involved in simulations had the same information as their peers working in the hands-on condition. No differences in battery behaviour were identified in the students' protocols. Nevertheless, hidden weaknesses of the simulation model could not be excluded completely. These could have influenced students' learning negatively.

To understand the reasons for the disadvantages German VET-participants faced when confronted with simulations, another study phase of the mode-comparing experiment (simulations vs. hands-on) was performed: Therein, students in the simulation condition were given the impression that they were performing traditional hands-on experiments (see "hidden simulations", as described in subsection 4.2.9). The simulation model employed was identical to the one used in phase 1, thus the second research phase helped to rule out problems with the simulation model as the basis of VET-students' disadvantages in phase 1.

Results of the second phase

In the second phase, one German run showed a slight trend towards advantages of *hidden simulations*, thus pointing in the opposite direction of the first phase. All other study runs of the second research phase showed no significant differences. The second phase overall showed no significant outcome. Both modes performed similar.

Outcomes were analysed separately depending on whether students had finished a VET before studies or not. No significant differences in students' performances between the modes were found in phase two. Students acquired equal knowledge in both conditions.

There were no significant differences in grades in VET and non-VET cohorts in the first phase. This supports the claim that the above-mentioned difference between hands-on experiments and simulations in the first phase was not caused by a bias in student competencies.

In the second phase, former VET participants performed somewhat better than their colleagues. That difference could be attributed to a general differences in competencies of the participating cohorts. The grade superiority of former VET participants is not common (see below, objective data of all THI STEM study programs), usually non-VET students have equal or slightly better grades (see section 5.7 and subsubsection 5.7.3.5).

Here, it is essential to mention that it was not possible to validate if non-VET from UAS differ significantly from traditional university students. It was challenging to get more of these universities to participate in the research. It could be argued that above-average students are more likely to study at traditional universities.

Hiding simulations in the second research phase worked well

The feedback on the experiments that students provided on moodle was analysed to see how well the simulations had been hidden from participants. The hands-on experiments in the first and second research phase were perceived very similarly: No statistically significant differences were detected in the perception of new insights gained while experimenting, difficulty and relevance. A slight (statistically not significant, $p = 0.089$) trend of fewer problems reported by the participants in the second phase was detected, which can be explained by the improvement of experiments between runs.

Analysis of the other secondary data sources (e.g. correlations of APE vs. the test performance or correlations of test performances in perceived hands-on experimentation with the subjective opinions on the learning modes) showed no significant differences between the two research phases. Across all runs of the second phase, not a single participant expressed doubt (verbally or in the students protocols) that the experimental results were measured on a real cell. The measurable differences between the modes in the first phase and the absence of these differences in the second phase suggest that hiding simulations in the second research phase worked well.

The simulation model of the first phase seems to be correct

The simulation model was thoroughly reviewed by experts. Nevertheless, weaknesses were possible. Based on the contrast between the phases, and the slight trend towards better learning with hidden simulations in the second phase, it was concluded that choosing a faulty simulation model could not have been the reason for the advantages of "hands-on" in previous runs. As the same simulation model was used in the first research phase, weaknesses of the simulation and its parameters are, on the level investigated in the laboratory lessons, highly unlikely.

Impact of the second research phase

The laboratory mode "hidden simulations" of the second study phase does not have practical relevance for teaching. It is unethical to misguide students and unpractical, as creating a working simulation that can't easily be discovered as such requires extreme efforts. In the study at hand however, this laboratory mode enabled the validation of the two hypotheses: Firstly, the simulation model and its parameters worked well. Secondly, the results of the first phase are likely based on perceptional or psychological effects. A further benefit of the second study phase resulted from gaining information through the improved and extended questionnaires.

The identified effect might be underestimated
VET test results across German runs of the first research phase showed a statistically significant difference in test performances favouring the hands-on laboratory mode. Assuming that students gained knowledge and skills during the laboratories, subsequently performed experiments and tests should have been easier to manage. Since the effect points towards better results in knowledge acquisition/retention when conducting hands-on laboratories compared to simulated ones, it is plausible that the differences in performance between groups would have been even higher, had the participants been allocated to either the hands-on or simulated laboratories across all experiments.

6.1.4 Simulations cause disadvantages for students with former VET education

The VET background determined whether or not differences between hands-on and simulated laboratories can be found. Test performances of students who had finished a German VET before studies scored fewer points after simulated experiments than after hands-on laboratories. With German participants who had not finished a VET program, no such differences were found.

As hidden differences in the simulations could be excluded from having been the reason for inferior learning results, psychological effects need to be considered to comprehend the effectiveness of the different laboratory modes. The students could have felt that simulated experiments were less relevant, and they might have lost some of the motivation to comprehend and remember what they had experienced and learned during the laboratory exercises.

6.1.5 The identified trend is likely due to the perception of the modes

When the simulated condition was hidden, no differences were found. Learning results were similar for all students, including the former VET participants. The two German runs of the second research phase (R6, R9) were performed with participants from the same study program as R1 and R2 of phase one (German, University of Applied Sciences, THI, B. Eng. program "Electrical Engineering and Electric Mobility"). Therefore, a change in the type of participants is unlikely the explanation of the missing differences.

It seems that as long as the students perceived tasks as hands-on experiments (e.g. "manipulating" real devices, a real device being tested) simulated experiments performed as good as hands-on experiments – a result which is in line with Lindsay's results [13].

The results suggest that the differences between the two conditions – hands-on and simulated – in phase one can be attributed to the perceived laboratory mode.

The smaller difference between conducted modes in the second research phase regarding the moodle feedback somewhat strengthens the "perception" thesis – since students' perception did not differ between modes in the second phase, no difference in the answers was to be expected.

The objective test results are supported by data from the indirect moodle questionnaire (section 5.5). In the second phase, none of the items showed significant differences between the modes, while differences were generally bigger between the modes of the first research phase (not statistically significant). For example, participants working on *overt simulations* in phase 1 were more likely to state that contents of the experiments were irrelevant to their future professional careers than their peers using *hidden simulations* in phase 2. When simulations were perceived as hands-on experiments, both modes led to the same opinions about the experiment (VET, non-VET and overall). Based on the absent differences in the second phase, one can conclude that the experiments were perceived identically.

6.1.6 It seems that the average participant "prefers" hands-on experiments over simulations

The general Qualtrics online survey (DS-E, page 67) included persons who had *not participated* in the battery laboratories. Statistically significant and strong trends towards better opinions about hands-on experiments were found for nearly all groups of participants, similar like [15]. Hands-on experiments were rated as being much more authentic and helpful for learning by the analysed groups, whereas simulations were generally viewed negatively. Even the effect sizes did not differ between many of the analysed participant characteristics (VET, self-stated amount of experience, semester of studies, and type of school). Based on the evaluations of the two experimental conditions hands-on vs. simulations in the present study, the overwhelming majority rated hands-on experiments to deliver more *authentic* experimental results and expected better learning results for themselves as well as for others. This was found to hold across country of origin, student or graduate status, and other variables.

When asked which laboratory mode they would recommend, more than 70% of participants favoured hands-on experiments in the university over simulations (remotely, at university or at home). To a certain extent, this outcome contradicts the surge of scientific literature promoting online education. Recent literature reports preference of a blending of both modes [22] (an option which was not offered in the questionnaire) or tends in favour of simulations [11]. This discrepancy between the results of the questionnaire and the literature could not be attributed to the high share of VET-participants in the study, as both VET and non-VET students reported similarly strong preferences for hands-on experiments.

In general, participants from a VET background tended to express more doubts about the authenticity and accuracy of laboratory experiments (DS-E, page 146). This contributed to the disadvantages for student learning with simulations within that group.

When measured indirectly by asking for feedback on the previous experiment (DS-D, page 131), *participants of the laboratories* preferred the hands-on mode. It should be noted that, to avoid any influence caused by students' preconceptions about the particular mode, DS-D-data were gathered by asking participants of the laboratories of the first phase about the previous experiments without mentioning that answers would be used to compare the modes. The comparison of hands-on laboratories with the simulation condition showed that the effect size regarding the student's subjective assessments of their gained knowledge (Table 5.41) was very similar to their objective test results of the first phase (Table 5.3).

The share of students disagreeing with an experiment's relevance was much higher in the feedback regarding simulated experiments. It seems that the feedback collected *after* the experiments and objectively measured learning success are correlated.

In both modes, knowledge transfer went well. General subjective opinions/beliefs on the learning mode *before* experimenting might have influenced learning through the respective learning mode and led to weaker results of simulations.

6.1.7 Former VET participants differ from non-VET students

Overall, students with a VET education before enrolment learned better with hands-on experiments than with simulations, as simulations turned out to be less useful for these students. Different approaches were followed to find out the reasons for the differing behaviour of former VET participants.

University database (DS-F)
The analysis of the THI university database in section 5.7 showed that former VET participants are, on average, two years older than their colleagues who enrol directly after school. Considering the students of the German runs were generally young ($\approx$22 y), the age difference might have influenced their behaviour. Based on the data, the older a student, the more likely he/she fails a test. It seems that the aspect of being generally older does not negatively influence the performance of VET students compared to non-VET.

Moreover, VET graduates more seldom give up studies in the first semesters. In advancing their studies, former VET participants are also less likely to trigger mandatory un-enrolment by repeatedly failing a test than their colleagues without VET education.

Furthermore, employment rates for technicians in Germany and especially in the region are very high. It needs to be considered that for VET graduates, consequences of not finishing the study programs are less severe – they have already completed an acknowledged degree, which allows them to enter the employment market at any time.

Generally, non-VET leave their study program more often without degree (for example to change to another program or university), while former VET-participants tend to finish their study programs successfully. VET participants probably select their study program with a higher awareness of their professional career development and personal interests.

Non VET-students more often fail the fundamental lectures in the first semester than VET-participants. Students without former VET education might not have the same motivation for these subjects, having less background knowledge to understand the purpose of these basic subjects. Additionally, for non-VET, it is the first time experiencing a less regulated school-environment. They have to adapt to self-organise their learning in university, often with no mandatory presence and only a single written test at the end of the study module. VET-participants on the other hand have had the opportunity to learn self-organisation during their apprenticeship and work experience.

At graduation, the average marks of former VET-participants were lower than those of non-VET students. The opposite was true for the average duration between enrolment and graduation: On average, VET participants finished their studies faster. In both cases, the difference is statistically proven, but practically irrelevant because of the small effect size.

One can conclude that overall, there is no general difference between student groups with regards to learning success, although the results of this study show that they are "different" types of students. Some aspects may reflect a higher intrinsic motivation among VET-students to study the selected program due to a more unobstructed view of the study program's content and/or a better understanding of their own personal interests.

Amount of Practical Experience (DS-B)
The definition of APE can be found in section 4.3.

Students who completed a VET before university have a higher *Amount of Practical Experience* than their colleagues who enrolled directly after school. That might be related to their age, but also the experiences gained during VET-training and subsequent work in the industry. APE correlated with the “Realistic” category of Holland’s well known RIASEC topology.

As mentioned above, the results of VET after *hands-on* experiments do not differ much from the results of students without Vocational Education and Training, but their results with *simulated* experiments were worse (R2).

For VET students, a higher *Amount of Practical Experience* was related to better results in the tests in both modes, while for the non-VET group, no statistically significant trend was found between their *general* performance and their APE: The lower the *Amount of Practical Experience* of former VET-participants, the worse they performed on average in the tests. VET offers many chances to gain practical experience, so a participant’s low APE despite a VET degree can be an indicator of low motivation to gain practical experience. Thus, these people might also participate in laboratories with less motivation, leading to worse results in the tests.

When comparing the success of both learning modes for individual students (using *superior learning mode*), a high APE was somewhat connected to better learning results with *simulated* experiments in contrast to hands-on experiments. This was found for the German (VET and non-VET: Spearman rank order correlation, $p \approx .10$), but not for international participants. This difference is surprising and can’t be explained by the present data.

RIASEC/Personality (DS-C)
Holland’s RIASEC-typology describes personal interest on a general level using six personality types (Realistic, Investigative, Artistic, Social, Enterprising, and Conventional). Two well-established questionnaires (Germany specific) test sets for the RIASEC-typology were used in later study runs (see section 4.4). Results from these questionnaires (see section 5.4) were used to shed further light on the respective personalities of German VET and non-VET UAS students.

Students with completed VET scored higher values in the “Realistic” category (which was strongly correlated to the self-developed APE index) than non-VET, while scores in the “Investigative” category were smaller. It needs to be taken into account that the questionnaires were created to match a person’s personality to the requirements of future jobs. Based on the questionnaire, the former VET students have shown less congruence with professions requiring a university education, based on their lower “Investigative” values. Nevertheless, their similar performance to non-VET students at the university shows that this would not be a fitting assessment.

The two categories (R, I) correlated only for non-VET – if a participant showed a high I value, he/she also showed a high R value. This correlation was absent for former VET-participants. That could have been a statistical effect to the saturation of the R scale for VET participants, as nearly all VET scored very high to maximum R values. Thus, no conclusions based on the absence of that correlation with VET can be formulated.

On the one hand, the presence of a VET, high Amount of Practical Experience, and the RIASEC-category "Realistic" correlated positive. They seemed to be dependent factors. On the other hand, regarding the superior learning mode (Figure 5.16), APE and VET act clearly as independent factors. It would be interesting to compare the "Investigative" category with the superior learning mode (hands-on vs. simulations), as it could be a precursor for the differing performance of VET and non-VET. The I factor generally had lower values for VET participants, while it was independent of APE. Unfortunately, the RIASEC data were only collected from German runs in the second research phase, where simulations were hidden (perceived as hands-on), and no differences in student learning were discovered.

The detected RIASEC saturation affirms the assumption that the RIASEC questionnaires used was too broad to capture students' interests in a science-based topic. This underlines the necessity of using the self-developed APE scale. Nevertheless, the proposed Amount of Practical Experience (APE) dimension still need to be improved. Especially the APE-M scale showed ceiling effects on sub-scales. Therefore, further items that allow to better distinguish subjects' practical experience would need to be added.

Qualtrics Survey (DS-E)

In the survey, which included answers from persons who had not participated in the laboratory, the general feedback on the authenticity of simulated experiments was worse than on hands-on experiments. In particular, participants with a completed VET education expressed even more doubts regarding the authenticity of laboratory experiments than non-VET students. While the difference between the two modes was identical for VET and non-VET (paired questions regarding both modes), persons with completed VET were particularly likely to doubt the authenticity of the results of simulated experiments, when compared to the attitudes they held towards hands-on experiments, and compared to the attitudes held by non-VET participants towards both modes. These doubts can make the experimental results less relevant. When simulations were hidden, these possible extra doubts were absent, which could have led to the achievement of similar results in both modes in the second phase.

As these answers were given before participating in any experiment or sourced from non-participants, the data gives hints on deviating matter of attitudes of the two groups, which might have led to the differing results in the first phase (hands-on vs. simulations).

General experiences from VET education

VET have more practical experiences, which means they likely have used expensive and potentially dangerous devices in their previous education – belonging to their company. For example, someone working in an automotive workshop could have had the experience that even a short-circuit fault in a car harness does not cause the battery to explode immediately. Thus, this group of participants might be less afraid to work with or harm expensive and potentially dangerous devices, while the non-VET group has to rely on the (overly cautious) safety instructions of the labs. These previous experiences may have contributed to the tendency this group's *superior learning mode* showed towards hands-on, while non-VET's *superior learning mode* tended more towards simulations.

This is somewhat contrary to the assumption that students feel more responsible during hands-on experiments because they can damage laboratory equipment. This assumption can explain better student learning via the hands-on mode. The danger of damaging equipment could force the students to take extra care, and thus, get more involved in hands-on experiments than in simulated ones. Based on that idea, a low APE should be connected with a superior learning mode in direction of hands-on, which was not found *generally*. It was found for both German subgroups (VET and non-VET) separately (For the international runs, no statistically significant trend was found.) Thus, this idea, which is somewhat contrary to the above-mentioned, would only explain the within-group trends. However, all students performing hands-on experiments were informed about the safety switch off box, which protected all devices and participants from harm with almost 100% certainty.

German non-VET students at UAS differ to German students at traditional universities. Similarly, one needs to consider that the VET students (students at the UAS who finished a VET education before studies) are not equivalent to the average successful VET participant. These students are ambitious enough to return to secondary school education and invest several years into their bachelor degree. Considering that this group already has an acknowledged qualifying degree, it seems that engineering students with VET degree have some intrinsic motivation for pursuing a deeper understanding of technical topics.

6.1.8 APE and RIASEC – independent from VET

Besides the above-mentioned points regarding VET vs. non-VET, further hints of different backgrounds and behaviours of students at German UAS were found. German UAS students seem to differ from the international groups.

Amount of Practical Experience (DS-B)

The higher the *Amount of Practical Experience* of German (VET and non-VET) participants, the more successful *simulations* are for teaching – analysed in comparison to the success of *hands-on experiments* for individual participants (superior learning mode, as shown in Figure 5.16).

This was a surprising aspect. On the one hand, the opposite would be plausible, people with high practical experience can be assumed to be familiar with certain devices, and thus may tend to profit from having these devices in front of them.

On the other hand, the data is consistent: low performing VET tended towards hands-on experiments (see Figure 5.4), and low performances were connected to low APE.

Former VET-Participants with high APE showed no difference in learning success between the modes (the superior learning mode value was neutral, as shown in Figure 5.16), while VET with weak APE tended towards better learning with hands-on experiments.

International students showed the opposite behaviour, a high APE was connected to generally bad test performances. It was impossible to pinpoint the reasons for the differing behaviour. The difference seemed to be a general trend, as it was independent of the mode.

As discussed above, while international students with high Amount of Practical Experience (APE) showed weaker results in the knowledge tests, a high Amount of Practical Experience tended to be helpful for the German UAS participants to perform well in the knowledge tests. Figure 5.9 through Figure 5.15 make it clear that the probability of a German UAS participant combining little practical experience with good test performances is low, while this combination described most international participants. Reasons for this unexpected result could be not investigated further, as the study provided no data.

It is important to mention that, based on the results of the study, one needs to be careful when analysing results for “German students”. They form by no means a homogeneous group and should be split strictly between students from UAS and traditional universities and further between VET and non-VET experienced students.

Personality / RIASEC (DS-C)
Holland’s RIASEC-typology describes personal interest on a general level using six personality types. Germany specific test sets for the RIASEC-typology were used in later study runs (see section 4.4). These question sets were used only in German runs, meaning this part cannot contribute to the analysis of differences between the participating German UAS students and internationals. Also, the published difficulties are only viable for comparison to German participants. For further studies it is recommended to include equivalent question sets for all runs. For results which are biased due to a ceiling effect (e.g. “Realistic” in this study, see section 5.4), more specific items need to be included to be able to characterise participants better.

6.1.9 Relevance

Students who claimed that they gained new insights (section 5.5) also tend to believe that the execution of the experiment will help them in their future professional life outside university. Students reported more insights and more perceived relevance after hands-on experiments compared to simulated experiments. In the second phase the connection between perceived insights and test performance was measured. Students who stated that an experiment had delivered new insights tended to produce better test results in the correlated test; no dependency on VET was found.

In the first phase, 71% of participants stated to have gained comprehension, but only 30% rated the conduction of the experiments as fruitful for their professional life. The same was true in the second phase, albeit with improved values (89% new comprehensions vs. 50% relevance). Based on these numbers, one can conclude that the students do not consider all of the gained insights relevant to their profession. The somewhat (not statistically significant) stronger correlation between both points in the simulations mode suggests that teachers should be prepared to explain the relevance of an experiment in simulated laboratories more carefully.

On the one hand, regardless of the learning mode, only 11% of students in the experiment (Table 5.40) stated that they perceived no benefit for their future professional lives from the experiments. On the other hand, the share of students wholly agreeing with the experiments' relevance was only 36%, which was somewhat disappointing. Increasing the share of students who feel that the experiments are relevant, will improve the success in reaching learning outcomes. A detailed look at negative responses reveals advantages after experiments which were perceived as hands-on (modes were: hands-on and hidden simulations). These experiments received a much lower share of negative feedback (7%) than overt simulated experiments (20%). This fact could be explained by a specific group of participants being de-motivated during simulations but not hands-on experiments, leading to better results of hands-on experiments in the tests.

This leads to the question if the objective outcomes can support the idea that the differences that were identified were partly caused by a specific group of students deeming simulated experimentation especially irrelevant. Unfortunately, the code words were not available with DS-D, so linking subjective feedback on simulated experiments with data from objective test results was not possible.

In the second phase, when it was possible to evaluate the correlations with knowledge test results, a dependency ($\rho = .36$) between stated relevance and test performance was found (Figure 5.20), which was statistically significant ($p < .001$). Students who rated experiments as relevant tended to perform better in the concerning test. The difference between VET and non-VET students was not significant, it was a general trend.

In that phase, both learning modes were perceived as hands-on, thus eliminating bias towards simulations as a possible factor influencing student learning. Nevertheless, if an experiment was stated more relevant, the student tended to perform better in the second phase. This correlation supports the claim that creating relevance will increase learning quality, e.g. by giving the example for cold cranking a car on a winter night before cooling a battery to negative Celsius temperatures. The lower perceived relevance of simulations may explain their low performance in the first phase.

Another effect is visible in Figure 5.4 with hands-on vs. simulation data from the first phase: For well-performing students, the difference between modes was small, but for weaker students, the hands-on experiment seemed to be the better performing teaching method. Both aspects support – not finally prove – the assumption that, for a specific group of students (low performers), it is essential to demonstrate relevance, and this could be more easily achieved by experimenting in the hands-on mode.

In the second phase the experiments (all perceived as hands-on) were stated more likely relevant by participants who had a high Amount of Practical Experience and by participants who had a high “Investigative” RIASEC value. Especially the APE-“Electronics” value had a high correlation. This group also stated to have gained more new insights while experimenting. It seems logical that the practically experienced students had more ideas on how to apply the procedure they had learned while conducting the experiments (or editing the laboratory reports) compared to their colleagues with lower APE. Persons with high “Investigative” RIASEC values like “to work with data and to immerse themselves in mental or scientific problems.” It seems logical that this group judges the conduction of laboratory experiments (more) relevant.

In order to fully comprehend the effectiveness of different laboratory modes, these psychological effects need to be considered. Perhaps the absence of a real physical device and/or a real battery was perceived differently by students with and without VET degree. The VET graduates could have felt that simulated experiments were less relevant, and they might have lost the motivation to comprehend or remember what they had experienced during the laboratory exercise.

Relevance and doubted accuracy of experimental results
Participants with a previous VET education generally expressed more doubts about the authenticity of laboratory experiments, compared to non-VETs. In the surveys after conduction, doubts were mostly expressed about the authenticity and accuracy of simulated experiments, while hands-on experiments were reported as more relevant. This can explain one of the reasons for the disadvantages of simulations for student learning. Inaccurate results could be seen as irrelevant, which are remembered less, and offer less motivation for further studies at home.

6.2 Limitations

Several study limitations need to be addressed.

Obviously, the conducted study was a case study (a specific instance that was used to illustrate a more general principle, see [115]) and thus the outcomes contribute only to knowledge of the same combination of learning circumstances (topic is batteries, most students are German, and so forth).

To deliver a general answer to the research questions Q2 and Q3 (for example to give teachers general ideas of whether to start new laboratories simulated or hands-on), it is necessary to use the methodology in many different disciplines and topics.

The advantage of the proposed strict methodology is that the limits for generalisation of the conducted case study can be clearly named.

Domain

This research presents the outcome of a study that was carried out in a *local* access domain. It compares the test results of a hands-on laboratory with that of a virtual laboratory (simulated on a local computer). As the study at hand focuses on a strict methodological approach, a remote laboratory condition was not included, even though many students favour online education and recent literature reports equal or better learning with remote laboratories (for details, see subsection 1.1.1).

Working in local domain meant working as a team in the university, in the same room as the instructor. The location was chosen according to the laboratory mode: hands-on lessons were conducted in a chemistry or electronics laboratory environment, while simulated experiments (if not hidden) were conducted in a computer pool. Theoretically it is possible – but unlikely – that all found differences are based on the differences in learning environment during teaching. Nonetheless, these environments are linked to the laboratory mode. It seems therefore fruitless to further investigate this aspect.

Usage of devices was mediated

There is a general trend towards mediation through computer interfaces in hands-on laboratories [14, 46]. In the study a PC was used to control the experiment and record the gained data.

The battery topic in combination with experiments with extremely low physical interaction in the hands-on mode does not allow universal transfer of the results.

When looking at the different conditions, it must be noted that students monitored currents and voltages displayed on a computer screen in *both* modes. During hands-on laboratory exercises, students additionally had the option to check currents and voltages that were displayed on the measurement equipment and were able to inspect the hardware they are engaged with. But similarly to the students conducting simulations, they had to trust that the displayed currents and voltages were correct. Moreover, the experiments for the content areas B, C, and D did not involve any physical interaction with the equipment during hands-on experiments, as measurement equipment was controlled by software.

Greater differences in learning between groups can be expected when more interaction with the experimental setups is needed.

Tests

Due to the strict methodology, tests (see chapter H) requested knowledge covered by both learning modes. Special knowledge which could be covered only by one of the modes (e.g. wiring high current lines in special way to reduce inductive coupling) was not focused in the lessons and tests. The tests focused on students' conceptual understanding/knowledge. They were not made to judge students' ability to conduct or design scientific experiments involving batteries in the future.

Knowledge loss

An emphasis was placed on keeping an equivalent time span between the theoretical lectures and the corresponding experiments to equate the influence of possible knowledge loss for the compared groups.

Many studies are based on results collected directly after the experiments. However, other studies have found dependencies on the time between the laboratory and corresponding test [39, 57].

The gap in this study was held at approximately one to two weeks, in order to account for some knowledge lost over time.

However, due to time constraints, the results still cannot provide information on long time knowledge retention.

Employing black box simulations to learn vs. learning to simulate

It is important to mention that the simulated laboratory in this study employed a teacher-prepared black-box simulation model. The sessions related to the research (do not confuse with the content of simulation workshop E, section G.10) were not about learning to simulate, but about battery behaviour and how to determine it using a proxy for the non-existing hardware. As the model was not created by the students but externally provided, a certain mistrust in the model might be a given. The students had no information about how the real world system was mapped, and where there were gaps and simplifications. An approach of creating the model with the students (white-box model) could change students' behaviour towards and relationship with the model, as the limits of applicability to real world behaviour would be clearer. VET participants have had more chances to experience cases in which theoretical formulas do not cover all effects in the real world, and thus this aspect might have influenced their attitudes more than those of non-VET participants.

Participants

The group of study participants who contributed to the case study data limits generalisation.

High share of participants from German UAS with high share of VET

Most of the analysed data is based on students from one German University of Applied Sciences enrolled in the same B. Eng. program "Electrical Engineering and Electric Mobility" (R1, R2, R6, and R9). A significant number of these participants had completed a VET program (subsection 3.2.1) before enrolment and the observed effect was particularly strong among these students.

The overall statistically significant difference detected in the first phase may be attributed to the high share of VET graduates in this part of the study, 59% of the students participating in the German B. Eng. 2017 run (R2) had completed vocational training. Due to the comparatively small number of participants in the internationally mixed runs (46%, R3-R5) compared to the German runs (54%, R1, R2), the overall outcome of the first phase comparing hands-on with simulated laboratories largely reflect results from German B. Eng. students. As fewer content areas were covered in the international runs, this limitation becomes stronger when evaluating weighted for tests.

Putting emphasis on the summarised evaluation of the different run types is not recommended, as participant characteristics differed based on other data sources. Differences were found between international and German B. Eng participants in the Amount of Practical Experience (see Table 4.3). Nevertheless, these other runs with mixed student background showed a slight trend towards better knowledge acquisition with hands-on laboratories as well.

UAS non-VET students vs. German traditional/full university non-VET students
The situation at UAS is extraordinary compared to traditional universities in Germany. Students with the best grades in standard secondary education usually choose traditional universities (for example studying human medicine, law, etc.) instead of studying at a UAS. In the Bavarian school system in particular, school selection in the early years (often based on marks in elementary school, see Figure 3.1) determines a student's future. Thus, the results are biased and cannot be generalised for full German universities. The presented data naturally compares the groups of students enrolled at UAS.

VET vs. non-VET
It needs to be noted that the outcomes addressing VET vs. non-VET are based on the test results of only one research run (R2, 30 students), as the other runs of the first research phase dealt with internationals (R4-R5) or did not collect VET information (R1). More data collection is necessary to allow for generalisation on German VET student behaviour.

Gender of participants
In all runs the majority of participants were male (>75%). Results are therefore not generalisable to student populations with a higher ratio of female participants.

International participants
Some data based on international programs with a good mixture of student backgrounds (engineering disciplines and home countries). On the one hand, the participants of the summer school runs arrived in Germany shortly before the conduction of the labs, thus the participants were in a special phase of cultural adjustment. On the other hand, there is no denying the participants were in an extraordinary situation, as they participated in the study during a stay in a foreign country. The groups have to be seen pre-filtered, for reasons like financial background to be able to visit Germany and pre-selection of home universities and THI based on study marks.

Another limitation that needs to be considered is the participants' motivation. Students from different study runs participated in different learning contexts (normal study semester vs. summer school). As the international programs were obviously visited by motivated students who took it upon themselves to travel to Germany for their studies, their motivation to get the most out of the learning experience was possibly higher compared to students from the other groups. While the participants' motivation was not controlled for, the mean scores and standard deviations of their test results do not indicate substantial differences between international students and the German groups. In addition, no extrinsic incentives were offered to the participants, so it is reasonable to conclude that their motivation was based on their own interest in the subject.

Hints which support generalisation in a limited way

Additionally, several hints were found which support generalisation in a limited way.

- When comparing participants' answers to the published difficulties in the RIASEC test manuals, they had "normal" values (for students, age etc.) in all categories, which gives a hint that the participants can be similar to the average German students.

- In the survey for participants' and other persons' subjective opinions on the learning modes (DS-E) a clear tendency towards better opinions regarding hands-on experiments was found for all groups (including data from traditional universities and countries other than Germany). Furthermore, the answers showed stronger trends among the groups which were not the major share in the main study with objective results. When assuming that these subjective opinions correlate with student learning, it is a hint that these outcomes could be somewhat generalised, for example for employed engineers.

    1. Participants who were still enrolled differed, as they tended *less* strongly in the direction of hands-on experiments compared to the other groups (teachers, engineers, etc.). This could be due to the age or role at the university.
    2. For countries other than Germany, a stronger trend towards hands-on was detected in two items (pair 1 and Q2), while the rest of the items did not differ statistically significantly.
    3. No major differences between UAS and traditional/full university students were found.

    As in the main study, most data came from German students, effects (when generalised) might be stronger or equal for other groups. Following this argument, it needs to be considered that subjective (data of this chapter) and objective data (main study) should be compared only very carefully, as of being of a different nature.

- With regards to the age of the students, the electric mobility bachelor program is representative of the common electrical engineering bachelor programs at THI.

Chapter 7

Conclusions and recommendations

7.1 Conclusions

First research phase (hands-on vs. simulations)

In several study runs in 2016 and 2017 that involved a total of 129 engineering students, test results related to knowledge acquisition as a result of conducting laboratory exercises in different modes were collected. A counterbalanced within-subject research methodology was applied. It was focused on the comparison of two laboratory modes: hands-on and simulation. In order to reduce interference, a diverse range of other variables was controlled for in both groups, such as the accompanying lectures, experimental instructions, teachers, learning objectives, tests, and many others.

Across all study runs of the first phase, students' test performances differed statistically significantly between the modes (R1-R5, $t(371) = 2.33$, $p = .021$, Cohen's $d = .24$). Students learned more effectively while engaged in hands-on laboratories compared to simulated laboratories.

Vocational Education and Training (VET)

The results of the German 2017 study run (R2) indicate a significant impact of the laboratory mode on the students' performance, depending on whether a participant had completed a VET program or not. In the hands-on condition, German participants who had completed such a program before university enrolment performed similarly well to their peers who enrolled directly after high school. In contrast, simulated experiments had less favourable effects on the learning of VETs than on that of non-VETs.

This finding suggests a need to investigate the utilisation of laboratories depending on the students' educational and industry background (VET). It is possible that engineering educators need to consider offering students different approaches to laboratory experiments based on individual prior learning and practical experience.

More data collection is necessary to come to a definitive conclusion on the reasons for these differences.

Second research phase (hands-on vs. hidden simulations)

Significant differences between the influence on learning of simulation and hands-on laboratories in the local domain were discovered in the first phase of the study. Hands-on laboratories produced better test results, especially for former VET-participants. The second phase of this study aimed to validate the first phase and examine the role of students' perception of the laboratory condition on their learning success.

This time, the learning environment and equipment used was the same in both conditions ("hands-on" and "hidden simulation"). Subjects involved in simulated laboratories were led to believe that they were conducting hands-on experiments. Under these circumstances, no significant differences were found in students' test results.

VET behaviour differed depending on the learning mode

After analysing all results from knowledge tests of the two research phases, one main conclusion can be drawn: If a student completed a German VET before enrolment, *perceived* simulated experiments are not as helpful as *perceived* hands-on experiments for learning in laboratories. Other groups also tended to learn better with hands-on experiments, but differences between modes were minor. Generally, no major differences between the modes were found.

When teaching former VET participants (who make up a large share of students at German UAS [28]), educators need to be aware of this aspect. Using simulations led to disadvantages for the average former VET participant; after hands-on experiments, this group showed similar results in the knowledge tests compared to their peers who did not have this educational background. No group showed significant disadvantages after using hands-on experiments.

7.1.1 Methodology

This study showed that the described methodology is applicable to focus on the comparison of two laboratory learning modes. With the instructions and learning objectives being identical and avoiding to change cooperative learning effects, results of student learning in the different modes were less influenced by interfering factors. The full list of excluded influences is given in the appendix on page 239.

A case study was performed, and while no significant outcome could be derived for overall generalisation, a specific outcome was identified: German UAS students with a complete VET education seem to tend much more towards better learning with hands-on experiments, while students without such an education behave differently. This seems to be a particularity of that particular group of students also when compared to international students.

Identification of relevant subgroups

Nonetheless, this outcome gives hints that many scientific studies may miss trends by not identifying certain properties causing the differences. For example, if the collection of VET data were not included in this study, the most significant outcome would have been missed. This leads to the question whether results of other studies where similar learning was demonstrated for different modes could be explained by a lack of subcategorisation regarding (relevant) subgroups in their setting. Crucially, this question is valid for any study, including the present one. It is possible that another unidentified subgroup or type of student produced a strong trend.

When designing similar studies, relevant subgroups need to be identified from the start to allow for individual analysis of the test results. The information regarding VET was not recorded during R1. This missing information lowers the study quality.

General influence of the learning modes

The study does not determine the best learning mode for battery lessons specifically; rather the goal was to evaluate the global influence of the learning modes on laboratory learning. Therefore, it is essential to mention that nearly identical experimental procedures were used in both modes. It is expected that the results would differ if the conditions were optimised separately according to their educational potential, for example using time-lapse technology in simulation or training optimal wiring in hands-on experiments.

Influence of physical interaction was nearly excluded
It is reasonable to assume that differences in test results between laboratory conditions in phase one would have been even higher if the experimental procedure had included physical interaction. However, this would have altered the hands-on condition, making it impossible to directly compare results of hands-on and simulated laboratories.

The strict methodology

On the one hand, the new methodology can be considered successful, as it was specifically designed to deliver more generalisable outcomes regarding the learning mode. By excluding other influences, the effect between the modes was minor (compared to many studies in literature, not avoiding these influences). It seems that only a small effect of the selected modes on student learning can be found.

On the other hand, the outcome still cannot be generalised. The overall outcome was weak to neutral, and a stronger trend was only detected when comparing specific subgroups. That is the nature of a case study. Study limitations were already discussed in detail in section 6.2. In order to get a generally applicable answer to the question which of the compared learning modes is more fruitful in teaching students in laboratories, more data collection is necessary.

German VET outcomes are also relevant in the international context
For the German engineering education research community, the study provides strong evidence that the results of scientific studies generally need to differentiate UAS students more precisely between “German-UAS-VET” and “German-UAS-non-VET”.

At the same time, the percentage of participants who completed a German VET program before university enrolment was high, suggesting that further research with students from diverse educational backgrounds is needed to make more accurate conclusions on the efficacy of hands-on and simulated laboratory modes. One of the important aspects might be the identified trends in the “Realistic” and “Investigative” RIASEC categories.

With German UAS students, former vocational training influenced student learning: an aspect which requires investigation in an international context. It is likely that the graduates of systems comparable to the German VET show the same bias in other countries.

Changes/enhancements to the methodology during the research

The methodology employed in the case study was improved/enhanced between the runs. Based on the outcomes of earlier runs, additional items (e.g. VET) and questionnaires (e.g. RIASEC) were included, developed, and tested. That is a natural process of research. Thus, some data sources were available too late to correlate them with results of the first research phase – the phase where the significant differences between modes were identified. For example, an analysis of the possible influence of the opinion on learning modes *before* conducting the experiments on student learning was not possible.

Although not all data sources were as fruitful as the knowledge test results, absent differences between analysed groups also give insight towards understanding the overall results better.

Recommended improvements regarding the Moodle questionnaire (DS-D)

Asking the participants to rate the difficulty of the previous experiment was not particularly fruitful. Requesting a description of the laboratory difficulty in a free text answer may help to gain more in-depth insight.

The same is true for the question regarding new insights during laboratory work. Identifying the precise insights students gained may be beneficial for the improvement of the experiments. Therefore, future surveys should also ask for the specific insights gained, and whether students consider these insights useful outside the university.

Having the questionnaires enhanced to investigate which content students missed, and which content they estimated as crucial for their future profession, would allow for adapting the experiments to increase relevance in the following runs and thus improve knowledge transfer.

Additionally, it might help to collect feedback from industry-professionals in an additional external questionnaire in order to improve the learning objectives and lessons. Mentioning this procedure in the laboratory instructions might also increase the (real and perceived) relevance for students.

Group creation using the "Amount of Practical Experience" dimension (DS-B)

In this study, cohorts of the German runs were split based on the Amount of Practical Experience (APE) value (please see for details to Table 4.4), as a correlation between APE and superior learning mode was assumed. The target was to create two comparable groups with similar group interaction.

Nevertheless, it was found that the correlation between VET and the superior learning mode was much stronger. The APE had an influence on the superior mode, but only *within* the VET and non-VET groups. Thus, for future studies, it is recommended to prioritise this item for the group split or, if no information can be collected, perform a fully randomised study. As a stable connection between VET and RIASEC "Realistic" was identified, further research would be necessary to base the group split upon the "Realistic" dimension.

The discussed methods used to distribute the students can also be applied in other situations. It might be feasible to provide students with the learning mode that is most suited to their predisposition.

Evaluation of the knowledge tests: superior learning mode (weighted per student) and test performances (weighted per knowledge test)

The results of knowledge tests (DS-A) were evaluated from two perspectives:

1. Comparing the average group results, based on t-tests comparing the two conditions weighted by the number of evaluated tests. This method had the advantage that runs with higher test density also counted more heavily in the overall result. Nonetheless, as mentioned before, the overall calculation should be doubted because of the differing results found for different student groups and characteristics. Thus, changing the number of students of specific types within the studied group would yield different results. However, this method allowed for the inclusion of students who had not participated in all the tests correctly.

2. Calculating the students' *superior learning mode* by comparing individuals' average test results of both modes, a single value for each participant expressing the individual tendency could be defined and allowed analysing correlations with other dimensions (e.g. APE). The disadvantage of this perspective on the data was that, if a student missed a test in the single-crossover-runs, their data were lost entirely. In the case of double-crossover-runs, some tests were intrinsically weighted double.

Overall results of both methods lead to the same outcomes. This similarity supported the overall outcomes and also justified the effort undertaken to gain a double view on the data.

Using knowledge tests to measure the success of laboratory teaching

Lindsay [13] reviewed laboratory protocols and thereby determined the quality of the work and focus of the students in the respective conditions. This method of quality control can be misleading, as laboratory protocols can be copied/modified by the participants to reduce their work load. To avoid this, findings of this study are based on written knowledge tests, which were performed at the beginning of the subsequent lesson. However, the setup of these knowledge tests might have also influenced students' learning success for two reasons: First, students were aware they would be tested for the study (although participation was anonymous and scores had no influence on their marks for the class). Thus, they may have been better prepared for subsequent lessons. Second, these tests made their knowledge gaps conscious. This might have motivated them to fill the gaps by asking questions in subsequent lessons. This possible influence of (even anonymous, unmarked) testing on learning success was not controlled for in the study, and might be an interesting point for further research.

7.1.2 Recommendations for educators, selecting laboratory modes

It was found that the learning environments which seemed to be connected to real-world implications (action and reaction were perceived "hands-on" in *hands-on* and *hidden simulation*) produced statistically significantly better student learning.

Therefore, results suggest opting for hands-on laboratories when deciding between hands-on experiments and *identically performed* (not necessarily optimised) simulated experiments for two reasons:

1. The hands-on condition resulted in better test scores.

2. Simulations seem disadvantageous to students from VET educational background.

However, the overall effect sizes were small.

A more general result of this research is that specific subgroups of students (e.g. VET) can have specific preferences (e.g. for laboratories conducted *hands-on* rather than simulated). These preferences may depend on the types of students in a class (see [71]) and can have a statistically significant impact on learning success.

The findings suggests a need to investigate the utilisation of laboratories depending on the students' educational and industry background. It is possible that engineering educators need to consider offering students laboratory experiments in different modes, depending on their prior learning and practical experience.

Using simulations to replace hands-on experiments

The solicited effect when excluding the influence of interface, teaching materials, learning objectives/experimental approach, supervision, learning synchrony and location was statistically significant but small.

Given that universities often struggle with resource restrictions, the small size of the disadvantage of the simulated mode detected in the case study permits the use of the simulation variant, e.g. for unexpectedly large class sizes.

In the questionnaires, participants favoured locally conducted experiments conducted in groups, even when they were given the option to remotely conduct simulated experiments at any time and place.

Depending on circumstances (like have to teach remotely or on a low budget), educators may be unable to perform physical experiments with their students. To support practical learning, educators should consider simulations instead of dropping experiments entirely. The same is true for part-time study programs, often completed while working full or part time, with reduced attendance at the university. Here, voluntary laboratories to be conducted at home or the workplace can offer students opportunities to improve their understanding. This is much more easily, both economically and logistically, achieved with simulations than with hands-on laboratories. A recent (and extreme) example is the closing of schools and universities during the COVID-19 pandemic in 2020 and 2021.

The employed simulations were not optimised
The small disadvantage of the simulated mode must be seen under an additional perspective: For the study, desired learning outcomes were selected in a way to be equally achievable in both modes. For example, things like the movement of electrons and ions in the battery cell were not included in the laboratory. Furthermore, to ensure equal experiments, possible advantages of simulations such as time dilation were not used. Examples for learning outcomes which would benefit from time-manipulation during experimentation include those of experiments B and D (pages 301 and 322 in the appendix).

Relevance and authenticity
In the surveys, doubts were mainly expressed about the authenticity and accuracy of simulated experiments, while hands-on experiments were considered more relevant. From the overall outcomes of the study and secondary data sources, it becomes clear that demonstration/explanation of the relevance of experiments, e.g. by the instructor, is essential for the success of particular groups. This could be more easily achieved by experimenting hands-on. Thus, when experimenting with simulations, efforts to create relevance and trust in the experimental results have to be more intense.

Is there practical relevance for teaching with hidden simulations?

It needs to be pointed out, that, besides research purposes, there is no practical advantage to tricking students into thinking they are using devices which they are not. As an example, consider a combination of cost-saving simulations with cheap "fake" devices (only a display).

1. It can be costly to create simulations that not only to mimic the identical experimental outcome, but are practically indistinguishable from real experiments. The "fake" devices would probably not be extraordinarily cheaper than real ones. The cost advantage simulations have over real experiments would be nullified.

2. The chosen topic “batteries” was supportive for hiding simulations, which many other fields aren’t. While it did not occur in the study, it is conceivable for students to uncover hidden simulations, which may destroy trust in the teacher, and in future laboratories. Based on the results of this study, perception matters, so casting doubt on authenticity/accuracy might lower students’ learning success for the duration of their studies, and can directly affect their results.

Based on these points, it can be concluded that hidden simulations have no practical relevance and should not be used to educate students.

7.2 Summary of contributions

In this thesis, the efficiency of two laboratory learning modes – hands-on and simulated experiments – was investigated, delivering the following key contributions:

- The key, novel, aspect of the study is the cross-over like methodology using the exact same setup and teaching material. The gained results are of theoretical interest to identify the potential differences in learning results.

  The presented methodology was focused on the learning modes of the laboratory experiments. The target was a comparison of the success in student learning after the same laboratory-experiment, executed either in reality or in a simulation.

  In actual teaching situations, educators would adapt the learning activities to take advantage of the benefits of a particular learning environment. This is also the approach considered in most of the existing literature (refer to section 2.7), where researchers replaced and compared existing hands-on experiments with newly created simulations. They improved both experiments (hands-on and simulated) independently to achieve optimal learning in the respective mode and situation. For example, students learned in groups in the university (hands-on), but alone at their work place (e.g. using simulations). As the aspect “learning modes” was mixed with other influences (in this case supervision, cooperative learning effects, distance learning, instructional papers), this research compared combinations of certain aspects, but could not isolate the influence of the laboratory learning mode.

  For this research, the laboratory experiments were developed in a way that every step in the students’ experiment was identical (see chapter B). This allowed for differences in student learning to be solely attributed to the use of either hands-on devices or simulation.

  The designed methodology allows a repetition with other applicable learning topics from other disciplines.

- The research concerns the particular situation in Germany, where an occupational education program (Vocational Education and Training, VET) is institutionalised as an option for higher education. Former VET participants, enrolling in STEM subjects after their VET, were analysed separately, giving hints on their unique needs and preferences in laboratory teaching.
- The research was enhanced by a second phase, supporting the equivalence of the employed simulation model.
- In contrast to Euan Lindsay's study [13],
  - A learning subject (battery cells) was used, where even in the hands-on mode most learning objectives are not tangible and audible (e.g. current, voltage).
  - A crossover approach was employed to bias the data to be able to exclude influences from group creation (see subsection 4.2.1).
  - Group creation and group interaction were excluded from influencing the study.
  - Lindsay's study, evaluating students' laboratory protocols, delivered insight into the student's change of attitude depending on the laboratory mode the experiments were held in. In this study, objective tests were employed to measure learning success.
- The developed laboratory sessions (refer chapter G) might not only be an interesting base for university teachers to develop own battery laboratories:
  - Many young people in Germany take part in the dual system of vocational training (for background knowledge see subsection 3.2.1). Schools for VET participants teach basic knowledge and particular information that not every company (like a car workshop) can instruct. These practice-orientated schools have a high demand for ready to use school units including teaching materials.
  - Even further education outside the university system, for example to become a foreman or technician, is possible in Germany. These courses offer a natural continuation of VET-affiliated schools at a higher theoretical and practical level.
  - The developed practical experiments and equipment can be used at other educational institutions, such as training departments of industrial corporations.
- Since 2015, the candidate constructed the hardware and firmware of the student battery test system, which is described in section F.2. The devices were optimised and qualified while conducting the PhD-research. The research of the second phase would have not been possible without the option to reprogram parts of the firmware freely.

- The thesis delivers difficulty values which allow other researchers to compare their participants to German UAS students, both with and without VET degree, for well-known question sets (subsection 4.4.4 and subsection 4.4.3).
- The thesis delivers objective statistical data about VET and non-VET students at German UAS, which was not available previously. This data allows comparisons by other researchers with their own statistics.
- A new question set to determine the practical experiences of STEM students (APE) was developed (see section 4.3) and validated with electric mobility students (see section 5.3). APE shows reliability in that target group. This new dimension was employed to create equivalent group environments for the relevant runs of the research investigation.
- The presented work delivers statistical RIASEC data about German UAS students. As the employed questions sets are standardised, the results allow other researchers to compare their participants/students with those from this study.

Chapter 8

Future work

Students' perception of a real existing specimen in a laboratory environment produced better understanding/knowledge retention. Nevertheless, the effect strength depended on the students' past: Learning experiences gained through a VET program are often practical. It was of particular interest how participants from such a background were affected by laboratory modes that differed in their factual as well as perceived relationship to the physical world. As the absence of a real device and a real battery was perceived differently by students with and without VET degree, a differential approach for teaching laboratories depending on the students' background and its implications for educators should be considered for the future design of learning laboratories.

Given the increasing push towards simulated and remote experiments [11], research on the efficiency of the learning modes is a very timely topic.

The key aspect of the study was to present and verify a new methodology similar to a crossover trial to evaluate the effectiveness of different learning modes. In the performed case study, computer-simulated laboratories were evaluated in comparison with hands-on exercises using a battery basics practical course. The methodology was able to contribute to new insights. Data sources were added and require a rerun of the study to elaborate the correlation of these data sources with the first study phase.

During the process and evaluation of this study, different aspects which require further research were noted:

Further studies for different student types

It is recommended that further studies utilising the counterbalanced research methodology should be carried out in other engineering fields with different learning objectives to thoroughly validate the methodology and the results reported in this study.

Furthermore, it is necessary to apply the methodology at other types of higher education institutions, as results show that the educational experiences and background of the students (e.g. the completion of a vocational education program before studies or practical experiences before enrolling in the study program) have influence on the study outcomes.

One of the most exciting outcomes of this study was the difference for students from a VET background. This topic certainly requires further investigation. Future studies should elaborate on the role of the Amount of Practical Experience, RIASEC results, and VET in student learning. It is particularly recommended to investigate the aspect of vocational training in an international context and with students at traditional German universities.

The educational pathway of former VET participants to enrolment at a university can differ a lot (see chapter 3). It is of interest to see if/how these individual pathways influence university learning.

Excluded factors (held constant in the case study) might have generally more influence on student learning than the learning mode

Based on the stronger effects found in the literature (see section 2.6), when compared to the small overall effect size detected in the case study comparing the learning modes, some of the excluded factors in this study might have a more significant impact on students' learning than estimated before.

The interference of the mode-effect with one or a combination of these factors could be an explanation for the varying results regarding the learning mode in non-isolated research.

Much remains to be done to obtain an overall impression. Just as this study focused on the learning mode by excluding influences, it is recommended to investigate the quantitative impact of those other factors (listed in chapter B) with a similarly intense focus on a single one.

Two factors are of particular interest:

Influence of group interaction

The APE dimension (DS-B) was utilised in this study to distribute students equally into the crossover groups. This method controlled for an overall similar group environment during the experimentation. VET had more impact on the superior mode compared to APE, while APE was valid *inside* the VET/non-VET categories.

It is recommended to conduct further research on this topic, for example by arranging the crossover groups as well as the small working teams in a way that these consist of former VET participants only, while others are made up of students without VET degree. Doing so will assure that team interactions depend on the different behaviour and might allow further insights into the influence of VET.

Learning environment and location

In the study at hand, hands-on experiments were conducted in a scientific laboratory environment, while the simulations were performed in a computer pool. In the laboratory, students had to stand, while in the computer pool the students sat at desks. When an educator needs to select the learning mode, these circumstances are correlated. Thus, for this study, these arrangements were seen as a part of the learning mode, and could not be separated.

It would be interesting to conduct a similar study investigating the influence of working position (see [117]) and environment of students on the success of simulated laboratory learning. It might be, that a more "scientific" environment (like a chemistry laboratory) leads to a different perception of simulated experiments than a computer pool.

Further investigation of the reasons why non-VET and VET behaviour differed

In the first research phase, it was found that local domain hands-on laboratory experimentation was more beneficial for student learning compared to simulations in the local domain (DS-A). For students who completed a VET before joining university, the hands-on mode lead to better test results. The second research phase helped acknowledge these outcomes. It supported the importance of the perceived learning mode for the VET group. There was an observable difference in behaviour between VET and non-VET students in otherwise identical situations.

On the one hand, the Moodle (DS-D) and Qualtrics (DS-E) survey might be able to explain the general trend towards hands-on, as the results suggest a psychological bias towards hands-on. On the other hand, no strong reasons were detected which would explain the differing behaviours of former VET and non-VET students at university. In the Qualtrics survey (DS-E, section 5.6), both participant groups (non-VET and VET) rated hands-on learning better than simulations. Most learners believe in the superiority of hands-on compared to in-person simulations. Looking to the rated differences between both modes for both groups, no trend was detected comparing VET and non-VET participants. It would be interesting if the correlation between perceived relevance and proven test performance is stronger for VET after simulated experiments. Unfortunately, the data collected in the study did not allow for that evaluation.

Analysing the university data (DS-F), some statistically significant differences were found between VET and non-VET (e.g. grades). However, these differences were practically so small that a relevant bias can be doubted, which supports further investigation of psychological reasons for the effect.

The correlation of APE (DS-B) and the superior learning mode (DS-A) showed the weak trend that students with high APE tend towards better test results after *simulations* – separately for both VET and non-VET (see Figure 5.16, Spearman rank order correlation $\rho \approx .5$, $p \approx .10$), with a bias between groups. That slope was similar in both groups and also cannot explain the significant difference regarding the learning mode between both groups.

On one hand, the difference in the RIASEC outcomes (DS-C) between VET and non-VET (VET being more "Realistic" and less "Investigative") might describe the two groups well. On the other hand, the nature of employed question directly describes the past of these groups.

Further study is required to determine why the identically negative perception (DS-D, DS-E) of one mode (*perceived* simulated experiments) *had an effect* on student learning for students with VET-degrees, but *not* for non-VET students. In order to comprehend the effectiveness of the different laboratory modes more fully, psychological effects should be considered. Perhaps the absence of a physical device and/or a real battery were perceived differently by students with and without VET degree.

In the following subparagraph, *one* of the possible explanations is discussed further.

"Acceptance" of theoretical results without practical verification
Students can see simulated results as resulting from models purely based on theoretically valid formulas. If participants doubt the underlying paradigm, they also question the experimental results of simulations in general, and they might need to invest more time/effort to "accept" the validity of externally given theoretical explanations without practical verification.

Such behaviour (distrusting theory/formulas) is not very compatible with the school system. For example, when calculating the kinematics of a cannonball fired down a hill in physics, the "right solution" in schools is usually based on simplified formulas. However, real-world application can depend on more than these calculations: For example, when a VET-participant calculates the necessary diameter of a wire for a specific current in school, he/she might experience (in the company or at a customer's site) that this wire – which was theoretically correct – is not wide enough, as it is too small to fit tightly into the required clamps. For this situation, the theoretical formula was not helpful. Non-VET are less likely to encounter this type of problem before joining the university (see chapter 3).

Thus, the slight disadvantages of simulations for former VET-participants may be explained by opinions towards the theoretical models taught in school.

Several possible reasons come to mind:

- Students without a VET degree probably have gained easier acceptance of theoretical results due to their intense exposure to theoretical paradigms, which large parts of the knowledge taught in schools are based on.

- Another explanation can be that this higher inclination towards theoretical schooling of non-VET is a personality characteristic – already present *before* deciding on a particular secondary school education – and perhaps one of the reasons for the decision. Persons with lower acceptance for theoretical teaching/learning tend to choose the more practical VET education, while persons who more readily accept theory are more likely to succeed in standard secondary school education.

- VET participants have had more chances to experience differences between simplified theories and complex observable properties, which in turn might have lowered their acceptance of simulations, seen as a simplified model of reality.

- The frequent use of tools and measurement devices during their vocational training could have been habit-forming for VET students, which positively influenced their learning using real-world tools.

The study's data does not allow for the extraction of the reasons for the particularities exhibited by VET students and can not provide a conclusive answer to these questions. It is recommended to design appropriate studies to investigate these reasons further.

Of course, other explanations for the differences in VET behaviour are possible. It is conceivable, for example, that a very general trend exists for VET to be more easily influenced by perceived settings.

Holland's RIASEC-typology

Holland's RIASEC-typology showed clear trends regarding VET. Former VET participants had very high "Realistic" values. Their "Investigative" values were lower than to their non-VET colleagues'. Unfortunately, these question sets were not used in the first research phase. Thus, correlations with a better performing learning mode were not possible.

As the topology is commonly available worldwide, it would be natural to continue the research based on these questions sets. It is possible that the "Realistic" and "Investigative" categories are correlated with the preference of a learning mode.

Depending on the question sets, it might be necessary to increase the difficulty of certain scales to avoid saturation. For example, when using standard sets with university level engineering students, high I- and R-values are to be expected, which frequently led to saturation effects in the present study.

Students' preconceptions about the learning modes

Unfortunately, opinions on the learning modes (DS-E) were not collected from DS-A-participants until late in the study. Response rates (esp. ones including code-words) were low. Therein, only four former VET participants' footnotes were available for direct correlation. Thus, analysing whether students' opinion on learning modes *before* the experiments caused differences in students' test performance by calculating direct correlations was not possible in this study. Assuming such a correlation exists, it would be very interesting to test if this correlation was stronger among VET than non-VET students.

Correlations between general test performance after conducting laboratories and APE

The APE dimension only correlated with the test performance of some groups of participants. In this dimension, German UAS and international groups showed opposite behaviour (see section 5.3). With German UAS students a high APE is connected to better performance, while opposite effect was detected for international students (see 113).

It might be interesting to establish a research study to further investigate these reasons. It would also be very relevant to include participants of German traditional university students in that research. The same goes for different cultures/nations, as the “international” runs presented a very diffuse nature of participant.

The in-subject crossover-like methodology might reduce the effect size

Since the effect points towards better results in knowledge acquisition/retention when conducting hands-on laboratories compared to simulated ones, it is plausible to argue that the differences in performance between groups would have been more significant, had the participants been allocated to either the hands-on or simulated laboratories across all experiments (“Carry-Over-Effect”, see [114]).

This calls for a cross-over-less comparison between hands-on laboratories and simulated laboratories, which is needed to gain insight into the differences in test results and effect sizes if students only participate in one of the two laboratory conditions.

In that setup, instead of performing the cross-over between the laboratory sessions, each of the groups works in one of the compared modes only. This allows advantages and disadvantages to accumulate, possibly making effect sizes larger than in the present study. These effect sizes may be more relevant for teachers when deciding about the desired learning mode for their teaching laboratory.

There would definitely be some ethical concerns to be addressed with such a proposal. These concerns largely stem from the fact that participants are also students, and the previously collected results suggest a certain student group (the VET students, selected to conduct all experiments using simulations) might be disadvantaged by the proposed study format. However, this is a constant peril with research in educational fields.

Also, it would be difficult to convince other universities about these guest lectures: since simulations tend to be perceived as less encouraging for students when selecting additional activities next to their normal study load, universities may be hesitant to invite a guest program in which some students only engage in simulated laboratories.

Influence of the teacher/instruction on the perception of the authenticity of simulations and learning success

It seems that it does not matter, if students really “work” with real devices, as long as they think they control a real experiment [13].

The second phase also raised questions about the skill level required of students to understand the differences between real and simulated results in the first research phase. As discussed, students’ doubts on the accuracy and authenticity of the performed experiments and their results might influence the success of laboratory teaching.

On the general online questionnaire (DS-E) 70% of the negative free text comments on simulations expressed doubts on accuracy, and 20% of the positive free text comments on hands-on expressed in some way that hands-on experimentation is authentic, and includes nature and real life.

Participants with a former VET education generally had more doubts about the authenticity of laboratory experiments (Pair 2, DS-E).

The second phase had shown that these doubts on accuracy were a preconception. Simulations can model the necessary behaviour well, especially since the teacher knows the crucial points investigated by the students' experiments. These preconceptions might have resulted in the different success of the laboratory modes.

It would be interesting to investigate if subjective bias can be deemed to be the reason for the differing performance of the modes. Unfortunately, it was not possible to correlate the answers in the general online questionnaire to the test results in this study, as the online questionnaire was first conducted in the second study phase, when the simulation was hidden. This aspect should be investigated with further runs similar to the first research phase.

Assuming that such a bias against simulated laboratories influenced the differences in student performance discovered in this study, it is recommended that educators explain to students (in laboratory manuals or the parallel lectures), that the employed simulations mimic the behaviour of the real experiment so well, that it is practically impossible to discriminate between these two modes, before performing the simulated experiments.

The teacher needs to devote substantial time to convince the participants that the simulations are authentic/accurate – and can contribute to learning as effective as hands-on experiments, for example through explanation of the outcomes of the present study.

Of course, the effectiveness of that priming would have to be tested as well. For example, by priming only one of two groups and then testing for opinions on simulations. This can be checked using a questionnaire before starting another case study with knowledge-test (without crossover, all simulations) to analyse if educators can reduce the disadvantages of simulations by explicitly promoting authenticity and applicability.

Diving deeper, it should be researched how best to achieve this goal of convincing students of a simulations' authenticity.

Chapter 9

References

[1] European Credit Transfer and Accumulation System. https://ec.europa.eu/education/resources-and-tools/european-credit-transfer-and-accumulation-system-ects_en.

[2] A. Hofstein and V. N. Lunetta. The laboratory in science education. Foundations for the twenty-first century. *Science Education*, 88(1):28–54, 2004.

[3] G. Faucher. The role of laboratories in engineering education. *International Journal Mechanical Eng. Education*, 13(3):195–198, 1985.

[4] B. Boesch. Representing in the student laboratory. *International Journal for the Historiography of Science*, 5:34–48, Dec. 2018.

[5] C. Unseld and G. Reucher. University types: universities of applied science. http://dw.com/p/Ovf8, 2010. Accessed: 05.05.2010.

[6] J. Roschelle, R. D. Pea, C. M. Hoadley, D. N. Gordin, and B. Means. Changing how and what children learn in school with computer-based technologies. *The Future of Children*, 10(2):76–101, 2001. DOI: 10.2307/1602690.

[7] S. Höntzsch, U. Katzky, K. Bredl, F. Kappe, and D. Krause. Simulationen und simulierte Welten. Lernen in immersiven Lernumgebungen. In M. Ebner and S. Schön, editors, *Lehrbuch für Lernen und Lehren mit Technologien*. 2013.

[8] J. Dahn and G. M. Ehrlich. Lithium-ion batteries. In *Linden's Handbook of Batteries 4th Edition*, pages 26.1–26.79. McGraw-Hill, 2011.

[9] R. Heradio, L. de la Torre, D. Galan, F. J. Cabrerizo, E. Herrara-Viedma, and S. Dormido. Virtual and remote labs in education. a bibliometric analysis. *Computers & Education*, 98:14–38, Aug. 2016.

[10] J. R. Brinson. Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories. A review of the empirical research. *Computers & Education*, 87:218–237, 2015.

[11] S. M. Zvacek and M. T. Restivo. Guidelines for effective online lab assignments – contributions to the discussion. In *Proceedings of the IEEE Global Engineering Education Conference 2018*, 1442–1446, Santa Cruz de Tenerife, Spain, 2018.

[12] L. D. Feisel and A. J. Rosa. The role of the laboratory in undergraduate engineering education. *Journal of Engineering Education*:121–130, 2005.

[13] E. D. Lindsay. *The Impact of Remote and Virtual Access to Hardware upon the Learning Outcomes of Undergraduate Engineering Laboratory Classes*. PhD thesis, Department of Mechanical & Manufacturing Engineering, The University of Melbourne, Australia, 2005.

[14] J. Ma and J. V. Nickerson. Hands-on, simulated, and remote laboratories: a comparative literature review. *ACM Computing Surveys*, 38(3), 2006.

[15] J. Corter, J. Nickerson, S. Esche, C. Chassapis, S. Im, and J. Ma. Constructing reality: a study of remote, hands-on, and simulated laboratories. *ACM Transactions on Computer-Human Interaction*, 14(2), 2007.

[16] N. S. Edward. Evaluation of computer based laboratory simulation. *Computers & Education*, 26(1-3):123–130, 1996.

[17] E. McAteer, D. Neil, N. Barr, M. Brown, S. Draper, and F. Henderson. Simulation software in a life sciences practical laboratory. *Computers & Education*, 26(1-3):102–112, 1996.

[18] S. A. Engum, P. Jeffries, and L. Fisher. Intravenous catheter training system: computer-based education versus traditional learning methods. *American J. Surgery*, 186(1):67–74, 2003.

[19] V. Mathiowetz, C.-H. Yu, and C. Quake-Rapp. Comparison of a gross anatomy laboratory to online anatomy software for teaching anatomy. *Anatomical sciences education*, 9(1):52–59, 2016.

[20] C. J. Keller, N. D. Finkelstein, K. K. Perkins, and S. J. Pollock. Assessing the effectiveness of a computer simulation in introductory undergraduate environments. In *Proceedings of the Physics Education Research Conference*, Syracuse, NY, USA, 2006. AIP Conf. Proceedings.

[21] N. D. Finkelstein, W. K. Adams, C. J. Keller, P. B. Kohl, K. K. Perkins, N. S. Podolefsky, and S. Reid. When learning about the real world is better done virtually: a study of substituting computer simulations for laboratory equipment. *Physical Review Special Topics -– Physics Education Research*, 1(1), 2005. DOI: 10.1103/PhysRevSTPER.1.010103.

[22] C. Tsihouridis, M. Batsila, D. Vavougios, and G. S. Ioannidis. The timeless controversy between virtual and real laboratories in science education – and the winner is... In M. E Auer., D. Guralnick, and I. Simonics, editors, *Teaching and Learning in a Digital World*, pages 1539–1550. Springer International Publishing AG, 2018. ISBN: 978-3-319-73203-9.

[23] D. Weisman. Incorporating a collaborative web-based virtual laboratory in an undergraduate bioinformatics course. *Biochemistry and Molecular Biology Education*, 38(1):4–9, 2015. DOI: 10.1002/bmb.20368.

[24] C. Gravier, J. Fayolle, B. Bayard, M. Ates, and J. Lardon. State of the art about remote laboratories paradigms – foundations of ongoing mutations. *International Journal of Online Engineering*, 4(1):19–25, Feb. 2008.

[25] M. Ogot, G. Elliott, and N. Glumac. An assessment of in-person and remotely operated laboratories. *Journal of Engineering Education*, 92(1):57–64, Jan. 2003.

[26] R. Heradio, L. de la Torre, and S. Dormido. Virtual and remote labs in control education: a survey. *Annual Reviews in Control*, 42:1–10, Aug. 2016. DOI: 10.1016/j.arcontrol.2016.08.001.

[27] K. Siebertz, D. van Bebber, and T. Hochkirchen. *Statistische Versuchsplanung*. Dec. 2017. ISBN: 9783662557426.

[28] C. Bendtsen and F. Steger. The share of graduates of the German Vocational Education and Training System in tertiary education. https://www.researchgate.net/publication/333508340_The_Share_of_Graduates_of_the_German_Vocational_Education_and_Training_System_in_Tertiary_Education, 2019. Accessed: 19.06.2019.

[29] Royal Society. History of the Royal Society – The story of the Royal Society is the story of modern science. https://royalsociety.org/about-us/history/. Accessed: 30.12.2021.

[30] O. A. Soysal. Computer integrated experimentation in electrical engineering education over distance. In *Proceedings of the 2000 Annual ASEE Conference*, pages 5.164.1–5.164.10, St. Louis, Missouri, USA. American Society for Engineering Education, June 2000. DOI: 10.18260/1-2--8230.

[31] L. D. Feisel and G. D. Peterson. A colloquy on learning objectives for engineering educational laboratories. In *ASEE Annual Conference and Exposition*. ASEE, Montreal, Ontario, Canada, June 2002.

[32] P. A. Kirschner and J. J. van Merriënboer. Do learners really know best? Urban legends in education. *Educational Psychologist*, 48(3):169–183, 2013.

[33] G. Olympiou. *Development of a framework of combination of virtual and real environments experimentation in Physics*. PhD thesis, Cyprus University, 2012.

[34] M. D. Dunnete and J. P. Campbell. Laboratory education: impact on people and organizations. *Industrial Relations A Journal of Economy and Society*, 8(1):1–27, May 2008.

[35] K. Radin-Salim, M. Puteh, and S. M. Daud. Assessing students' practical skills in basic electronic laboratory based on a psychomotor domain model. *Procedia - Social and Behavioral Sciences*, 56:546–555, 2012. ISSN: 18770428. DOI: 10.1016/j.sbspro.2012.09.687.

[36] A. N. Wendt. Is there a problem in the laboratory? *Frontiers in psychology*, 9, Dec. 2018. ISSN: 1664-1078.

[37] S. Sullivan, D. Gnesdilow, S. Puntambekar, and J.-S. Kim. Middle school students' learning of mechanics concepts through engagement in different sequences of physical and virtual experiments. *International Journal of Science Education*, 6(2):1–28, 2017. ISSN: 0950-0693. DOI: 10.1080/09500693.2017.1341668.

[38] M. Koretsky, C. Kelly, and E. Gummer. Student perceptions of learning in the laboratory: comparison of industrially situated virtual laboratories to capstone physical laboratories. *Journal of Engineering Education*, 100(3):540–573, 2011.

[39] J. J. Chini, A. Madsen, E. Gire, N. S. Rebello, and S. Puntambekar. Exploration of factors that affect the comparative effectiveness of physical and virtual manipulatives in an undergraduate laboratory. *Physical Review Special Topics – Physics Education Research*, 8(010113):1–12, Apr. 2012.

[40] J. P. Atanas. Is virtual-physical-virtual manipulatives in physics irrelevant within studio physics environment? *Athens Journal of Education*, 5(1):29–43, 2018. DOI: 10.30958/aje.5-1-2.

[41] F. V. Evangelou and K. Kotsis. Real vs virtual physics experiments: comparison of learning outcomes among fifth grade primary school students. a case on the concept of frictional force. *International Journal of Science Education*, 41(3):330–348, 2019. DOI: 10.1080/09500693.2018.1549760.

[42] S. Dormido. Control learning: present and future. In *Proceedings of the 15th Triennial World Congress*, volume 35, pages 71–93, Barcelona, Spain. International Federation of Automatic Control, 2002. DOI: 10.3182/20020721-6-ES-1901.01641.

[43] S. Dormido. Control learning: present and future. *Annual Reviews in Control*, 28(1):115–136, 2004.

[44] N. M. Lima, M. C. Viegas, and F. J. Garcia-Penalvo. Learning from complementary ways of developing experimental competences // aprendizaje a partir de maneras complementarias de desarrollar capacidades experimentales. *Education in the Knowledge Society*, 18(1):63, 2017. DOI: 10.14201/eks20171816374.

[45] V. J. Harward et al. The ilab shared architecture. a web services infrastructure to build communities of internet accessible laboratories. *Proceedings of IEEE*, 96(6):931–950, 2008.

[46] J. Tuttas and B. Wagner. The relevance of haptic experience in remote experiments. In *EdMEDIAWorld Conf. Educational Multimedia, Hypermedia, Telecommunications*, Denver, CO, USA, June 2002.

[47] E. D. Lindsay and M. C. Good. Effects of laboratory access modes upon learning outcomes. *IEEE Transactions on Education*, 48(4):619–631, Nov. 2005.

[48] F. Steger, A. Nitsche, A. Arbesmeier, K. Brade, H.-G. Schweiger, and I. Belski. Teaching battery basics in laboratories: hands-on versus simulated experiments. *IEEE Transactions on Education*, 63(3):198–208, Aug. 2020. DOI: 10.1109/TE.2020.2970554.

[49] C. Lattemann, S. Stieglitz, and S. Korreck. Lernen in Virtuellen Welten. In H. R. Hansen, D. Karagiannis, and H.-G. Fill, editors, *Business Services: Konzepte, Technologien, Anwendungen. 9. Internationale Tagung Wirtschaftsinformatik Band 2*, pages 431–440. 2009.

[50] T. Kangasniemi, S. Olkinuora, P. Joensuu, O. Natri, P. Qvist, M. Ketola, J. Brusin, H. Virtanen, M. Närhi, R. Jokela, E. Palomäki, H. Tiitu, and K. Nordström. In *Proceedings of the 5th International Conference on Computer Supported Education, 2013*, pages 661–672. CSEDU, 2013. ISBN: 978-989-8565-53-2.

[51] P. Qvist. *LabLife3D – a science-based 3D virtual learning environment – designing virtual laboratories: from pedagogical design to learning analytics*. Bachelor's thesis, Turku University of Applied Sciences, 2015.

[52] K. Lee. The future of learning and training in augmented reality. *InSight: A Journal of Scholarly Teaching*, 7:31–42, 2012. ISSN: 1933-4850.

[53] J. Helmer. *Second Life and Virtual Worlds*. Learning Light Limited, 2007.

[54] M. Shimba, M. P. J. Mahenge, and C. A. Sanga. Virtual labs versus hands-on labs for teaching and learning computer networking: a comparison study. *International Journal of Research Studies in Educational Technology*, 6(1):43–58, 2017.

[55] L. Stern, N. Barnea, and S. Shauli. The effect of a computerized simulation on middle school students' understanding of the kinetic molecular theory. *Journal of Science Education and Technology*, 17:305–315, 2008. DOI: 10.1007/s10956-008-9100-z.

[56] C. Sarabando, J. P. Cravino, and A. Soares. Improving student understanding of the concepts of weight and mass with a computer simulation. *Journal of Baltic Science Education*, 15(1), 2016.

[57] S. L. Hannel and J. Cuevas. A study on science achievement and motivation using computer-based simulations compared to traditional hands-on manipulation. *Georgia Educational Researcher*, 15(1), 2018.

[58] J. C. Domínguez, R. Miranda, E. J. González, M. Oliet, and M. V. Alonso. A virtual lab as a complement to traditional hands-on labs: characterization of an alkaline electrolyzer for hydrogen production. *Education for Chemical Engineers*, 23:7–17, 2018.

[59] N. Lima, C. Viegas, M. Zannin, A. Marques, G. Alves, M. C. Felgueiras, R. Costa, A. Fidalgo, S. Marchisio, F. Lerro, C. Merendino, J. B. da Silva, M. I. Pozzo, E. Dobboletta, I. Gustavsson, K. Nilsson, and F. Garcia-Peñalvo. Do students really understand the difference between simulation and remote labs? In *Proceedings of the 5th Technological Ecosystems for Enhancing Multiculturality (TEEM17)*, Cádiz, Spain. Association for Computing Machinery, Oct. 2017.

[60] S. M. Aziz and T. L. J. Ferris. A psychomotor skills extension to Blooms's taxonomy of education objectives for engineering education. In *Proceedings of the iCEER Conference on Exploring Innovation in Education and Research*, 2005.

[61] J. D. W. Andrews. Discovery and expository learning compared: their effects on independent and dependent students. *The Journal of Educational Research*, 78(2):80–89, 1984. DOI: 10.1080/00220671.1984.10885578.

[62] S. W. Reichmann and A. F. Grasha. A rational approach to developing and assessing the construct validity of a student learning styles scales instrument. *Journal of Psychology*, 87:213–223, 1974.

[63] D. S. Domin. A review of laboratory instruction styles. *Journal of Chemical Education*, 76(4):543–547, 1999. DOI: 10.1021/ed076p543.

[64] L. Buck, S. Bretz, and M. Towns. Characterizing the level of inquiry in the undergraduate laboratory. *Journal of College Science Teaching*, 38(1):52–58, Sept. 2008.

[65] D. S. Domin. A content analysis of general chemistry laboratory manuals for evidence of higher-order cognitive tasks. *Journal of Chemical Education*, 76(1):109–112, 1999. DOI: 10.1021/ed076p109.

[66] J. M. Chamberlain, K. Lancaster, R. Parson, and K. K. Perkins. How guidance affects student engagement with an interactive simulation. *Chemistry Edu. Res. & Practice*, 15(4):628–638, 2014.

[67] N. M. Webb. Student interaction and learning in small groups. *Review of Educational Research*, 52(3):421–445, 1982.

[68] R. E. Slavin. Research for the future: research on cooperative learning and achievement: what we know, what we need to know. *Contemporary Educational Psychology*, 21(1):43–69, 1996. URL: https://pure.york.ac.uk/portal/en/publications/research-on-cooperative-learning-and-achievement(a7191aef-4dfc-45e3-876a-e8936490c5c6).html.

[69] N. M. Webb. Peer interaction and learning in small groups. *International Journal of Educational Research*, 13(1):21–39, 1989.

[70] M. Phillips, L. H. Gildensoph, M. J. Myers, C. G. Norton, A. M. Olson, K. A. Tweeten, and D. D. Wygal. Investigative labs in biology: the importance of attending to team dynamics. *Journal of College Science Teaching*, 37(2):23–27, Nov. 2007.

[71] D. Stamovlasis, A. Dimos, and G. Tsaparlis. A study of group interaction processes in learning lower secondary physics. *Journal of Research in Science Teaching*, 43(6):556–576, 2006. ISSN: 00224308. DOI: 10.1002/tea.20134.

[72] J. Van der Laan Smith and R. M. Spindle. The impact of group formation in a cooperative learning environment. *Journal of Accounting Education*, 25(4):153–167, 2007.

[73] H.-W. Tien, Y.-S. Lin, Y.-C. Chang, and C.-P. Chu. A genetic algorithm-based multiple characteristics grouping strategy for collaborative learning. In D. K. W. Chiu, M. Wang, E. Popescu, Q. Li, R. Lau, T. K. Shih, C.-S. Yang, and D. G. Sampson, editors, *Advances in Web-Based Learning – ICWL 2013 Workshops*. Volume 8390, Lecture Notes in Computer Science, pages 11–22. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. ISBN: 978-3-662-46314-7. DOI: 10.1007/978-3-662-46315-4{\textunderscore}2.

[74] A. Alaie. High-achieving postbaccalaureate-student teaching assistants. *Journal of College Science Teaching*, 37(4), 2008.

[75] S. Nikolic, T. Suesse, T. Mccarthy, and T. Goldfinch. Maximising resource allocation in the teaching laboratory: understanding student evaluations of teaching assistants in a team-based teaching format. *European Journal of Engineering Education*, 42(6):1–19, Feb. 2017.

[76] G. F. Gobaw and H. I. Atagana. Assessing laboratory skills performance in undergraduate biology students. *Academic Journal of Interdisciplinary Studies*, 2016. ISSN: 22813993. DOI: 10.5901/ajis.2016.v5n3p113.

[77] I. Abrahams, M. J. Reiss, and R. M. Sharpe. The assessment of practical work in school science. *Studies in Science Education*, 49(2):209–251, 2013. ISSN: 0305-7267. DOI: 10.1080/03057267.2013.858496.

[78] E. R. Sanders, S. Toma, A. M. Hirsch, and P. Y. Lee. Transforming laboratory education in the life sciences. *Microbe*, 11(2):69–74, Feb. 2016.

[79] B. Campbell, F. Lubben, S. Mathabatha, and M. Rollnick. Measuring university entrants' previous practical experience: how valid are students' self-reports? *African Journal of Research in Mathematics, Science and Technology Education*, 4(1):87–95, 2000.

[80] G. P. Baxter, J. Pine, and R. J. Shavelson. Performance assessment in science. *Applied Measurement in Education*, 4(4):347–362, 1991.

[81] Office of Assessment, University of Colorado Denver. Electrical engineering undergraduate program assessment. http://www.ucdenver.edu/about/departments/assessmentoffice/templates/Pages/default.aspx, 2015. Accessed: 09.04.2017.

[82] A. B. Hunter, S. L. Laursen, H. Thiry, and T. J. Weston. Urssa: evaluating student gains from undergraduate research in the sciences. council on undergraduate research. *Council on Undergraduate Research*, 29:15–19, 2009.

[83] D. Anastasio, H. Chenette, G. Neumann, and T. Ribera. Assessing student learning in chemical engineering laboratory courses: a multi-dimensional approach. *Assessment Institute: IUPUI*, 2018.

[84] Z. B. Razali and J. P. Trevelyan. Can practical intelligence from a laboratory experience be measured? *Proceedings of the 2008 AaeE Conference, Yeppoon*, 2008.

[85] Z. B. Razali and J. P. Trevelyan. An evaluation of students' practical intelligence and ability to diagnose equipment faults. *Procedia - Social and Behavioral Sciences*, 56:42–51, 2012. ISSN: 18770428. DOI: 10.1016/j.sbspro.2012.09.630.

[86] European Commission - EACEA National Policies Platform - Eurydice. https://eacea.ec.europa.eu/national-policies/eurydice/content/germany_en.

[87] Autorengruppe Bildungsberichterstattung. *Bildung in Deutschland 2018. Ein indikatorengestützter Bericht mit einer Analyse zu Wirkungen und Erträgen von Bildung*. Deutsches Institut für Internationale Pädagogische Forschung, 2018.

[88] Y. Shavit and W. Müller. Vocational secondary education: where diversion and where safety net? *European Societies*, 2(1):29–50, 2000.

[89] D. DiBiase. Rethinking laboratory education for an introductory course on geographic information. *Cartographica*, 33(4):61–72, Dec. 1996.

[90] Paragraph 35 to 39 bayeug, bavarian law on education and schooling.

[91] BMBF, German Federal Ministry of Education and Research. *Report on Vocational Education and Training 2016*. German Federal Ministry of Education and Research, 2016.

[92] BMBF, German Federal Ministry of Education and Research. *Dual Training at a Glance*. German Federal Ministry of Education and Research, 2013.

[93] C. Wieland. Germany's dual vocational-training system: possibilities for and limitations to transferability. *Local Economy*, 30(5):577–583, 2015. DOI: 10.1177/0269094215589318.

[94] H. Solga, P. Protsch, C. Ebner, and C. Brzinsky-Fay. The German vocational education and training system. Its institutional configuration, strengths and challenges, Berlin/Germany: WZB Berlin Social Science Center, 2014.

[95] BIBB Datenreport 2019 / Vorabversion, Federal Institute for Vocational Training (BIBB), Bonn, Germany, 2019. Accessed 2019-05-16.

[96] K. Hoeckel and R. Schwartz. Learning for jobs. *OECD Reviews of Vocational Education and Training Germany*, 2010.

[97] S. Dummert. Betriebliche Ausbildungsbeteiligung 2006 bis 2017 – Ergebnisse aus dem IAB-Betriebspanel 2. Mai 2018, Nürnberg/Germany: Institut für Arbeitsmarkt- und Berufsforschung, 2018.

[98] P. Protsch. Segmentierte Ausbildungsmärkte. Berufliche Chancen von Hauptschülerinnen und Hauptschülern im Wandel. Opladen: Budrich UniPress, 2014. DOI: 10.3224/86388050.

[99] Ausbildung 2018. Ergebnisse einer DIHK-Online-Unternehmensbefragung – Unternehmen berichten über ihre Ausbildungssituation, Deutscher Industrie- und Handelskammertag, 2018.

[100] A. Hall and E. Krekel. Erfolgreich im Beruf? Duale und schulische Ausbildungen im Vergleich, Bonn, Germany, 2014.

[101] M. Baethge. Das deutsche Bildungs-Schisma. *SOFI-Mitteilungen*, 34:13–27, 2006.

[102] Hochschulkompass: Hochschulen in Deutschland, Offered by German Rectors' Conference. https://www.hochschulkompass.de/hochschulen.html. Accessed: 09.03.2019.

[103] B. Schmidt, A. von Hippel, and R. Tippelt. Higher education evaluation in Germany. *Research in Comparative and International Education*, 5(1):98–111, 2010.

[104] Qualifizierungsangebote im Bereich akademische Bildung. http://www.nque.de/de/datenbank_akademisch.php, NQuE Netzwerk Qualifizierung Elektromobilität. Accessed: 10.01.2018.

[105] M. Baethge, C. Kerst, M. Leszczensky, and M. Wieck. Zur neuen Konstellation zwischen Hochschulbildung und Berufsausbildung. *Forum Hochschule*, 3, 2014.

[106] VET in Europe. German Country Report 2016, 2016.

[107] J. Powell and H. Solga. Why are participation rates in higher education in Germany so low? Institutional barriers to higher education expansion. *Journal of Education and Work*, 24(1):49–68, 2011.

[108] N. Kracke, E. Middendorff, and D. Buck. *Beteiligung an Hochschulbildung, Chancen(un)gleichheit in Deutschland (DZHW Brief 3/2018), Updated Version*. Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), Hannover, 2018.

[109] Bundesagentur für Arbeit. Statistik der Bundesagentur für Arbeit: Berichte: Blickpunkt Arbeitsmarkt – Akademikerinnen und Akademiker, Nürnberg, 2019.

[110] Statistisches Bundesamt. Verdienste auf einen Blick, 2017.

[111] B. Christoph, L. U., and H. Stübner. Höhere Abschlüsse zahlen sich mit dem Alter zunehmend aus. Aktuelle Analysen aus dem Institut für Arbeitsmarkt- und Berufsforschung. *IAB Kurzbericht*, 13, 2017.

[112] Gehalt.de/Gehaltsvergleich.com. Gehaltsbiografie 2019. https://cdn.gehalt.de/cms/Gehaltsbiografie-2019.pdf, 2019. Accessed: 19.09.2019.

[113] B. Jones and M. G. Kenward. *Design and Analysis of Cross-Over Trials*. Chapman and Hall, London, 2014.

[114] T. J. Cleophas. Carryover effects in clinical research. In T. J. Cleophas, editor, *Human Experimentation*, pages 25–36. Springer, Dordrecht, 1999. ISBN: 978-0-7923-5827-5. DOI: 10.1007/978-94-011-4663-0_4.

[115] L. Cohen, L. Manion, and K. Morrison. *Research Methods in Education (6th edition)*. Routledge Falmer, London and New York, 2007.

[116] M. Pickering. Laboratory education as a problem in organization. *Journal of college science teaching*, 16(3):187–189, Jan. 1987.

[117] M. J. Larson, J. D. Le Cheminant, K. Hill, K. Carbine, T. Masterson, and E. Christenson. Cognitive and typing outcomes measured simultaneously with slow treadmill walking or sitting: implications for treadmill desks. *PLoS ONE*, 10(4), 2015.

[118] M. Kendall and A. Stuart. *The Advanced Theory of Statistics. Volume 2: Inference and Relationship, Section 20.31*. Griffin, 1973. ISBN: ISBN 0-85264-215-6.

[119] S. Hidi. Interest and its contribution as a mental resource for learning. *Review of Educational Research*, 60(4):549–571, 1990.

[120] J. L. Holland. *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments. 3rd Edition*. Psychological Assessment Research Inc., U.S, 1997. ISBN: 9780911907278.

[121] A. Krapp and M. Prenzel. Research on interest in science: theories, methods, and findings. *International Journal of Science Education*, 33(1):27–50, 2011.

[122] C. Bergmann and F. Eder. *AIST-R: Allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test (UST-R), Manual.* Beltz Test Gesellschaft / Hogrefe Verlagsgruppe, Göttingen, Germany, 2005.

[123] S. J. Fux, F. Stoll, C. Bergmann, F. Eder, and B. Hell. *explorix.de – Test-Set Ausgabe Deutschland, 4th ed.* Huber, Hogrefe AG, Bern, 2012. ISBN: 9783456837437.

[124] Qualtrics subscription software platform. https://www.qualtrics.com. Accessed: 07.02.2018.

[125] J. L. Holland. A theory of vocational choice. *Journal of Counseling Psychology*, 6:35–45, 1959. DOI: 10.1037/h0040767.

[126] S. J. Fux, F. Stoll, C. Bergmann, and F. Eder. *explorix – das Werkzeug zur Berufswahl und Laufbahnplanung – Manual, 4th ed.* Huber, Hogrefe AG, Bern, 2018.

[127] Moodle is a free and open-source learning management system (LMS) written in PHP. https://moodle.org. Accessed: 07.03.2016.

[128] A. L. Comrey. Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*, 56(5):754–761, 1988. DOI: 10.1037/0022-006X.56.5.754.

[129] A. P. Field. *Discovering statistics using SPSS, 3rd edition*. SAGE Publications, Los Angeles, London, 2009. ISBN: 9781847879073.

[130] B. G. Tabachnick and L. S. Fidell. *Using multivariate statistics (6th ed.)* Pearson Education, Boston, 2013.

[131] L. Guttman. Some necessary conditions for common-factor analysis. *Psychometrika*, 19(2):149–161, 1954.

[132] H. F. Kaiser. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1):141–151, 1960. DOI: 10.1177/001316446002000116.

[133] R. Schnell, P. B. Hill, and E. Esser. *Methoden der empirischen Sozialforschung, 7. Auflage*. Oldenbourg Wissenschaftsverlag, Vienna, 2005.

[134] F. Steger, A. Nitsche, H.-G. Schweiger, and I. Belski. Teaching battery basics in laboratories: comparing learning outcomes of hands-on experiments and computer-based simulations. In *Proceedings of the 27th Annual Conference of the Australasian Association for Engineering Education (AAEE 2017)*, Coffs Harbour, Australia, 2016.

[135] C. Nebl, F. Steger, and H.-G. Schweiger. Discharge capacity of energy storages as a function of the discharge current – expanding Peukert's equation. *International journal of electrochemical science*, 12(5):4940–4957, 2017. DOI: 10.20964/2017.06.51.

[136] W. Peukert. Über die Abhängigkeit der Kapacität von der Entladestromstärke bei Bleiakkumulatoren. *Elektrotechnische Zeitschrift*, 20:287–288, 1897.

[137] F. Steger, A. Nitsche, K. Brade, I. Belski, and H.-G. Schweiger. Teaching energy storages by means of a student battery cell test system. In *Proceedings of the 45th SEFI Annual Conference*, pages 169–176, Angra do Heroísmo, Portugal. SEFI – European Society for Engineering Education, 2017.

[138] F. Steger, A. Nitsche, K. Brade, and H.-G. Schweiger. Batterietestsystem für die Lehre in der Elektromobilität. *emobilitytec*, 2:58–61, 2017.

[139] F. Steger, A. Nitsche, C. Miley, H.-G. Schweiger, and I. Belski. Laboratory learning: hands-on versus simulated experiments. In *Proceedings of the 28th Annual Conference of the Australasian Association for Engineering Education (AAEE 2017)*, Manly, Australia, 2017.

[140] F. Steger, A. Nitsche, H.-G. Schweiger, and I. Belski. Hands-on experiments vs. computer-based simulations in energy storage laboratories. In *Proceedings of the 45th SEFI Annual Conference*, Angra do Heroísmo, Portugal. SEFI – European Society for Engineering Education, 2017.

[141] F. Steger, A. Peniche dos Santos, I. Belski, and H.-G. Schweiger. Impedance spectroscopy upgrade to a student battery cell test system. In *International Workshop on Impedance Spectroscopy*, 2017.

[142] F. Steger, A. Nitsche, K. Brade, I. Belski, and H.-G. Schweiger. Energiespeicher-Praktikum an der TH Ingolstadt: Reale versus simulierte Experimente. In *Didaktik Nachrichten, Sonderausgabe Forum der Lehre*, pages 90–97, Ingolstadt, Germany, 2018.

[143] E. Machuca*, F. Steger*, J. Vogt, K. Brade, and H.-G. Schweiger. Availability of lithium ion batteries from hybrid and electric cars for second use: how to forecast for germany until 2030. *Journal of Electrical Engineering*, 6(3):129–143, June 2018. DOI: 10.17265/2328-2223/2018.03.001.

[144] W. K. Kim, F. Steger, B. S. Kotak, P. V. R. Knudsen, U. Girgsdies, and H.-G. Schweiger. Water condensation in traction battery systems. *Energies*, 12(6):1171–1181, Mar. 2019. DOI: 10.3390/en12061171.

[145] F. Steger, A. Arbesmeier, A. Nitsche, K. Brade, H.-G. Schweiger, and I. Belski. Laboratory learning: influence of the perceived laboratory mode on learning outcomes. In *Proceedings of the 47th SEFI Annual Conference*, pages 1897–1907, Budapest, Hungary. SEFI – European Society for Engineering Education, Sept. 2019. ISBN: 978-2-87352-018-2. URL: https://www.sefi.be/wp-content/uploads/2019/10/SEFI2019_Proceedings.pdf.

[146] D. Pourtoulidou, F. Steger, U. Burger, H. Göllinger, L. König, E. Obermeier, H.-G. Schweiger, and A. Frey. Mission-framed project-based learning: integrating an electric powertrain into a motor glider. In *Proceedings of the 47th SEFI Annual Conference*, pages 879–888, Budapest, Hungary. SEFI – European Society for Engineering Education, Sept. 2019.

[147] F. Steger, K. Brade, A. Nitsche, and H.-G. Schweiger. Battery education at Technische Hochschule Ingolstadt. In *Advanced Battery Power Conference/Kraftwerk Batterie*, Aachen, Germany, Apr. 2019.

[148] H.-G. Schweiger, C. Nebl, F. Steger, K. Brade, and K. Böhm. Forschung zur Sicheren Elektromobilität an der Technischen Hochschule Ingolstadt. In *University publication of the Institute of Forensic Engineering and Research, University of Žilina*, Žilina, Slovak Republic, Mar. 2019.

[149] F. Steger, J. T. Krogh, K. Brade, L. Meegahapola, and H.-G. Schweiger. Predicting available charge and energy of lithium-ion cells. *(in preparation)*, 2021.

[150] G. Goertzel. An algorithm for the evaluation of finite trigonometric series. *American Mathematical Monthly*, 65(1):34–35, 1958. DOI: 10.2307/2310304.

[151] D. Aurbach. *The Role of Surface Films on Electrodes in Li-Ion Batteries*. Springer, Boston, MA, USA, 2002.

[152] J. Y. Son, P. Narguizian, D. Beltz, and R. A. Desharnais. Comparing physical, virtual, and hybrid flipped labs for general education biology. *Online Learning*, 20(3):228–243, 2016.

[153] BMUB. Wie klimafreundlich sind Elektroautos? http://www.bmub.bund.de/fileadmin/Daten_BMU/Download_PDF/Verkehr/emob_klimabilanz_2015_bf.pdf, BMUB German Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety. Accessed: 15.01.2016.

[154] *Nationale Plattform Elektromobilität: Fortschrittsbericht 2014 - Bilanz der Marktvorbereitung*. Berlin, Germany, 2014. Gemeinsame Geschäftsstelle Elektromobilität der Bundesregierung.

[155] H. K. Norian. A battery experiment. *International Journal of Electrical Engineering Education*, 52(1):14–21, 2015.

[156] C. Thanomsilp. STEM teaching in a chemistry laboratory "how to build a simple battery in the laboratory". *Engineering and Applied Science Research*, 45(2):154–157, Apr. 2018.

[157] R. Hermann. Elektrotechnik und Elektromobilität (B. Eng.) https://www.thi.de/elektrotechnik-und-informatik/studiengaenge/elektrotechnik-und-elektromobilitaet-beng/, THI, Aug. 2018. Accessed: 20.12.2018.

[158] R. Hermann. *Modulhandbuch Elektrotechnik und Elektromobilität (B. Eng.)* TH Ingolstadt, Germany, Jan. 2018.

[159] N. Holmes. Why traditional labs fail... and what we can do about it. In J. Mintzes and E. Walter, editors, *Active Learning in College Science*. Springer, Cham, Feb. 2020. ISBN: 978-3-030-33599-1. DOI: 10.1007/978-3-030-33600-4_18.

[160] B. Biswas, S. Das, C. Koley, and P. Purkait. *Electrical and electronics measurements and instrumentation*. McGraw Hill Education (India), New Delhi and Singapore, 2013. ISBN: 125902959X.

[161] S. Tumanski. *Principles of electrical measurement*. Series in sensors. Taylor & Francis, New York, 2006. ISBN: 9780750310383.

[162] P. Purkait, B. Biswas, S. Das, and C. Koley. *Electrical and Electronics Measurements and Instrumentation*. McGraw Hill, New Delhi, 2013. ISBN: 978-1-25-902959-2.

[163] X. C. Tong. *Advanced Materials for Thermal Management of Electronic Packaging*, volume 30 of *Springer Series in Advanced Microelectronics*. Springer New York, New York, NY, 1st edition, 2011. ISBN: 1441977597.

[164] Mersen. Behaviour and operation of the fuse. *Mersen Whitepaper*, 2012.

[165] A. Etschmaier, G. Mori, W. Wegscheider, and H. Wieser. Wirkungsweise von Lotwerkstoffen beim M-Effekt. *Materialwissenschaft und Werkstofftechnik*, 33(2):90–98, 2002. ISSN: 09335137.

[166] M. Creydt. *Diffusion in galvanisch aufgebrachten Schichten und Weichloten bei Temperaturen zwischen 23 und 212°C*. PhD thesis, Eidgenössische Technische Hochschule Zürich, Zürich, 1971.

[167] A. Wright and P. G. Newbery. *Elecric Fuses, 3rd Edition*. Inst. of Electrical Engineers, 2004.

[168] J. L. Carbonero, G. Morin, and B. Cabon. Comparison between beryllium-copper and tungsten high frequency air coplanar probes. *IEEE Transactions on Microwave Theory and Techniques*, 43(12):2786–2793, 1995.

[169] R. S. Timsit. Electrical contact resistance: Properties of stationary interfaces. *IEEE Transactions on Components and Packaging*, 22(85), 1999.

[170] R. Holm. *Electric Contact, 4th ed.* Springer, Berlin, 4th edition, 1967.

[171] Y. Y. Lau and W. Tang. A higher dimensional theory of electrical contact resistance. *Journal of Applied Physics*, 105(12):124902, 2009. ISSN: 0021-8979. DOI: 10.1063/1.3148289.

[172] H.-M. Fischer. *Voltage Classes for Electric Mobility*. German Electrical and Electronic Manufacturers' Association, 2013.

[173] L. Yang, Y. Luo, R. S. K. Moorthy, D. Rahman, W. Yu, and I. Husain. Design and test of a planarized high power density 100 kW SiC traction inverter with 1kV DC-link. In *2018 IEEE Energy Conversion Congress and Exposition (ECCE)*, Sept. 2018.

[174] W. Chou, A. Kempitiya, and O. Vodyakho. Reduction of power losses of SiC MOSFET based power modules in automotive traction inverter applications. In *2018 IEEE Transportation Electrification Conference and Expo (ITEC)*, June 2018. DOI: 10.1109/ITEC.2018.8450130.

[175] V. Reber. New possibilities with 800-volt charging. *Porsche Engineering Magazine*, 1:10–15, 2016.

[176] BMW Group. Technical specifications – BMW i3 (120 Ah). http://www.press.bmwgroup.com%2Fglobal%2Farticle%2Fattachment%2FT0284828EN%2F415571&usg=AOvVaw3_LQToi09oUKWeeRexF6sU, 2018. Accessed: 07.02.2019.

[177] BMW Group. Specifications – BMW i8 Coupe. https://www.press.bmwgroup.com/global/article/attachment/T0286588EN/419603, 2018. Accessed: 07.02.2019.

[178] smart. Smart fortwo electric drive brochure. https://www.smart.com/content/dam/smart/EN/PDF/smart_electric_drive_Broschuere_en_INT.indd.pdf, 2018. Accessed: 07.02.2019.

[179] Tesla. Model X Owner's Manual. https://www.tesla.com/sites/default/files/model_x_owners_manual_north_america_en.pdf, 2019. Accessed: 08.12.2019.

[180] Tesla. Model S Owner's Manual. https://www.tesla.com/sites/default/files/model_s_owners_manual_north_america_en_us.pdf, 2019. Accessed: 08.12.2019.

[181] Tesla. Model 3 Owner's Manual. https://www.tesla.com/sites/default/files/model_3_owners_manual_north_america_en.pdf, 2019. Accessed: 08.12.2019.

[182] VW. Der neue e-up! https://cdn.volkswagen.at/media/Kwc_Basic_DownloadTag_Component/86_brochures-1120-child-download-downloadTag/default/7e10e0bc/1524477712/e-up-preisliste-04-2018-web.pdf, 2018. Accessed: 07.02.2019.

[183] T. Blackburn. NextEV: Der Nio EP 9 — Technische Daten. https://www.bmb-consult.com/upload/files/EP_9_Technisches_Datenblatt_analog_media_pack_NextEV_Layout.pdf, 2017. Accessed: 07.02.2019.

[184] M. S. Naidu and V. Kamaraju. *High voltage engineering*. Tata McGraw-Hill, New Delhi, 3rd edition, 2004. ISBN: 9780070494640.

[185] P. Changhao and T. Cong. Study on isolation monitoring of high-voltage battery system. *Applied Mechanics and Materials*, 44-47:571–575, 2011.

[186] E. Kuffel, W. S. Zaengl, and J. Kuffel. High voltage engineering: fundamentals, 1984.

[187] P. Lamp. Requirements for batteries used in electric mobility applications. In R. Korthauer, editor, *Lithium-Ion Batteries: Basics and Application*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN: 978-3-662-53069-6. DOI: 10.1007/978-3-662-53071-9.

[188] A. Wiebelt and M. G. Zeyen. Battery thermal management. In R. Korthauer, editor, *Lithium-Ion Batteries: Basics and Application*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN: 978-3-662-53069-6. DOI: 10.1007/978-3-662-53071-9.

[189] D. Anseán, V. García, M. González, J. Viera, C. Blanco, and A. J.L. DC internal resistance during charge: analysis and study on LiFePO4 batteries. In *EVS 27*, Barcelona, Spain, Nov. 2013.

[190] B. Ratnakumar, M. Smart, L. Whitcanack, and R. Ewell. The impedance characteristics of Mars exploration rover li-ion batteries. *Journal of Power Sources*, 159(2):1428–1439, 2006. ISSN: 0378-7753.

[191] J. P. Schmidt, S. Arnold, A. Loges, D. Werner, T. Wetzel, and E. Ivers-Tiffée. Measurement of the internal cell temperature via impedance: evaluation and application of a new method. *Journal of Power Sources*, 243:110–117, 2013. ISSN: 03787753. DOI: 10.1016/j.jpowsour.2013.06.013.

[192] S. H. Ahmed, X. Kang, and B. Shrestha. Effects of temperature on internal resistances of lithium-ion batteries. *Journal of Energy Resources Technology*, 137:98–111, 2015. DOI: 910.1115/1.4028698.

[193] H.-G. Schweiger, O. Obeidi, O. Komesker, A. Raschke, M. Schiemann, C. Zehner, M. Gehnen Markusand Keller, and P. Birke. Comparison of several methods for determining the internal resistance of lithium ion cells. *Sensors*, 10(6):5604–5625, 2010. DOI: 0.3390/s100605604.

[194] ISO12405-1:2011(E) Electrically propelled road vehicles – Test specification for lithium-ion traction battery packs and systems – Part 1: High-power applications, Geneva: ISO copyright office, 2011.

[195] R. Dorn, R. Schwartz, and B. Steurich. Battery management system. In R. Korthauer, editor, *Lithium-Ion Batteries: Basics and Application*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN: 978-3-662-53069-6. DOI: 10.1007/978-3-662-53071-9.

[196] U. S. Kim, C. B. Shin, and C.-S. Kim. Modeling for the scale-up of a lithium-ion polymer battery. *Journal of Power Sources*, 189(1):841–846, 2009. ISSN: 03787753. DOI: 10.1016/j.jpowsour.2008.10.019.

[197] K. Onda, T. Ohshima, M. Nakayama, K. Fukuda, and T. Araki. Thermal behavior of small lithium-ion battery during rapid charge and discharge cycles. *Journal of Power Sources*, 158(1):535–542, 2006. ISSN: 03787753. DOI: 10.1016/j.jpowsour.2005.08.049.

[198] M. S. Rad, D. L. Danilov, M. Baghalha, M. Kazemeini, and P. H. L. Notten. Thermal modeling of cylindrical lifepo batteries. *Journal of Modern Physics*, 04(07):1–7, 2013. ISSN: 2153-1196. DOI: 10.4236/jmp.2013.47A2001.

[199] N. Damay, C. Forgez, M.-P. Bichat, and G. Friedrich. Thermal modeling of large prismatic LiFePO4/graphite battery – coupled thermal and heat generation models for characterization and simulation. *Journal of Power Sources*, 283:37–45, 2015.

[200] A. A. Pesaran. Thermal Characteristics of Selected EV and HEV Batteries. http://www.nrel.gov/transportation/energystorage/pdfs/long_beach_btm.pdf. Accessed: 07.02.2016.

[201] THT. THT: determining the specific heat capacity of a battery pack. http://www.inforlab-chimie.fr/doc/document_fichier_279.pdf. Accessed: 07.02.2016.

[202] C. Forgez, D. Vinh Do, G. Friedrich, M. Morcrette, and C. Delacourt. Thermal modeling of a cylindrical LiFePO4/graphite lithium-ion battery. *Journal of Power Sources*, 195(9):2961–2968, 2010.

[203] R. Benger, H. Wenzl, and H.-P. Beck. Electrochemical and thermal modeling of lithium ion cells for use in HEV or EV application. *World Electric Vehicle Journal*, 3(2):342–351, 2009. DOI: 10.3390/wevj3020342.

[204] A. A. Pesaran. Battery thermal management in EVs and HEVs issues and solutions, 2013.

[205] T. Woehrle. Lithium-ion cell. In R. Korthauer, editor, *Lithium-Ion Batteries: Basics and Application*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN: 978-3-662-53069-6. DOI: 10.1007/978-3-662-53071-9.

[206] S. Leuthner. Lithium-ion battery overview. In R. Korthauer, editor, *Lithium-Ion Batteries: Basics and Application*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN: 978-3-662-53069-6. DOI: 10.1007/978-3-662-53071-9.

[207] J. Janek and P. Adelhelm. Next generation technologies. In R. Korthauer, editor, *Lithium-Ion Batteries: Basics and Application*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN: 978-3-662-53069-6. DOI: 10.1007/978-3-662-53071-9.

[208] C. Graf. Cathode materials for lithium-ion batteries. In R. Korthauer, editor, *Lithium-Ion Batteries: Basics and Application*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN: 978-3-662-53069-6. DOI: 10.1007/978-3-662-53071-9.

[209] M. Kulkarni, A. Vaidya, and P. Karwa. I^2t derivation for electrical safety. In *IEEE Int. Transp. Electr. Conf. (ITEC)*, 2015.

[210] S. F. Schuster, M. J. Brand, P. Berg, M. Gleissenberger, and A. Jossen. Lithium-ion cell-to-cell variation during battery electric vehicle operation. *Journal of Power Sources*, 297:241–251, Nov. 2015.

[211] F. M. González-Longatt. Circuit-based battery models: a review. In *2 do congreso iberoamericano de estudiantes de ingeneria electrica*, 2006. URL: https://pdfs.semanticscholar.org/5a79/ca86eaea1de96f7c3ea1248ea4b28f55c2f1.pdf.

[212] W. S. Bergveld H. J. and. Kruijt and P. H. L. Notten. *Battery Management Systems — Design by Modelling*. Springer, Netherlands, 2002.

[213] S. S. Zhang, K. Xu, and T. R. Jow. Electrochemical impedance study on the low temperature of li-ion batteries. *Electrochimica Acta*, 49:1057–1061, 2004.

[214] L. Guzzella and A. Sciarretta. *Vehicle Propulsion Systems – Introduction to Modeling and Optimization*. Springer Science & Business Media, Berlin Heidelberg, 2012.

[215] O. Tremblay and A.-I. Dessaint L.-A.and Dekkiche. A battery model for the dynamic simulation of hybrid electric vehicles. In *IEEE Vehicle Power and Propulsion Conference 2007*, Arlington, USA, 2007.

[216] C. Meuser. *Physikalische und chemische Charakterisierung von Lithiumionenzellen*. PhD thesis, Fakultät Mathematik und Naturwissenschaften, Dresden, Germany, 2011, page 93.

[217] V. S. Arpaci. *Conduction Heat Transfer*. Addison-Wesley Publishing Company, Massachusetts, University of Michigan, 1966.

[218] A123. AMP20M1HD-A Datasheet, 2011.

[219] Matlab: estimate parameters from measured data. https://www.mathworks.com/help/sldo/gs/estimate-parameters-from-measured-data-using-the-gui.html.

[220] J. B. Carlin and L. W. Doyle. Mann-whitney test is not just a test of medians: differences in spread can be important. *Journal of Paediatrics and Child Health*, 39(4):309–311, 2003.

[221] G. V. Glass, P. D. Peckham, and J. R. Sanders. Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Researchs*, 42(3):237–288, 1972. DOI: 10.3102/00346543042003237.

[222] N. J. Salkind. *Encyclopedia of Research Design (Vol. 2)*. Sage, Los Angeles, USA, 2010.

[223] B. Rasch, M. Friese, W. Hofmann, and E. Naumann. *Quantitative Methoden 1, Einführung in die Statistik für Psychologen und Sozialwissenschaftler, 4. überarbeitete Auflage*. Springer, Berlin/Heidelberg, 2014. ISBN: 978-3-662-43523-6.

[224] T. Lumley, P. Diehr, S. Emerson, and L. Che. The importance of the normality assumption in large public health data sets. *Annual Rev. Public Health*, 23:151–169, 2002. DOI: 10.1146/annurev.publheath.23.100901.140546.

[225] A. Rubin. *Statistics for Evidence-Based Practice and Evaluation*. Belmont: Brooks/Cole, 3rd edition, 2013, pages 144–145.

[226] J. Cohen. *Statistical power analysis for the behavioural sciences (2nd edition)*. Erlbaum, Hillsdale, NJ, 1988.

[227] A. Hart. Mann-whitney test is not just a test of medians: differences in spread can be important. *British Medical Journal*, 323:391—393, 2001.

[228] L. J. Cronbach. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3):297–334, 1951.

[229] J. P. Peter. Reliability: a review of psychometric basics and recent marketing practices. *Journal of Marketing Research*, 16:6–24, 1979.

[230] J. C. Nunnally. *Psychometric Theory (2nd ed.) – Assessment of Reliability*. McGraw-Hill, New York, 1978.

[231] K. S. Taber. The use of Cronbach's alpha – when developing and reporting research instruments in science education. *Research in Science Education*, 48(6):1273–1296, 2018.

[232] R. A. Fisher. *Statistical methods for research workers*. Oliver & Boyd, Edinburgh, 1925.

[233] R. A. Fisher. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4):507–521, 1915. DOI: 10.2307/2331838.

[234] S. C. Choi. Tests of equality of dependent correlation coefficients. *Biometrika*, 64(3):645–647, 1977. DOI: 10.1093/biomet/64.3.645.

[235] E. C. Fieller, H. O. Hartley, and E. S. Pearson. Tests for rank correlation coefficients. *Biometrika*, 44(3-4):470–481, 1957. DOI: 10.1093/biomet/44.3-4.470.

[236] J. M. Andrade and M. G. Estévez-Pérez. Statistical comparison of the slopes of two regression lines: a tutorial. *Analytica Chimica Acta*, 838(4):1–12, 2014. DOI: 10.1016/j.aca.2014.04.057.

[237] R. B. Darlington and A. F. Hayes. *Regression Analysis and Linear Models: Concepts, Applications, and Implementation*. The Guilford Press, New York, 2017. ISBN: 9781462521135.

[238] R. Korthauer, editor. *Lithium-Ion Batteries: Basics and Applications*. Springer, Berlin, Heidelberg, 2018. ISBN: 978-3-662-53069-6. DOI: 10.1007/978-3-662-53071-9.

Appendix A

Publications

The following appendix details the research publications which were published during the candidature.

A.1 List of publications

A.1.1 Journal publications

1. C. Nebl, F. Steger, and H.-G. Schweiger, “Discharge Capacity of Energy Storages as a Function of the Discharge Current – Expanding Peukert’s equation,” *Intern. Journal of Electrochemical Science,* 2017

2. F. Steger, K. Brade, A. Nitsche, and H.-G. Schweiger, “Batterietestsystem für die Lehre in der Elektromobilität,” *emobility tec,* 2017

3. F. Steger, A. Nitsche, K. Brade, I. Belski, and H.-G. Schweiger, “Energiespeicher-Praktikum an der TH Ingolstadt: Reale versus simulierte Experimente,” *Didaktik Nachrichten,* 2018

4. E. Machuca*, F. Steger*, J. Vogt, K. Brade, and H.-G. Schweiger, “Availability of lithium-ion batteries from hybrid and electric cars for second use: How to forecast for Germany until 2030,” *Journal of Electrical Engineering,* 2018

5. W.-K. Kim, F. Steger, B. S. Kotak, P. V. R. Knudsen, U. Girgsdies, H.-G. Schweiger, “Water Condensation in Traction Battery Systems”, *Energies,* 2019

6. H.-G. Schweiger, C. Nebl, F. Steger, K. Brade und K. Böhm, “Forschung zur Sicheren Elektromobilität an der Technischen Hochschule Ingolstadt,” University publication of the Institute of Forensic Engineering and Research, University of Žilina, 2019

7. F. Steger, A. Nitsche, A. Arbesmeier, K. Brade, H.-G. Schweiger, and I. Belski, “Teaching Battery Basics in Laboratories: Hands-on versus Simulated Experiments,” *IEEE Transactions on Education,* 2020

8. F. Steger, J. T. Krogh, K. Brade, L. Meegahapola, and H.-G. Schweiger, "Predicting Available Charge and Energy of Lithium-Ion Cells," (in preparation)

A.1.2 Conference proceedings

1. F. Steger, A. Nitsche, H.-G. Schweiger, and I. Belski, "Teaching Battery Basics in Laboratories: Comparing Learning Outcomes of Hands-on Experiments and Computer-based Simulations," *27th Annual Conference of the Australasian Association for Engineering Education,* 2016

2. F. Steger, A. Nitsche, K. Brade, I. Belski, and H.-G. Schweiger, "Teaching Energy Storages by means of a Student Battery Cell Test System," *45th European Society for Engineering Education Annual Conference,* 2017

3. F. Steger, A. Nitsche, C. Miley, H.-G. Schweiger, and I. Belski, "Laboratory Learning: Hands-on versus Simulated Experiments," *28th Australasian Association for Engineering Education Annual Conference,* 2017

4. F. Steger, A. Nitsche, H.-G. Schweiger, and I. Belski, "Hands-on Experiments vs. Computer-based Simulations in Energy Storage Laboratories," *45th European Society for Engineering Education Annual Conference,* 2017

5. F. Steger, A. Nitsche, K. Brade, A. Arbesmeier, H.-G. Schweiger, and I. Belski, "Laboratory learning: influence of the perceived laboratory mode on learning outcomes," *47th Annual Conference of the European Society for Engineering Education,* 2019

6. D. Pourtoulidou, F. Steger, U. Burger, H. Göllinger, L. König, E. Obermeier, H.-G. Schweiger, A. Frey, "Mission-framed Project-based Learning and Teaching: Integrating an Electric Powertrain into a Motor Glider", *47th Annual Conference of the European Society for Engineering Education,* 2019

A.1.3 Conference contributions

1. F. Steger, A. Peniche dos Santos, I. Belski, and H.-G. Schweiger, "Impedance Spectroscopy Upgrade to a Student Battery Cell Test System," *10th international Workshop on Impedance Spectroscopy,* 2017

2. F. Steger, "Laboratory Learning: Hands-on Practicals versus Simulated Experiments", *Tag der digitalen Lehre,* 2019

3. F. Steger, K. Brade, A. Nitsche and H.-G. Schweiger, "Battery Education at Technische Hochschule Ingolstadt", *Advanced Battery Power Conference / Kraftwerk Batterie,* 2019

A.2 Contributions to research publications

In the following section, for each publication, a break-down of the main contributions made by the candidate is given.
Please refer also to https://researchrepository.rmit.edu.au.

F. Steger, A. Nitsche, H.-G. Schweiger, and I. Belski, "Teaching Battery Basics in Laboratories: Comparing Learning Outcomes of Hands-on Experiments and Computer-based Simulations," ***27th Annual Conference of the Australasian Association for Engineering Education,*** **2016, [134]**

Full paper in conference proceedings. 9 pages. Double-blind peer-reviewed. Contribution to paper 66%.

This paper was the first publication regarding the educational research, describing the methodology and the results regarding the learning outcome of the 2016 run R1 (DS-A). A weak (insignificant) effect towards better learning in the hands-on mode was reported.

- Conducted literature review.
- Responsible for the overall design of the study.
- Developed simulation model.
- Conducted experiments for model parametrisation.
- Created the devices and firmware for the experiments (many thanks to Alexander Nitsche for implementing the simulations in the computer control program).
- Wrote a structured abstract. Amended structured abstract according to co-authors' suggestions. Submitted abstract for inclusion at conference.
- Wrote the draft paper. Amended paper according to co-authors' suggestions. Submitted paper for inclusion at the conference. Amended paper according to suggestions provided by the reviewers. Submitted final version for inclusion at the conference.
- Preparation of slides.

C. Nebl, F. Steger, and H.-G. Schweiger, "Discharge Capacity of Energy Storages as a Function of the Discharge Current – Expanding Peukert's equation," ***Intern. Journal of Electrochemical Science,*** **2017, [135]**

Journal paper, peer-reviewed, 18 pages. Contribution 40%.

In this article, the discharge capacity of energy storages was investigated from low to very high (out of the allowed range of the datasheet) constant discharge rates. The first author conducted the measurements for his master's thesis. From low to intermediate discharge rates, these energy storage devices showed ideal Peukert behaviour [136], while at high discharge rates the cells provide less charge than predicted by Peukert's law. The paper proposes an empirical formula to describe that behaviour.

- Conducted a part of the literature review.
- Developed a (not published) empirical formula to describe the deviation of measurement results from ideal Peukert behaviour. The formula fit well, but it was not able to solicit meaningful parameters. Also, it was not compatible with the well-known Peukert law.
- Developed the (published) formula to describe the deviation from ideal Peukert behaviour, which solicited understandable physical parameters of a cell (bend position and slope). It expands the well-known Peukert's law to allow easy usage with well-known parameters. The extension makes Peukert's law compatible with very high discharge rates and is capable of describing the discharge behaviour of lithium-ion battery cells, electrochemical double-layer capacitors, and lithium capacitors from low to high discharge rates.
- Wrote the computer programs which allowed the first author to fit his data to the proposed formula and generate graphs.
- Amended draft paper. Added thoughts regarding the explanation of the formula and conclusions.

F. Steger, A. Nitsche, K. Brade, I. Belski, and H.-G. Schweiger, "Teaching Energy Storages by means of a Student Battery Cell Test System," ***45th European Society for Engineering Education Annual Conference,*** **2017, [137]**

Conference contribution and full paper in conference proceedings. 8 pages. Double-blind peer-reviewed. Contribution to paper 66%.

This publication gives an overview of the selected content taught in the battery laboratory. It states the reasons for developing a new test system and the functions of the battery test system. All three submodules (Potentiostat/Galvanostat, Thermostat, Safety switch-off module) are explained in detail. The available functions allow student experiment procedures and instructions which encourage reflection on the new topic in an environment that stimulates hands-on learning from experimental errors.

- Market research
- Conducted the literature review.
- Responsible for the overall design of the devices and laboratories.
- Wrote a structured abstract. Amended structured abstract according to co-authors' suggestions. Submitted abstract for inclusion at conference.

- Wrote the draft paper. Amended paper according to co-authors' suggestions. Submitted paper for inclusion at the conference. Amended paper according to suggestions provided by the reviewers. Submitted final version for inclusion at the conference.

- Preparation of slides.

- Presented research findings at the conference.

F. Steger, K. Brade, A. Nitsche, and H.-G. Schweiger, "Batterietestsystem für die Lehre in der Elektromobilität," ***emobility tec,*** **2017, [138]**

Journal paper, 4 pages. Contribution 50%.

This article in a German journal reports the study programs regarding electric mobility at the Technische Hochschule Ingolstadt, and the practice-orientated teaching in the battery laboratory. Additionally, it gives an overview of all functions of the battery test system.

- Responsible for the overall design of the devices and laboratories.

- Wrote a draft paper, included co-authors' parts. Amended paper according to co-authors' suggestions. Submitted paper for publication.

F. Steger, A. Nitsche, C. Miley, H.-G. Schweiger, and I. Belski, "Laboratory Learning: Hands-on versus Simulated Experiments," ***28th Australasian Association for Engineering Education Annual Conference,*** **2017, [139]**

Conference contribution and full paper in conference proceedings. 8 pages. Double-blind peer-reviewed. Contribution to paper 66%.

This paper reports the results of runs R1 and R2, which showed a weak and significant effect on better learning results in hands-on mode (DS-A). Additionally, the results of the feedback form on the experiments (DS-D) of that runs are presented. It was the first publication which reported results of correlation research based on the questions of the pre-questionnaire (DS-B) regarding marks and the question for Germany's Dual System of Vocational Education and Training.

- Conducted literature review.

- Responsible for overall design of the study.

- Wrote structured abstract. Amended structured abstract according to co-authors' suggestions. Submitted abstract for inclusion at conference.

- Wrote draft paper. Amended paper according to co-authors' suggestions. Submitted paper for inclusion at conference. Amended paper according to suggestions provided by the reviewers. Submitted final version for inclusion at the conference.

- Preparation of slides.

- Presented research findings at the conference.

F. Steger, A. Nitsche, H.-G. Schweiger, and I. Belski, "Hands-on Experiments vs. Computer-based Simulations in Energy Storage Laboratories," ***45th European Society for Engineering Education Annual Conference,*** **2017, [140]**

Conference contribution and full paper in conference proceedings. 8 pages. Double-blind peer-reviewed. Contribution to paper 66%.

This paper reports the learning result comparison (DS-A) of the first run of 2016 and results of the online survey conducted after all laboratory experiments (DS-D) to ask for student opinions on the learning modes.

- Conducted literature review.
- Responsible for the overall design and management of the study.
- Lab-teaching.
- Conducted the study (questionnaires, tests, evaluation).
- Wrote structured abstract. Amended structured abstract according to co-authors' suggestions. Submitted abstract for inclusion at the conference.
- Wrote draft paper. Amended paper according to co-authors' suggestions. Submitted paper for inclusion at conference. Amended paper according to suggestions provided by the reviewers. Submitted final version for inclusion at the conference.
- Preparation of slides.
- Presented research findings at the conference.

F. Steger, A. Peniche dos Santos, I. Belski, and H.-G. Schweiger, "Impedance Spectroscopy Upgrade to a Student Battery Cell Test System," ***10th international Workshop on Impedance Spectroscopy,*** **2017, [141]**

Conference contribution, the abstract was peer-reviewed before invitation to the conference. Contribution 85%.

This conference contribution concerned an impedance measurement function which was added to the battery test system in 2017. The function was necessary for the laboratory version for guest lectures at other universities to replace a tool (Hioki Hitester), to avoid the transport of these devices to other cities. It is also helpful when using the self-developed devices for scientific purposes. The paper presented the extensions which were necessary to be able to measure impedance at each frequency between 60 Hz and 10 kHz with the battery test system. The research additionally reports the reached measurement quality regarding small lithium-ion battery cells.

- Conducted literature review.
- Addons and changes to the firmware of the devices to allow for impedance spectroscopy > 60 Hz.

- Conducted the experiments for validation of the devices accuracy.
- Wrote short abstract. Submitted abstract for inclusion at conference.
- Wrote extended abstract (2 pages) for the conference. Amended paper according to co-authors' suggestions.
- Preparation of slides.
- Presented research findings at the conference in Chemnitz.

F. Steger, A. Nitsche, K. Brade, I. Belski, and H.-G. Schweiger, "Energiespeicher-Praktikum an der TH Ingolstadt: Reale versus simulierte Experimente," ***Didaktik Nachrichten,*** **2018, [142]**

Journal paper based on a conference workshop contribution, peer-reviewed, 7 pages. Contribution 80%.

The German paper discusses the results of the first phase with the main focus of the target group of the journal and conference, German university teachers.

- Submitted draft paper based on the main research of this thesis.
- Amended paper according to suggestions provided by the reviewers of the journal.
- Submitted final paper.
- Held workshop at the associated conference.

E. Machuca*, F. Steger*, J. Vogt, K. Brade, and H.-G. Schweiger, "Availability of lithium-ion batteries from hybrid and electric cars for second use: How to forecast for Germany until 2030," ***Journal of Electrical Engineering,*** **2018, [143]**

Journal paper, peer-reviewed, 14 pages. * the two first authors contributed equivalent to the publication. Contribution 40%.

This research concerned a method of forecasting the available lithium-ion car battery systems for second use and recycling. The study included batteries from accidents and ageing of electric cars in Germany.

- Responsible for the part "car ageing" in terms of literature research and data acquisition. (Enrique Machuca worked on the assumptions regarding the first registrations of electric driven cars, Johanna Vogt did the estimations for accident rates.)
- Based on the literature, developed and parameterised a model for car ageing and life span.
- Developed the overall mathematical model to forecast market behaviour based on the combined data of all co-authors.

- Developed the mathematical system and equations to derive output data.
- Took over responsibility for the publication after E. Machuca left THI.
- Amended draft paper. Submitted draft paper.
- Amended paper according to suggestions provided by the reviewers of the journal.
- Submitted final paper.

W.-K. Kim, F. Steger, B. S. Kotak, P. V. R. Knudsen, U. Girgsdies, H.-G. Schweiger, "Water Condensation in Traction Battery Systems", *Energies,* 2019, [144]

Journal paper, peer-reviewed, 17 pages. Contribution 30%.

Lithium-ion traction battery systems of hybrid and electric vehicles, like all other components and systems of a vehicle, must have a high level of durability and reliability. Battery systems get heated while in the application. To ensure the desired life span and performance, most systems are equipped with a cooling system.

The changing environmental condition in daily use may cause water condensation in the housing of the battery system. In this study, three system designs were investigated to compare different solutions to deal with pressure differences and condensation:

1. a sealed battery system,
2. an open system and
3. a battery system equipped with a pressure compensation element (PCE).

These three designs were tested under two conditions: (a) in normal operation and (b) in a maximum humidity scenario.

The amount of condensation in the housing was determined through a change in relative humidity of air inside the housing. Through PCE and available spacing of the housing, moisture entered into the housing during the cooling process.

While applying the test scenarios, the gradient-based drift of the moisture into the housing contributed maximum towards condensation. Condensation occurred on the internal surface for all three design variants.

- Support while conducting a small part of the experiments.
- Taking over responsibility for the publication after Woong-Ki Kim left THI.
- Review of measurement data, validation (as far as it was possible).
- Creation of a rough model to explain the effects and identification of the gradient-driven water transport as the main reason (based on the data).
- A lot of rework on the draft paper.

- Added schemes to explain the test bench.
- Language corrections.
- Enhancing the team with new members to be able to finalise the paper. Organising the rest of the work.
- Submitted draft paper.
- Amended paper according to suggestions provided by the reviewers of the journal.
- Submitted final paper.

F. Steger, "Laboratory Learning: Hands-on Practicals versus Simulated Experiments", *Tag der digitalen Lehre,* 2019

Conference contribution, Oral presentation, approximately 30 minutes. Contribution to the introductory part 95%.

Introductory part of a workshop held by Hans-Georg Schweiger ("Einsatz digitaler Labor-Praktika in der Weiterbildung", "Digital lab teaching for executive education"). THI, 22.5.2019.

- Extraction of relevant information of the main thesis research for executive education.
- Preparation of slides.
- Oral presentation.

F. Steger, A. Nitsche, K. Brade, A. Arbesmeier, H.-G. Schweiger, and I. Belski, "Laboratory learning: influence of the perceived laboratory mode on learning outcomes," *47th Annual Conference of the European Society for Engineering Education,* 2019, [145]

Conference contribution and full paper in conference proceedings. 10 pages. Double-blind peer-reviewed. Contribution to paper 66%.

This paper presents the first results of the second phase of the study (R5 – R8). In order to ensure that the mode of a laboratory experiment will not be an influencing factor, all participants used hands-on equipment for laboratory experiments. Subjects from one group used hands-on equipment. Subject from the other group only thought that they were conducting hands-on experiments. Their equipment was 'modified' to display the results of simulated experiments. The outcomes of the second phase of the study supported the "perception" hypothesis. No statistically significant differences in student learning of the subjects from the second phase of the experiment were discovered.

- Conducted literature review.

- Responsible for overall design and management of the study.
- Modification of the used devices and firmware to allow the "hidden simulations" (many thanks to Alexander Nitsche for also adding the necessary functions in the computer control program).
- Lab-teaching.
- Conducted the study (questionnaires, tests, evaluation).
- Wrote a structured abstract. Amended structured abstract according to co-authors' suggestions. Submitted abstract for inclusion at conference.
- Wrote draft paper. Amended paper according to co-authors' suggestions. Submitted paper for inclusion at conference. Amended paper according to suggestions provided by the reviewers. Submitted final version for inclusion at the conference.
- Preparation of slides.
- Presented research findings at the conference (Session: Integrated learning environments for the digital native learners, Monday, 16/09/2019, 10:05 - 11:45).

D. Pourtoulidou, F. Steger, U. Burger, H. Göllinger, L. König, E. Obermeier, H.-G. Schweiger, A. Frey, "Mission-framed Project-based Learning and Teaching: Integrating an Electric Powertrain into a Motor Glider", *47th Annual Conference of the European Society for Engineering Education,* 2019, [146]

Conference contribution and full paper in conference proceedings. 9 pages. Double-blind peer-reviewed. Contribution to paper 35%.

Technische Hochschule Ingolstadt (THI) started implementing *mission-framed* project-based learning (PBL) in 2015. Instead of receiving a task which can be completely finished in a typical one-semester period, students work on a small part of a complex and interdisciplinary system. More genuine to professional life, they are required to meet a strict time-frame with expected results. Such mission-framed projects create an industrial working environment for project modules to deepen practice-oriented technical knowledge.

As a case study, THI initiated the cross-faculty project "E-Falke (e-falcon)". The goal of this project was the substitution of the internal combustion engine of a glider aeroplane with an electric motor. Over 130 students were involved in this project, cooperating in small groups for the fulfilment of pre-defined goals, such as the feasibility evaluation of the project, ground bench testing, integration of the system into the frame and the official flight certification from the German Federal Aviation Office.

Like in conventional projects, students learned the handling of technical tasks in teams and developed their personal competences. The extensive technical complexity of the aeroplane required practical knowledge and thorough documentation for ensuing groups. This challenged the students and highlighted the need for intensive supervision and technical support during the semester projects.

This paper outlines the diverse goals of all stakeholders (teachers, university, external partners) and the advantages and disadvantages experienced during the implementation of this mission-framed teaching method. The lessons learnt through the years with planned optimisations for the ongoing project are presented to allow other educators to improve PBL.

- Supervising most of the student projects regarding E-Falke.
- Initial idea to publish the gained experiences on “mission-framed” PBL.
- Wrote structured abstract. Amended structured abstract according to co-authors' suggestions. Submitted structured abstract for inclusion at conference.
- The abstract was accepted for full paper submission at the 46th European Society for Engineering Education Annual Conference in Copenhagen (Sep. 2018), but due to shortage of time in the whole team, it was decided to present the full paper in the following year. Finally, Despoina Pourtoulidou took over responsibility for next year's publication. Despoina Pourtoulidou interviewed all stake-holders of the project, including the relevant lecturers (F. Steger, U. Burger, H. Göllinger, L. König, E. Obermeier, H.-G. Schweiger, A. Frey) and wrote the draft paper.
- Rework of the draft paper.
- Amendments on the final paper.
- Rework of the slides.
- Presented research findings at the conference (Session: New complexity quest in engineering sciences, Monday, 16/09/2019, 16:15 - 17:15).
- Participated in the subsequent panel discussion on project based learning and teaching.

F. Steger, K. Brade, A. Nitsche and H.-G. Schweiger, “Battery Education at Technische Hochschule Ingolstadt”, *Advanced Battery Power Conference / Kraftwerk Batterie,* 2019, [147]

Oral conference contribution. The extended abstract was peer-reviewed. Contribution 90%.

This presentation reported the contents taught in the battery laboratories and experiences with the laboratory gained at THI. It was optimised for battery experts, which is the target group of that conference. It presented the contents of the battery laboratory to allow other teachers to start with similar laboratories. Also, the laboratory mode impact on the different types of students was given.

- Extraction of relevant information of the main thesis research for battery experts.
- Preparation of slides.
- Oral presentation.

H.-G. Schweiger, C. Nebl, F. Steger, K. Brade und K. Böhm, "Forschung zur Sicheren Elektromobilität an der Technischen Hochschule Ingolstadt," University publication of the Institute of Forensic Engineering and Research, University of Žilina, 2019, [148]

University publication. 13 pages. Contribution 15%.

The publication sums up the research activities regarding safe electro-mobility at the UAS Ingolstadt.

- Author of the part regarding the laboratory for electro-mobility, including the available devices and research methods.
- Amendments to the draft versions of other authors.

F. Steger, A. Nitsche, A. Arbesmeier, K. Brade, H.-G. Schweiger, and I. Belski, "Teaching Battery Basics in Laboratories: Hands-on versus Simulated Experiments," *IEEE Transactions on Education,* 2020, [48]

Journal article. 11 pages. Peer-reviewed. Contribution to paper 66%.

This paper presents all results of the first phase of the study, including results of summer schools and research in other faculties and universities (up to run R5). Statistically significant differences in student learning in the first phase of the experiment were discovered. Hands-on experiments performed better than computer simulations at teaching battery basics.

- Conducted literature review.
- Responsible for overall design and management of the study.
- Creation of the used devices and firmware to allow the experiments (many thanks to Alexander Nitsche for adding all functions in the computer control program).
- Lab-teaching.
- Conduction of the study (questionnaires, tests, evaluation).
- Wrote draft paper. Amended paper according to co-authors' suggestions. Submitted paper to the IEEE journal. Amended paper according to suggestions provided by the reviewers twice.

F. Steger, J. T. Krogh, K. Brade, L. Meegahapola, and H.-G. Schweiger, "Predicting Available Charge and Energy of Lithium-Ion Cells," (in preparation), [149]

Journal article. ≈12 pages. Contribution to paper 66%.

Design and operation of performant and safe electric vehicles depend on precise knowledge of the state of their electrochemical energy storage system. Optimisation via simulation and proper function of battery management systems often rely on equations and discrete-time battery models that correctly emulate the battery characteristics. Among available models, electric-circuit models have shown to be especially useful to describe the electrical characteristics of batteries. To overcome existing drawbacks, such as the need for discrete-time simulations or the usage of look-up tables, a set of equations has been developed that solely rely on the open-circuit voltage and the internal resistance of a battery in relation to its state of charge. These values can be obtained from typical cell data sheets or can be easily extracted through standard measurements. The proposed equations allow for direct analytical determination of available discharge capacity and the available energy content depending on the discharge current, as well as the Peukert exponent. The fidelity of the proposed system was validated experimentally using a 18650 2.5 Ah NMC lithium-ion cell, and results are in close agreement with the datasheet.

- Initial concept for the methodology and research.
- Responsible for overall design and management of the study.
- Development of the mathematical system and equations.
- Creation of the battery test plan for parametrisation of the equations.
- Parametrisation of battery.
- Creation of the battery test plan for validation of the calculated results.
- Instructed the assistants who performed the above-mentioned experiments.
- Evaluation of the simulation and experimental results.
- Identified the main reasons for the Peukert effect, which are covered by the proposed equations.
- Wrote a draft paper. Amended paper according to co-authors' suggestions.
- Repeated the experiments with other battery cells to collect a broader database
- Work on this research is ongoing

Appendix B

Aspects which were held identical in the compared modes (Summary)

This appendix lists the identified aspects which were intentionally covered by the methodology in the case study to be identical in the compared modes.

- Content of the laboratories
  - Identical learning objectives in all modes
  - Experimental approach, procedure, and experimental results were the same in both modes
  - The same experiment was conducted (hands-on/simulated).
  - Trough the usage of the safety-switch-off-module (see subsection F.2.4) lessons were created without further safety limitations, students were able to work with trial and error both in the simulations and with hands-on experiments.
  - The experiments were developed for both modes simultaneously, rather than finishing one mode and "transferring" the procedure to the other.
- Teaching materials
  - Mode-specific instructions were completely avoided (in contrast to e.g. [15–18]).
  - A single set of instructions was used, which did not include any mode-specific hints, which might disturb the comparison. Thus, the level of guidance and investigation was the same.
- Accompanying lectures were always held in front of all student of a run at the same time to avoid different knowledge gain from the theoretical background.
- The same graphical user interface was used to control simulations and hands-on experiments.

- Modes of learning
  - Interaction with equipment was identical in the compared modes (sensor experiments).
  - No domain change, the comparison hands-on vs. simulation was not mixed with in-person (local at university) vs. remote teaching, so effects from distance learning were not mixed with mode-effects (in contrast to e.g. [15, 16]).
  - The study did not mix synchronous/asynchronous learning.
- Controlled group interaction
  - To ensure the same cooperative learning conditions, working team partners did not change for the whole duration of laboratory work [67, 69, 72]. Also, the group of students was constant during all runs.
  - In many runs, both groups conducted the laboratory simultaneously in the same room.
  - Compared groups were formed to be "similarly practical" (regarding APE) to ensure similar group environments in the German runs.
- Supervision
  - In all runs, the same instructor supervised the labs. Laboratory assistants did not change within single study runs and were always responsible for both compared groups.
  - As the intensity of guidance can influence student learning [66], instructions were written to foster stand-alone work, which reduced further influences of supervision to a minimal level.
  - The amount, scope and type of supervision while experiments were held similar in both modes (in contrast to e.g. [15], and [16]).
- Simulation model
  - The simulation model was a black-box model to avoid additional insights caused by a visible model which may influence the research outcomes.
  - Simulation time in the context was equivalent to hand-on experiments. As the simulation ran continuously from beginning of the lesson through to the end, simulated cell behaviour also covered aspects regarding the global process of experimentation (e.g. the cell cooling down after a high current profile, even when no experiment was running, equivalent to the real cell in hands-on experiments).
  - In the proposed research, possible "enhancements" specific to one mode, that did not allow for transfer to the other, were avoided (This was not the norm in previous studies: [15], for example, provided "colour coded

    stress and strain values" and additional options like changing materials in simulation, which were not available in the real experiment).
  - The simulation model simulated exactly all observed effects of the real battery cell, including behaviour of the hands-on devices which are used in hands-on mode (e.g., measurement noise, quantisation, etc.). Quality of the simulation was checked in the second research phase.

- Devices and the simulation were accessed through the same user interface.

- Testing the learning

  - Identical tests were employed for both groups of a run.
  - Tests were graded by the same person.
  - Tests were mixed before corrections to avoid a "drift" during grading. Later the results were assigned to the learning mode.
  - The study was not asking for understanding directly after the sessions, in order to include information on knowledge retention. In order to equalise the influence of time on the ability to remember, equal periods were targeted between experimentation and the associated tests for both groups.
  - The environment while writing tests was equivalent. All tests were filled in a sitting position in a computer pool or standing in the lab. That mode did not change during a research run; see Table 4.4.
  - Tests targeted for knowledge and understanding; see section 2.8.
  - All tests were held in written form.

Appendix C

Research ethics – letter of approval

This research and usage of student time to conduct parts of the research was approved by the Faculty of Electrical Engineering and Computer Science at Technische Hochschule Ingolstadt (Deans Letter shown in Figure C.2, Prof. Dr. Wolf-Dieter Tiedemann) and the College Human Ethics Advisory Network of RMIT, Melbourne (Approval shown in Figure C.1). The HREC/CHEAN Approval Number for this research was "ASEHAPP 18-16".

Access to the university data-base to collect objective data was approved by the Vice President for Teaching, Students and Alumni at THI (Vice Presidents Letter shown in Figure C.3, Prof. Dr. Michaela Regler) and the College Human Ethics Advisory Network of RMIT, Melbourne by an amendment to the above-mentioned approval (shown in Figure C.4).

A more detailed information on the topic research ethics is presented on page 79.

RMIT UNIVERSITY

RMIT University

Science Engineering and Health

College Human Ethics Advisory Network (CHEAN)

Plenty Road
Bundoora VIC 3083

PO Box 71
Bundoora VIC 3083
Australia

Tel. +61 3 9925 [redacted]
Fax +61 3 9925[redacted]
• www.rmit.edu.au

30th March 2016

Professor Iouri Belski
[redacted]
School of Engineering
RMIT University

Dear Professor Belski

ASEHAPP 18-16 BELSKI-STEGER Battery Test System for Practical Student Education regarding Energy Storage Systems

Thank you for submitting your amended application for review.

I am pleased to inform you that the CHEAN has approved your application for a period of **2 Years** from the date of this letter to **30th March 2018** and your research may now proceed.

The CHEAN would like to remind you that:

All data should be stored on University Network systems. These systems provide high levels of manageable security and data integrity, can provide secure remote access, are backed up on a regular basis and can provide Disaster Recover processes should a large scale incident occur. The use of portable devices such as CDs and memory sticks is valid for archiving; data transport where necessary and for some works in progress.
The authoritative copy of all current data should reside on appropriate network systems; and the Principal Investigator is responsible for the retention and storage of the original data pertaining to the project for a minimum period of five years.

Please Note: Annual reports are due on the anniversary of the commencement date for all research projects that have been approved by the CHEAN. Ongoing approval is conditional upon the submission of annual reports failure to provide an annual report may result in Ethics approval being withdrawn.

Final reports are due within six months of the project expiring or as soon as possible after your research project has concluded.

The annual/final reports forms can be found at:
www.rmit.edu.au/staff/research/human-research-ethics

Yours faithfully,

Dr Linda Jones
Chair, Science Engineering & Health
College Human Ethics Advisory Network

Cc CHEAN Member: Margaret Lech School of Engineering RMIT University
Student Investigator/s: Fabian Steger [redacted] School of Engineering RMIT University
Other Investigator/s: Dr. Lasantha Meegahapola School of Engineering RMIT University

Figure C.1: Ethics Approval Letter

Technische Hochschule Ingolstadt
Fakultät für Elektrotechnik und Informatik

Technische Hochschule Ingolstadt • Postfach 210454 • 85019 Ingolstadt

To whom it may concern

Besucheranschrift
Esplanade 10
85049 Ingolstadt
www.thi.de

Prof. Dr. Wolf-Dieter Tiedemann
Dekan

Tel +49 841 9348-
Fax +49 841 9348-

Unser Zeichen:

Ingolstadt, 12.02.2016

Battery Test System for Practical Student Education regarding Energy Storage Systems

The aim of Mr. Steger's experiments is to prove (or disprove) that his newly developed hands-on training on batteries leads to a better student comprehension of battery concepts than the MatLab-based training approach we used previously.

The Faculty of Electrical Engineering and Computer Science grants full support to Mr. Steger's research, including

- the permission to recruit students (corresponding to item 2.5.1 of the ethics form),
- the permission to use class time (according to item 2.5.2.2 of the ethics form), and
- the permission to use on-campus locations such as the Energy Storage Systems Laboratory C301 and several computer laboratories for practical education and accompanying surveys (in accordance with item 6.1 of the ethics form).

Prof. Dr. Wolf-Dieter Tiedemann
Dean

Figure C.2: Support Letter Dean

Technische Hochschule Ingolstadt
Die Vizepräsidentin

Technische Hochschule Ingolstadt • Postfach 210454 • 85019 Ingolstadt

RMIT University
Human Research Extension of HREC/CHEAN
Project

Besucheranschrift
Esplanade 10
85049 Ingolstadt
www.thi.de

Prof. Dr. jur. Michaela Regler
Vizepräsidentin

Tel +49 (0) 841 / 9348-
Fax +49 (0) 841 / 9348-

Ingolstadt, 15.05.2019

Recommandation for Fabian Steger

Dear Mary Duffy,

The aim of Mr. Steger's experiments is to compare laboratory work conducted in two different learning modes: first as a practical hands-on exercise and secondly using computer-based simulations. The outcomes significantly differ between participants that completed a vocational education and training (VET) before matriculating at university and the students that matriculated directly after school.

As in our university the share of VET-participants is high, the reasons for this differing behavior are of high interest.

The university grants full support to Mr. Steger's research, including the analysis of anonymized data mentioned in the amendment form.

Best regards

Prof. Dr. jur. Michaela Regler
Vicepresident für Teaching, Students and Alumni

Esplanade 10, D-85049 Ingolstadt
Tel +49 (0) 841 / 9348-
Fax +49 (0) 841 / 9348-

www.thi.de

Figure C.3: Support Letter Vice President

RMIT UNIVERSITY

College Human Ethics Advisory Network (CHEAN)
College of Science, Engineering & Health (SEH)
NHMRC Code: EC00237

Notice of Approval

Date:	**21 May 2019**
Project number:	**ASEHAPP 18-16**
Project title:	**'Battery Test System for Practical Student Education regarding Enegy Storage Systems**
Risk classification:	**Low Risk**
Investigator(s):	**Prof Iouri Belski, Mr Fabian Steger, Dr Lasantha Meegahapola**
Approval period:	From: **30/03/2016** To: **31/05/2019**

I am pleased to advise that your amendment/extension request has been granted ethics approval by the Science, Engineering and Health College Human Ethics Advisory Network (SEH CHEAN), as a sub-committee of the RMIT Human Research Ethics Committee (HREC). Ethics approval is extended until **30/07/2019**.

Terms of approval:

1. **Responsibilities of investigator**
   It is the responsibility of the above investigator/s to ensure that all other investigators and staff on a project are aware of the terms of approval and to ensure that the project is conducted as approved by the CHEAN. Approval is only valid whilst the investigator/s holds a position at RMIT University.
2. **Amendments**
   Approval must be sought from the CHEAN to amend any aspect of a project including approved documents. To apply for an amendment please use the 'Request for Amendment Form' that is available on the RMIT website. Amendments must not be implemented without first gaining approval from CHEAN.
3. **Adverse events**
   You should notify HREC immediately of any serious or unexpected adverse effects on participants or unforeseen events affecting the ethical acceptability of the project.
4. **Participant Information Sheet and Consent Form (PISCF)**
   The PISCF and any other material used to recruit and inform participants of the project must include the RMIT university logo. The PISCF must contain a complaints clause.
5. **Annual reports**
   Continued approval of this project is dependent on the submission of an annual report. This form can be located online on the human research ethics web page on the RMIT website.
6. **Final report**
   A final report must be provided at the conclusion of the project. CHEAN must be notified if the project is discontinued before the expected date of completion.
7. **Monitoring**
   Projects may be subject to an audit or any other form of monitoring by HREC at any time.
8. **Retention and storage of data**
   The investigator is responsible for the storage and retention of original data pertaining to a project for a minimum period of five years.

Figure C.4: Amendment for accessing objective data and extension

Appendix D

Amount of Practical Experience, individual items (DS-B)

The following appendix presents the analysis of the responses to individual APE test set items for determining the Amount of Practical Experience. The main analysis based on the identified APE-dimensions can be found in section 5.3.

D.1 Item difficulties

To aid other researchers wishing to use the same or similar items in their research, the difficulties of all items are reported in Table D.1. The analysis was performed for German (VET and non-VET) and international participants separately.

Differences of VET and non-VET

According to the results of a non-parametric Mann-Whitney U-Test, VET-participants had more often experienced the employed items (see Table D.2). More than 50% of the items show significant trends towards higher experience among former VET-participants, while not a single item exhibited a trend towards higher experience among non-VET students.

Table D.1: APE: Difficulties

		German UAS				International		all	
		VET		non-VET					
Students		47		35		55		181	
Item		Mean	Mdn	Mean	Mdn	Mean	Mdn	Mean	Mdn
APE-1	Realised a function using a self-made circuit diagram.	.16	.5	-.70	-1.5	.22	.5	-.10	-.50
APE-2	Transferred a circuit diagram to a PCB.	.18	.5	-.01	.5	-.10	-.5	-.02	-.50
APE-3	Soldered electronic parts with a soldering iron.	1.48	1.5	1.47	1.5	.92	1.5	1.29	1.50
APE-4	Configured and assembled a desktop-PC.	.20	.5	-.16	-.5	.32	.5	.12	.50
APE-5	Made a thread on a bore hole.	1.46	1.5	.59	1.5	-.03	.5	.73	1.50
APE-6	Programmed a micro controller.	-.44	-.5	-.33	-.5	-.07	-.5	-.17	-.50
APE-7	Assembled a model kit.	.97	1.5	.84	1.5	.23	.5	.73	1.50
APE-8	Sawed off a pipe.	1.44	1.5	.93	1.5	.54	.5	1.04	1.50
APE-9	Searched an error in the electric system of a car, found it, and fixed it.	.82	1.5	-.21	-.5	-.52	-1.5	.00	-.50
APE-10	Changed a car's tires.	1.46	1.5	1.16	1.5	.59	1.5	1.11	1.50
APE-11	Produced a metal part from a technical drawing.	1.27	1.5	-.21	-.5	-.22	-.5	.42	1.50
APE-12	Changed a motor's oil.	.78	1.5	.07	.5	.20	.5	.35	1.50
APE-13	Designed an analogue low pass-filter of 2nd order for a given requirement.	-.16	-.5	-.33	-.5	-.71	-1.5	-.48	-.50
APE-14	Simulate the frequency response of a simple RC element in SPICE.	.59	.5	.59	.5	-.45	-.5	.26	.50
APE-15	Read and explain a circuit diagram containing an operational amplifier.	.37	.5	.10	.5	.20	.5	.22	.50
APE-16	Explain to others what a C compiler does in principle.	.20	.5	.16	.5	-.31	-.5	.06	.50
APE-17	Charge a lead-acid battery of a common car with a laboratory power supply.	.90	1.5	.04	.5	-.01	.5	.40	.50
APE		.69	.74	.23	.32	.05	.03	.35	.38
APE-E		.29	.39	.09	.17	.00	-.06	.13	.17
APE-ED		.50	.50	.40	.50	-.07	-.10	.27	.30
APE-EC		.07	.00	-.24	-.25	.06	.00	-.02	.00
APE-M		1.14	1.38	.40	.50	.10	.13	.60	.75
APE-MC		1.00	1.50	.26	.50	.07	.00	.46	.75
APE-MD		1.39	1.50	.43	.50	.11	.17	.73	1.17

Table D.2: APE: individual APE items, non-parametric comparison of VET and non-VET answers. The table is sorted by effect size.

		non-VET N = 39		VET N = 48					
Item	Content	M Rank	Sum	M Rank	Sum	U	Z		p
APE-11	Produced a metal part from a technical drawing.	31.5	1229	54.1	2599	449	-4.83	**	.000
APE-5	Made a thread on a bore hole.	35.5	1385	50.9	2443	605	-4.09	**	.000
APE-9	Searched an error in the electric system of a car, found it, and fixed it.	33.2	1295	52.8	2534	515	-3.90	**	.000
APE-17	Charge a lead-acid battery of a common car with a laboratory power supply.	34.0	1325	52.2	2504	545	-3.59	**	.000
APE-8	Sawed off a pipe.	38.2	1489	48.7	2340	709	-3.13	**	.002
APE-1	Realised a function using a self-made circuit diagram.	35.1	1368	51.3	2461	588	-3.11	**	.002
APE-12	Changed a motor's oil.	36.8	1434	49.9	2395	654	-2.77	**	.006
APE-10	Changed a car's tires.	40.8	1592	46.6	2237	812	-2.25	*	.024
APE-4	Configured and assembled a desktop-PC.	38.7	1508	48.3	2321	728	-1.84	†	.066
APE-15	Read and explain a circuit diagram containing an operational amplifier.	39.7	1548	47.5	2281	768	-1.52		.127
APE-13	Designed an analogue low pass-filter of 2nd order for a given requirement.	41.3	1610	46.2	2219	830	-.95		.343
APE-3	Soldered electronic parts with a soldering iron.	43.3	1688	44.6	2141	908	-.77		.442
APE-2	Transferred a circuit diagram to a PCB.	41.8	1631	45.8	2197	851	-.76		.446
APE-16	Explain to others what a C compiler does in principle.	42.2	1644	45.5	2184	864	-.64		.523
APE-14	Simulate the frequency response of a simple RC element in SPICE.	43.1	1682	44.7	2147	902	-.31		.756
APE-7	Assembled a model kit.	43.3	1689	44.6	2139	909	-.27		.786
APE-6	Programmed a micro controller.	44.5	1737	43.6	2092	916	.18		.855

Note. † = p < .10. * = p < .05. ** = p < .01.
U = Mann-Whitney U, Z pos. = more experience of non-VET.

D.2 Test performances

A Spearman rank-order correlation coefficient was computed to assess the relationship between the individual APE items and the students' overall test performance. The results are presented in Table D.3 in the right column. The significant correlations (APE-14, APE-9, APE-7, APE-8 and APE-6) were distributed over all APE-dimensions. All of them were positive, showing that better average test results of students were connected to a higher Amount of Practical Experience.

D.3 Superior learning mode

A Spearman rank-order correlation coefficient was computed to assess the relationship between the individual APE items and the students' superior learning mode. The results are presented in Table D.3 in the left column. Three (APE-14, APE-16, APE-3) of the four significant correlations were part of the APE-ED dimension, all of them positive, explaining the overall correlation between the APE-ED dimension and better learning results in hands-on experiments.

As the results of German-VET, German-non-VET and internationals were mixed for this analysis, the results have to be seen critically.

Superior learning mode, grouped for VET and non-VET

A Spearman rank-order correlation coefficient was computed to assess the relationship between the individual APE items and the students' superior learning mode, grouped for VET and non-VET students. The results are presented in Table D.4. All statistically significant correlations in individual items suggested that students who had *not* previously had the experience performed better in tests after hands-on experiments compared to simulated experiments (SLM). Nearly all items showed this trend, both for VET and non-VET students.

Table D.3: APE: Spearman rank-order correlation between superior learning mode, students' overall test performance and all 17 APE items. All runs, first study phase.

[a] Item		superior learning mode (TP hands-on - simulated) ρ		p	N	TP all modes ρ		p	N
APE-ED		.22	†	.061	71				
APE-14	Simulate the frequency response of a simple RC element in SPICE.	.22	†	.060	71	.18	*	.020	161
APE-16	Explain to others what a C compiler does in principle.	.24	*	.043	71	.01		.921	161
APE-15	Read and explain a circuit diagram containing an operational amplifier.	.08		.486	71	.09		.246	161
APE-13	Designed an analogue low pass-filter of 2nd order for a given requirement.	.10		.403	70	-.03		.733	160
APE-3	Soldered electronic parts with a soldering iron.	.21	†	.082	71	-.04		.584	161
APE-MC		.05		.657	71				
APE-9	Searched an error in the electric system of a car, found it, and fixed it.	.07		.589	71	.13	†	.095	161
APE-17	Charge a lead-acid battery of a common car with a laboratory power supply.	.18		.142	71	.08		.326	161
APE-12	Changed a motor's oil.	-.10		.392	71	.08		.291	161
APE-10	Changed a car's tires.	.16		.175	71	.05		.552	161
APE-7	Assembled a model kit.	-.03		.790	70	.15	†	.057	160
APE-MD		.19		.122	71				
APE-11	Produced a metal part from a technical drawing.	.20	†	.099	71	-.01		.904	161
APE-5	Made a thread on a bore hole.	.19		.110	69	.05		.549	159
APE-8	Sawed off a pipe.	-.06		.627	71	.16	*	.038	161
APE-EC		.08		.493	71				
APE-1	Realised a function using a self-made circuit diagram.	.05		.705	71	-.03		.744	161
APE-2	Transferred a circuit diagram to a PCB.	.08		.503	71	.05		.570	161
APE-4	Configured and assembled a desktop-PC.	-.10		.429	71	-.03		.667	161
APE-6	Programmed a micro controller.	.05		.667	71	.15	†	.058	160

Note. † = p < .10. * = p < .05. ** = p < .01. Positive "superior learning mode"-values describe advantages of hands-on.

[a] As the results of German-VET, German-non-VET and internationals were mixed for this analysis, these results have to be seen critically.

Table D.4: APE: Spearman rank-order correlation between the superior learning mode and all the APE items and dimensions, grouped for VET. R2, first study phase.

Item		VET N = 16 ρ		VET N = 16 p	non-VET N = 9 ρ		non-VET N = 9 p
APE		-.42		.103	-.59	†	.097
APE-E		-.23		.392	-.31		.417
APE-M		-.66	**	.005	-.67	*	.049
APE-ED		-.18		.501	.08		.847
APE-14	Simulate the frequency response of a simple RC element in SPICE.	.19		.534	.37		.328
APE-16	Explain to others what a C compiler does in principle.	.05		.852	-.23		.560
APE-15	Read and explain a circuit diagram containing an operational amplifier.	-.23		.387	.00		.999
APE-13	Designed an analogue low pass-filter of 2nd order for a given requirement.	-.39		.136	.08		.839
APE-3	Soldered electronic parts with a soldering iron.		a			a	
APE-MC		-.62	*	.011	-.80	*	.010
APE-9	Searched an error in the electric system of a car, found it, and fixed it.	-.50	*	.049	-.89	**	.001
APE-17	Charge a lead-acid battery of a common car with a laboratory power supply.	-.33		.206	-.16		.689
APE-12	Changed a motor’s oil.	-.53	*	.034	-.82	**	.007
APE-10	Changed a car’s tires.	-.31		.246	-.14		.725
APE-7	Assembled a model kit.	-.20		.469	-.54		.135
APE-MD		-.34		.205	-.17		.660
APE-11	Produced a metal part from a technical drawing.	-.34		.205	.17		.670
APE-5	Made a thread on a bore hole.	-.14		.605	-.48		.186
APE-8	Sawed off a pipe.		a		-.14		.725
APE-EC		-.14		.617	-.35		.351
APE-1	Realised a function using a self-made circuit diagram.	-.14		.597	-.64	†	.065
APE-2	Transferred a circuit diagram to a PCB.	-.08		.783	-.17		.666
APE-4	Configured and assembled a desktop-PC.	-.16		.567	-.69	*	.039
APE-6	Programmed a micro controller.	-.20		.452	-.34		.367

Note. † = p < .10. * = p < .05. ** = p < .01.
Positive “superior learning mode”-values describe advantages of hands-on.
[a] = All participants of the VET/non-VET group answered identically.

Appendix E

RIASEC test set questions (DS-C)

The following appendix presents the test set questions as used in German. They were additionally translated to the English language.

E.1 Explorix [123]

E.1.1 Realistic

E.1.1.1 Interests

- elektrische Geräte installieren (anschließen, einbauen)
  install electrical equipment
- mit einem/einer guten Mechaniker/-in oder Techniker/-in zusammenarbeiten
  work with a good mechanic or technician
- aus Holz ein Büchergestell zimmern
  build a wooden bookcase
- ein Praktikum in einer Werkstatt machen
  do an internship in a workshop
- bei der Renovierung einer Wohnung mitarbeiten
  collaborate in the renovation of an apartment
- einen Kurs für Auto-Mechanik besuchen
  take a course in car mechanics
- ein Fahrrad reparieren
  repair a bicycle
- die Verstärkeranlagen für ein Popkonzert einrichten
  set up the PA system for a pop concert
- Maschinen oder Werkzeuge mit Motor bedienen
  operate motor powered machines or tools

- im Garten oder in der Landwirtschaft arbeiten
  work in the garden or in agriculture

- Metall bearbeiten oder etwas aus Metall herstellen
  work with metal or make something out of metal

E.1.1.2 Capabilities

- ich kann eine Stichsäge, Drehbank oder Schleifmaschine bedienen
  I can operate a jigsaw, lathe or grinding machine

- ich kann bei einem Auto das Öl oder die Reifen wechseln
  I can change the oil or tires of a car

- ich kann gut mit einer Nähmaschine oder Bohrmaschine arbeiten
  I can work well with a sewing machine or drill

- ich kann Möbel nach einer Skizze zusammenbauen oder reparieren
  I can assemble or repair furniture according to a sketch

- ich kann einfache Elektrik-Reparaturen ausführen
  I can do simple electrical repairs

- ich kann einen Fahrradschlauch auswechseln
  I can change a bicycle tire's inner tube

- ich kann mit einer Axt Holz zerkleinern
  I can chop wood with an axe

- ich kann ein verstopftes Abflussrohr reinigen
  I can clean a clogged drain pipe

- ich kann einfache Gegenstände aus Holz herstellen
  I can make simple objects from wood

- ich kann die Wände einer Wohnung tapezieren
  I can wallpaper an apartment

- ich kann einen Traktor, Bus oder Lkw fahren
  I can drive a tractor, bus or truck

E.1.2 Investigative

E.1.2.1 Interests

- wissenschaftliche Bücher oder Zeitschriften lesen
  read scientific books or magazines

- in einem Forschungsinstitut oder Labor arbeiten
  work in a research institute or laboratory

- einen Biologiekurs (oder Biologie als Wahlfach) besuchen
  attend a biology course
- ein Experiment mit Chemikalien durchführen
  conduct an experiment with chemicals
- sich mit einer wissenschaftlichen Theorie auseinandersetzen
  engage with a scientific theory
- einen Physik- oder Chemiekurs besuchen
  attend a physics or chemistry course
- eine neue Programmiersprache erlernen
  learn a new programming language
- Quellen und Dokumente über eine geschichtliche Epoche analysieren
  analyse sources and documents about a historical epoch
- an einem geistes- oder naturwissenschaftlichen Projekt arbeiten
  work on a project in the humanities or natural sciences
- Gesteinsschichten geologisch untersuchen
  geological investigation of rock layers
- über längere Zeit an der Lösung eines Problems arbeiten
  work on the solution of a problem over a longer period of time

E.1.2.2 Capabilities

- ich kann Algebra anwenden, um mathematische Probleme zu lösen
  I can use algebra to solve mathematical problems
- ich kann ein Experiment oder Projekt durchführen
  I can carry out an experiment or project
- ich verstehe den Begriff «Halbwertszeit» bei einem radioaktiven Element
  I understand the term "half-life" for a radioactive element
- ich kann eine mathematische Funktion grafisch darstellen
  I can graphically represent a mathematical function
- ich kann einfache chemische Formeln interpretieren
  I can interpret simple chemical formulas
- ich verstehe, warum Satelliten nicht auf die Erde stürzen
  I understand why satellites do not crash to earth
- ich kann einen wissenschaftlichen Bericht schreiben
  I can write a scientific report

- ich kenne die Urknall-Theorie über die Entstehung des Universums
  I know the big bang theory about the origin of the universe
- ich kann die Rolle der DNS in der Genetik erklären
  I can explain the role of DNA in genetics
- ich kenne Theorien über die Entstehung verschiedener Gesteinsarten
  I know theories about the formation of different types of rock
- ich kann einen Sachverhalt mit wissenschaftlichen Formeln beschreiben
  I can describe a situation with scientific formulas

E.2 AIST [122]

E.2.1 Realistic

- mit Maschinen oder technischen Geräten arbeiten
  work with machines or technical equipment
- Metall/Holz bearbeiten, etwas aus Metall/Holz herstellen
  working with metal/wood, making something from metal/wood
- Arbeiten verrichten, bei denen man sich körperlich anstrengen muss
  perform work that requires physical exertion
- in einen Computer Teile einbauen
  install parts into a computer
- Konstruktionspläne zeichnen
  draw construction plans
- elektrische Geräte oder Anlagen bauen
  build electrical devices or installations
- auf einer Baustelle arbeiten
  work on a construction site
- Servicearbeiten durchführen (reinigen, instandhalten, reparieren)
  perform service work (cleaning, maintenance, repair)
- etwas nach einem Plan oder einer Skizze anfertigen
  produce/construct something according to a plan or a sketch
- untersuchen, wie etwas funktioniert
  investigate how something works

E.2.2 Investigative

- sich mit unerforschten Dingen beschäftigen
  deal with unexplored things

- in einem Versuchslabor Experimente durchführen
  carry out experiments in an experimental laboratory

- etwas genau beobachten und analysieren
  observe and analyse something closely

- das Verhalten von Tieren oder Pflanzen untersuchen
  study the behaviour of animals or plants

- über längere Zeit an der Lösung eines Problems arbeiten
  work on the solution of a problem over a longer period of time

- chemische, physikalische oder biologische Versuch durchführen
  carry out chemical, physical or biological experiments

- ein Computerprogramm entwickeln
  develop a computer program

- die Ursachen eines Problems erforschen
  investigate the causes of a problem

- herausfinden, was man mit einem Computerprogramm alles tun kann
  find out what you can do with a computer program

- wissenschaftliche Artikel lesen
  read scientific articles

Appendix F

Used devices and simulations

This chapter of the appendix presents the employed cells and devices. Furthermore, it describes how they were represented in simulations.

F.1 Hands-on lithium-ion cells

Two different cell types (accumulators, or secondary cells) were employed in the laboratories:

- First, a 600 mA h $LiMn_2O_4$ based lithium-ion accumulator (see Figure F.1 and Table F.1 for details).
- Secondly, a $LiFePO_4$ based type of the same capacity (see Figure F.2 and Table F.2 for details).

Table F.1: Data of the lithium-ion $LiMn_2O_4$ cell employed in the laboratory

Cell type	IMR 14430
Manufacturer	Efest
Colour	red
Nominal capacity	600 mA h ± 20 mA h
Nominal voltage	3.65 V
Charge end voltage	4.20 V ± .05 V
Discharge end voltage	3.00 V
Max. charging current	1 C
Max. discharging current steady / pulse	3.5 C / 5.8 C
Cathode	$LiMn_2O_4$
Protective devices	none
Diameter	14.14 mm ± .05 mm
Height	42.65 mm ± .10 mm
Weight	15 g
Charging method	CC-CV

The cell types were prepared for Kelvin connection by welding 7 mm × 38 mm × 0.15 mm nickel strips to both poles to ensure that lead and contact resistances did not influence the measurements. The welding spots were in the centre of the nickel strips. The force and sense leads were connected to the cells by soldering stranded

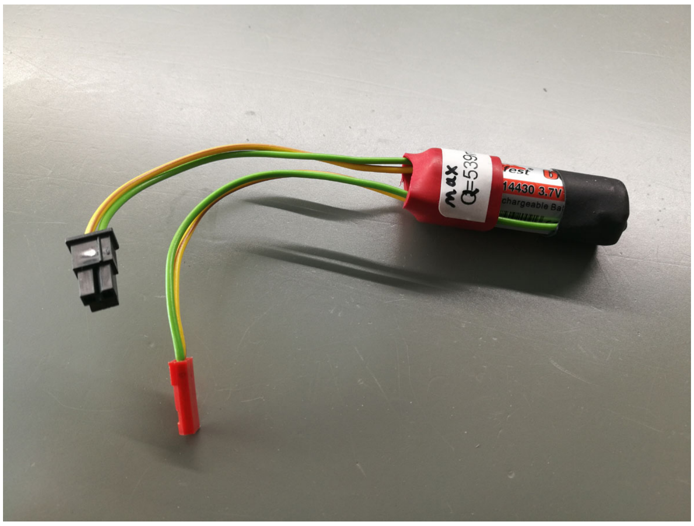

Figure F.1: Lithium-ion $LiMn_2O_4$ cell employed in the laboratory

wires to the nickel strips. Voltage drop on the sense lines caused by the current on the nickel strips was avoided by using the opposite side related to the force lines at the nickel strip for connection. The specimen thus defined the cell plus two welding connections. To allow for secure handling, plug connections fitting to the hands-on devices were prepared. Open contacts were covered by shrinking tubes to avoid unintended short circuits by students, e.g. when placing the cell on a metal surface.

Table F.2: Data of the lithium-ion $LiFePO_4$ cell employed in the laboratory

Cell type	IFR 14500
Manufacturer	unknown
Colour	blue
Nominal capacity	600 mA h ± 20 mA h
Minimal capacity	580 mAh
Nominal voltage	3.20 V
Charge end voltage	3.65 V ± .05 V
Max. charging current	1 C
Discharge end voltage	2.0 V
Max. discharging current steady / pulse	3 C / 10 C
Working temperature range charging	0 °C to 45 °C
Working temperature range discharging	−20 °C to 60 °C
Cathode	$LiFePO_4$
Protective devices	none
Diameter	14.2 mm ± .1 mm
Height	50.2 mm ± .2 mm
Weight	16 g
Charging method	CC-CV

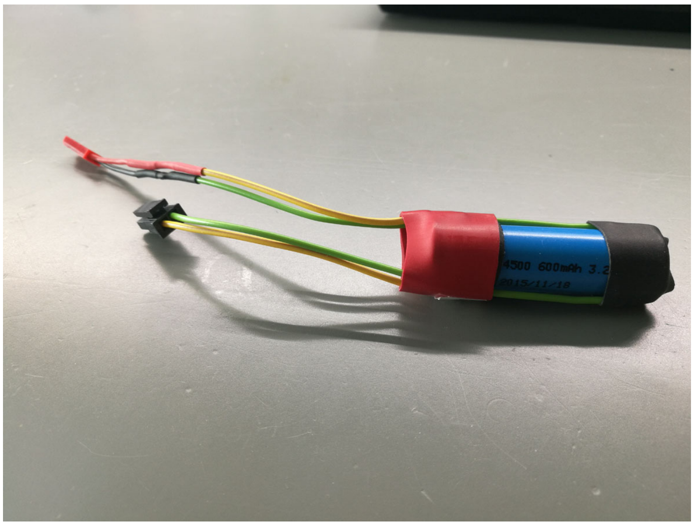

Figure F.2: Lithium-ion $LiFePO_4$ cell employed in the laboratory

F.2 Custom-made hands-on devices

After the initial design of the laboratory experiments, a specification sheet for the necessary battery test system was created. No fitting device was found on the market.

Ready-to-use industrial battery cell test benches and temperature cabinets are generally large and costly devices. Using such bulky and costly equipment makes it practically impossible to conduct laboratory work in small learning groups. Thus, the development of small-sized and flexible cell test systems allowing safe and effective individual learning experiences for the students was initiated. The development was funded by the German Federal Government's Showcase Program "Academic Education Initiative on Electro-Mobility". After several development steps, funding by the Faculty of Electrical Engineering and Computer Science allowed a complete teaching laboratory at THI to be equipped with 13 battery test systems. The main development and production of the devices for the laboratory was finalised before the educational research for the thesis started. The result is shown in Figure F.3.

The developed bench-top test system supports temperature-dependent experiments with lithium-ion cells and other types of energy storage cells, while a safety shut-off module ensures a high degree of safety during operation at all times. The combination of all functions required for an energy storage laboratory in a compact design is a unique feature of this test system. The described battery test system allowed students to conduct the hands-on experiments efficiently. Using this system, the students can carry out practical experiments on accumulator cells in a safe, flexible learning environment.

Finally, the custom-made battery test bench allowed for the second research phase, where firmware changes were applied to hide the simulations (see subsection 4.2.9).

The rest of the section is structured as follows: First, a short overview of the components of the test bench is given. Later, the individual components are discussed in detail.

Galvanostat

Firstly, this battery test system allows accurate current and voltage measurement, and also acts as a powerful and responsive electrical sink/source. The challenge, aside from the technical aspects, was to find the best compromise between cost and accuracy. For details, see subsection F.2.2.

Thermostat

Secondly, students must be able to comprehend the strong temperature dependence of battery cells. This can be achieved using industrial temperature cabinets, but they are space-consuming and optimised for testing larger test objects. Therefore, they are unable to change temperature quickly enough to suit the timing needs of laboratory work at university. These problems were solved by integrating a small, inexpensive,

Figure F.3: Employed hands-on devices, on the photo the device including an example cell is shown.

and extremely fast responding thermo-electric cooler into the battery test system. For details, see subsection F.2.3.

Safety switch of module

Thirdly, lithium-ion battery cells are potentially dangerous. A safety shut-off monitoring module was created to conduct hand-on laboratories safely. The tutor can parametrise this module to allow students more autonomy in controlling the battery system without the need for an increased number of supervisors. Ensuring full safety for the hands-on participants also meant that this aspect was not a limitation for the lesson design, as hands-on laboratories were just as safe as simulated ones. For details, see subsection F.2.4.

F.2.1 Battery test system / housing

The battery test system is based on a 19-inch tabletop case of three rack unit height. Its housing consists of three modules with standardised plug-in technology according to IEC 60297 3 101. It allows simple adaptation of the test system to the desired functional range or future developments through connectors according to IEC 60603 2.

The functions of the three modules that are shown in Figure F.4 and Figure F.5 cover all needs of the energy storage laboratory.

An ARM Cortex M4 processor was used as an arithmetic unit in all modules. The housing – including all sub-modules and integrated power supply – weighs 11.5 kg.

Production by students
As even the production of the tester was used to educate students, the circuit boards were designed by the researcher in a way to enable easy manual assembly by (paid) student assistants in a separate project. Thus, most components are in the relatively large SMD format 0805.

Figure F.4: Battery test system for hands-on laboratories. Students strengthen their basic understanding of electrochemical storage and learn about the behaviour of cells.

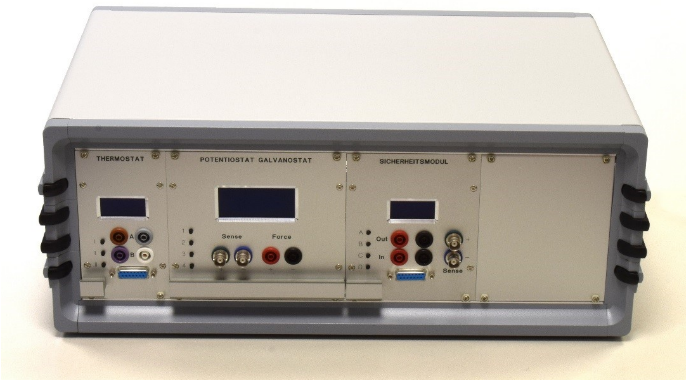

Figure F.5: Battery test system, from left to right: thermostat, potentiostat/galvanostat, safety switch-off module

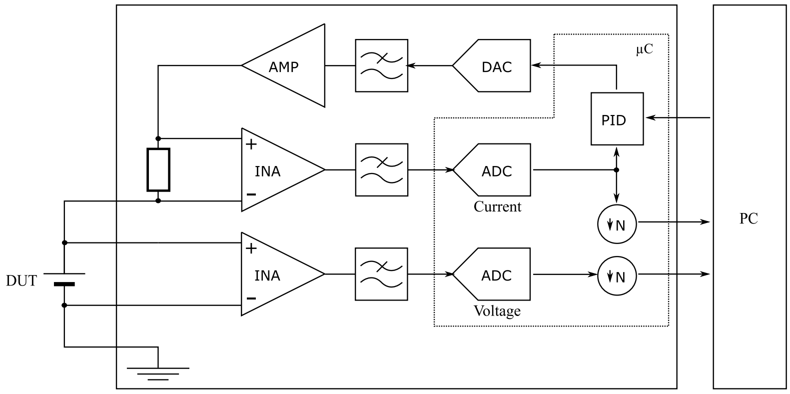

Figure F.6: Simplified working principle of the galvanostat. Actual current and voltage are converted into digital values using two synchronised ADCs. After downsampling, the values are submitted to the controlling computer. The current is compared with the target current by software which also controls the DAC.

F.2.2 Galvanostat

The central element of the battery test bench is a potentiostat/galvanostat module. It allows conducting of all laboratory experiments of chapter G with small single cells (amp-hour capacity up to 600 mAh). With limited C-rates, it also supports experiments with cells of bigger amp-hour capacity.

Primarily, currents and voltages can be applied to the cell and be measured time-discrete. The voltage range is symmetrical for measurements up to $\pm 12\,\mathrm{V}$, with currents up to 8 A for discharging and 4 A for charging.

Apart from the technical implementation, finding a good compromise between cost and production efforts and measuring accuracy was challenging. A test system should be easy to build and to maintain, and at the same time, it should be accurate enough to provide a realistic test environment for the students. Therefore, the focus was laid on optimised measuring accuracy and not on the exact analogous injection of current and voltage.

If the system injects inaccurately, the resulting errors will automatically be corrected by the measured data (see Figure F.6). When calibrated for voltage measurements the error is below 10 mV, and for current measurements the error stays below 5 mA in the specified temperature range (21 °C $\pm$ 5 °C).

The down-sampled data rate for the time-discrete transfer to the computer is adjusted automatically in the range between 312 Hz and 625 Hz. Figure F.7 presents the concept to provide different data-rates (all of them correctly anti-aliased) for different purposes. For example, besides the right data transfer rate, higher frequencies are used to control the voltage or current in a control loop, and lower frequencies are used to provide data for the LC-display on the front panel of the device.

The galvanostat/potentiostat may be controlled from a laboratory computer via serial communication (USB). C# and Java-based control programs are available as well as a LabVIEW driver.

While the C# and the LabVIEW software were developed by the author of the the-

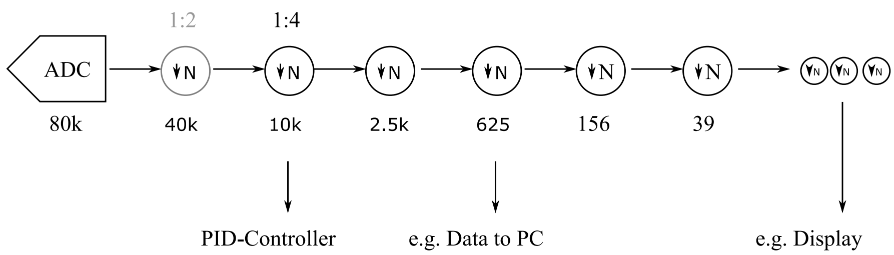

Figure F.7: Concept for down-sampling. Before reducing samples, in each stage, a low-pass filter is implemented to avoid aliasing.

sis, the Java-based program was written by a master student (master's thesis) before the first laboratory run R1. The Java program was used for the student laboratory, because it was designed similar to software used for industrial test equipment regarding architecture and user interface. It allows learners to freely program the procedures (including loops and jump conditions) to implement battery experiments. The conformity to industrial test benches prepares the students for professional life. The user interface of the program was also used for simulations and is presented in section F.4.

Impedance Spectroscopy

To remove the needs for additional devices, the firmware of the battery cell test system was enhanced to allow frequency-dependent injection and measurement. In addition to the DC component, an AC voltage of lower amplitude can be overlaid for impedance measurements (adjustment steps are presented in Figure F.8). To adjust the amplitude, the PID-controller seen in Figure F.6 is operated based on the computed amplitude instead of the direct time-discrete data used in the pure DC-mode. In the impedance mode, the device determines the impedance's magnitude for frequencies up to 10 kHz and its phase for frequencies up to 5 kHz. In this mode, the test system offers a continuous frequency range with 1 Hz resolution.

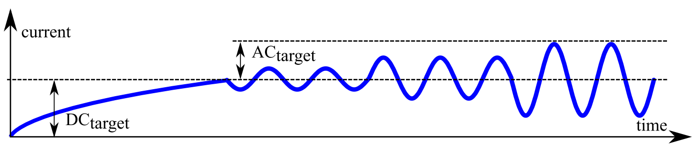

Figure F.8: For injection of currents to perform impedance measurements, the galvanostat first adjusts the DC current (in most cases 0 A), and in a later step increases the AC part in the requested frequency until the requested amplitude is achieved.

The cell test system determines the impedance' magnitude up to 10 kHz, and the phase up to 5 kHz. The enhanced firmware has a fine resolved frequency range, whereas many other devices work only on a small number of discrete frequencies. This was achieved by means of a combination of

1. sampling rate modification in the range between 40 kHz and 80 kHz,

2. flexible access to the output data of a ladder of down-sampling-steps

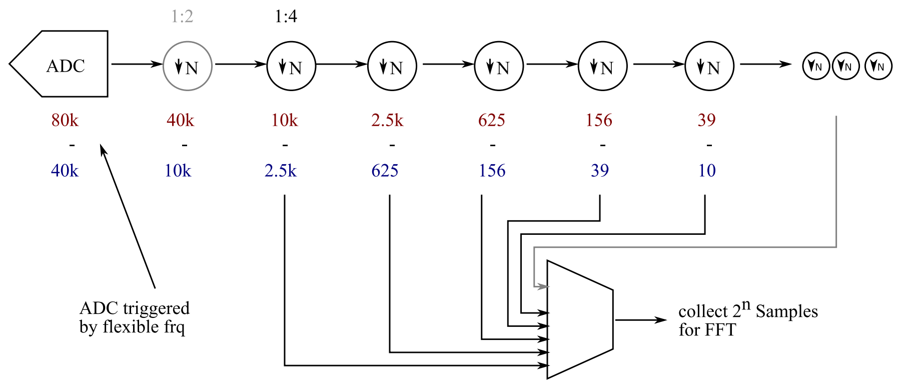

Figure F.9: Concept for down-sampling, with an additional flexible sampling rate of ADC and DAC, and data source switch to adapt to different target frequencies for the impedance measurement function. Before reducing samples, in each stage, a low-pass filter is implemented to avoid aliasing.

3. modification of the length/size of the discrete Fourier transformation (DFT) window and

4. the employed coefficient of the DFT.

The principle is presented in Figure F.9.

Applying an additional window function was not necessary, as the injected frequency is known, and DAC, ADC, and FFT-window are synchronised all the time.

To reduce the computing effort of the device's micro-controller, as on option, only the main spectral line is computed by the Goertzel-algorithm [150].

Reached results

The impedance measurement was validated against a Zahner ZENNIUM electrochemical workstation supported by the Zahner PP241 4-quadrant power potentiostat (results shown in Figure F.10). Uncalibrated, the self-developed device showed acceptable results for student laboratories (difference $< \pm 7\%$ for impedance, $< \pm 4°$ of phase, tested with a 600 mAh $LiMn_2O_4$-Cell in the range from 100 Hz to 8 kHz, at 1 A amplitude).

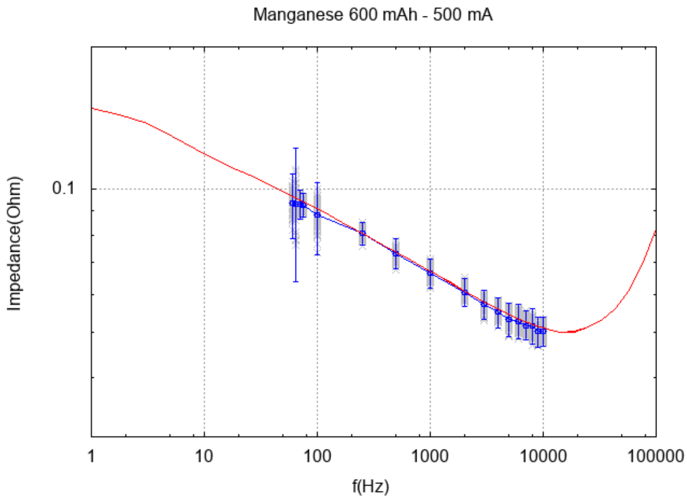

Figure F.10: Result of impedance spectroscopy gained with the developed device (blue). The error bars are plotted in blue. A reference measurement was performed with a Zahner ZENNIUM device (highly accurate, red)

F.2.3 Thermostat

A major challenge for electric mobility is the strong temperature dependence of parameters of electrochemical energy storage systems [8]. Thus, it is important that students can comprehend the effects of high and low temperatures on the performance of battery cells. Unfortunately, this cannot be achieved using industrial temperature cabinets, which are optimised for testing larger test objects. These cabinets are too big, expensive and do not allow changing temperature of small single cells quickly enough to conduct a laboratory within two to three hours.

Besides the requirements on the equipment, the waiting time for a homogeneous temperature distribution within the cells must be considered for lesson preparation. The cooling performance is especially essential to be able to carry out various temperature dependent measurements during a lesson. For example, to conduct several tests on the internal resistance of a cell according to ISO 12405-1 (subsubsection G.7.1.2) within a teaching unit of three hours, cooling from room temperature to -18 °C was required within 20 minutes. The achieved speed was analysed and is presented in Figure F.11.

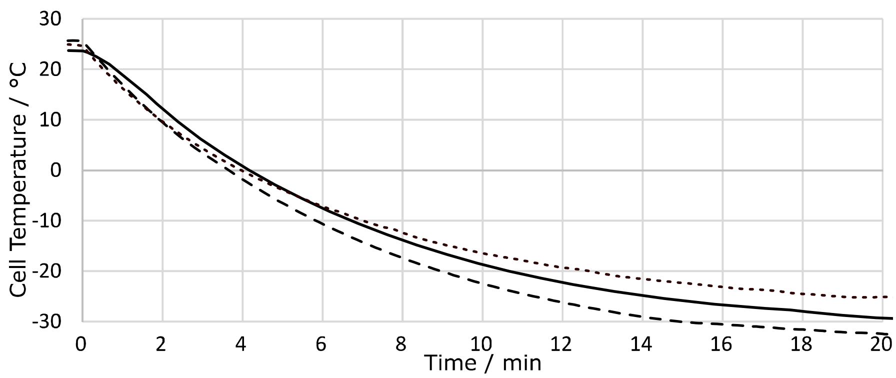

Figure F.11: Max. cooling performance of three exemplary heating/cooling systems, T_{env}=25 °C

To avoid the disadvantages of convection cooling or fluid thermostats and to solve the problem of minimal time, a small, cost-effective, and rapidly responding two-stage thermo-electric cooler as shown in Figure F.12 was integrated into the battery test system.

Cooling and heating are based on Peltier elements, which allow temperature control in a range from -25 °C to 50 °C. The working principle is presented in Figure F.13.

To achieve rapid cooling (Figure F.11), the test cell is in direct contact with the metal on four sides. The chamber is made for cells with maximum dimensions of 90 mm $\times$ 34 mm $\times$ 30 mm. The cooling stages are driven by two modulated H-Bridges, providing up to 250 W of power. Heat flux losses are reduced by a foam cube, which insulates the cooling device from the environment. The chamber temperature is measured with an absolute error below 1 K.

This external cooling system weighs 5 kg, its dimensions (26 cm $\times$ 26 cm $\times$

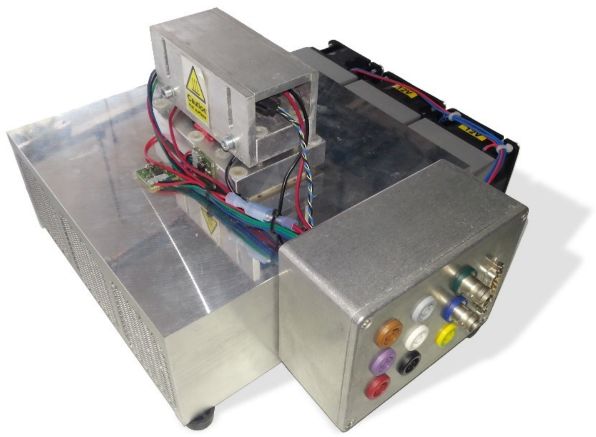

Figure F.12: External heating and cooling system

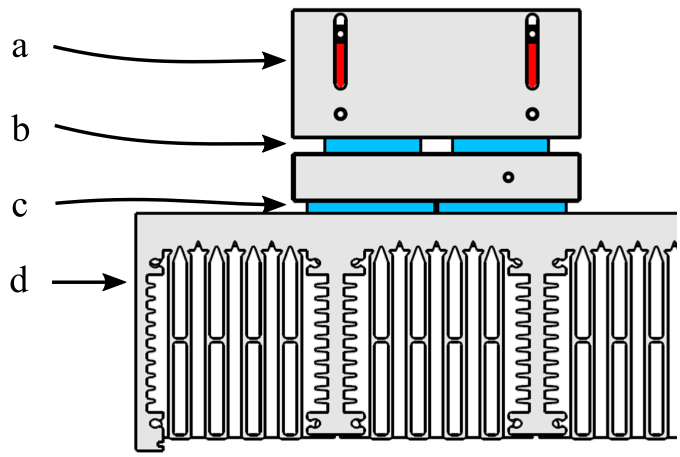

Figure F.13: External heating and cooling system – Working principle a) cell chamber with direct contact cooling, b) upper Peltier stage, c) lower Peltier stage, d) heat sink with fans

20 cm) allow for easy storage while not in use.

The thermostat supports remote control, utilising the galvanostat/potentiostat module as mediator between the USB-communication and the CAN-based bus between all the modules.

F.2.4 Safety-switch-off module

Lithium-ion battery cells are potentially dangerous. In the case of mistreatment, they may burst or burn [8]. To ensure students' safety at all time, experiments must be either greatly simplified or prepared in step-by-step instructions, which the student has to follow without any degree of freedom. However, this style of instructions would preclude students from gaining skills in planning their experimental work which are at the core of the engineering profession.

To enable students to gain skills in planning experimental work, a custom-made safety switch-off module was integrated into the test system. This module is pre-parameterised by the instructor and allows the students full autonomy in controlling the battery system, while minimising the need for supervision.

The module monitors three parameters of the test specimen:

- current (± 12 A),
- voltage (± 12 V),
- temperature.

If the determined values of these parameters fall outside of the range permitted by the instructor, the module automatically disconnects the cell from the test system (the arrangement is shown in Figure F.14). It is important to note that the switch-off value may be configured to be pulse-length dependent, as most battery cells allow pulses outside the cell's specified range for a short time (an $I^2 \cdot t$ behaviour can be configured).

Since this safety switch-off module cannot be influenced – intentionally or not – by a student, the student has full freedom to program the potentiostat/galvanostat. He/she may gain valuable learning experience through trial and error.

The existence of the safety switch-off module allowed the hands-on experiments to be designed identical to the simulated experiments, as safety was ensured at all time.

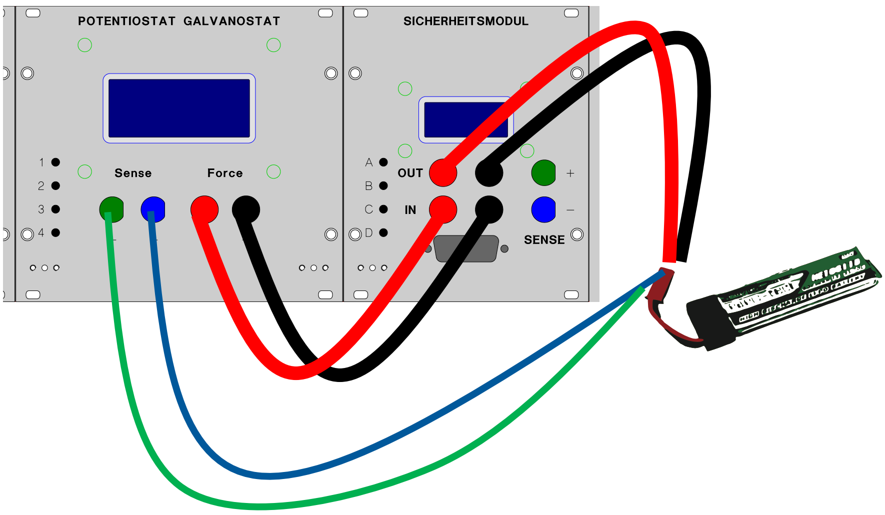

Figure F.14: Connecting the safety switch-off module. The connection supports Kelvin arrangement. While the voltage sense lines can be attached to both devices, the force lines are wired in serial connection to allow the module to interrupt the current.

F.3 Simulation model

To conduct the study, a simulation model was prepared. It was employed to replace the hands-on devices and cells with simulated experiments in the first phase. In the second research phase, the same simulation model was employed – but hidden from the students' perception.

To guarantee that the experimental results achieved in the simulation mode were similar to those of the hands-on mode, battery cells and the behaviour of the employed hands-on devices were analysed in order to parametrise the underlying simulation model[1].

Black box simulation
Only the input and output data were visible to students. The model itself, as well as the cell parameters and internal computed values of the simulated cell, had not been released to students (black box model). Such an arrangement ensured that all participants involved in simulations had the same information as their peers working in the hands-on condition.

Steady running real-time simulation
The simulation of cell behaviour always started after opening a graphical user interface regardless of whether experimental procedures were running or not. Thus, simulated cell behaviour also included aspects regarding the global design of the experiment (e.g. the cell cooling down for an appropriate time between two experiments).

F.3.1 Cells

The battery cells were analysed in order to parametrise the underlying cell simulation model (see Figure F.15).

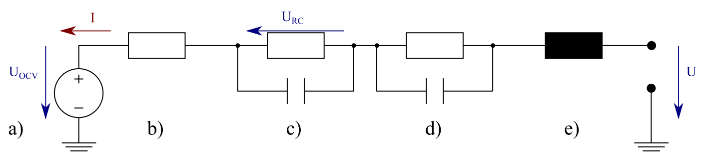

Figure F.15: Simulation Model of the electrical characteristics of the battery cell. a) Open Circuit Voltage b) Internal Resistance c) & d) Cathode and anode double-layer capacitance and charge transfer resistance e) Inductance. Such models are widely used in battery simulation/characterisation to meet the frequency-dependent behaviour of cells (e.g. [151]). In the actual case, the model was improved by state-of-charge and temperature-dependent behaviour of b).

The part of the simulation model which cared about the behaviour of the simu-

[1]The programming of the simulation was performed by Alexander Nitsche, and was clearly not a part of this thesis, see the acknowledgements, page iii. The characterisation of the cells (measurements and deriving parameters) was performed by the author. The employed simulation model was developed in cooperation with Alexander Nitsche.

lated lithium-ion cells consisted of:

- A base model for the state of charge dependency of the open-circuit voltage.
- A model of the internal resistance, which depends on the temperature, the state of charge, the direction of the current plus a dependency on the intensity of the current in case of low temperatures.
- One RC-branch to simulate double layer and charge transfer resistance behaviour with static parameters, as described in subsection G.10.1.
- A thermal battery model based on the heat created by losses and a model for the heat transfer to the thermostat, as described in subsection G.10.1.
- A model for inductive behaviour, which had little to no influence on the student experiments.

The model was prepared for both types of employed cells. All student experiments were conducted several times using the real devices and cells as well as using the simulations. The parameters were optimised to reduce deviations in overall behaviour as well as focused onto the results of the student laboratory experiments. As a result, simulations closely imitated the actual behaviour of battery cells.

Furthermore, before the laboratories, small deviations regarding the capacity $\pm 5\%$ and the internal resistance $\pm 5\%$ were finally added to the configuration files of the different computers, to avoid identical outcomes of different groups' experiments.

F.3.2 Simulation of the hands-on devices

Galvanostat/potentiostat

The simulations of the galvanostat/potentiostat included a small de-calibration of the measured voltage and current offset, as well as quantisation and noise (compare [12, p. 128, 152]) of the measured voltage and current. The quantisation and noise parameters were compared to the results derived by the real devices.

As a result, in the second phase, it was not possible to distinguish the hidden simulation set-ups from hands-on set-ups, even for the lecturers.

Thermostat

The thermostat was simulated based on the Peltier base equation for conducted heat flow and a simulation of the heat capacity of the metal parts of the thermostat.

Safety switch-off module

The simulated safety switch-off-module acted on the same break conditions as the real device. In the second research phase, the simulation reacted on the open/close button and also operated the relays to emit the same sound like in the hands-on condition when opening/closing.

F.4 The client GUI controlling both

As part of a research master's thesis[2] basic software was developed to control the experiments.

The GUI is structured very similarly to industrial test benches. It allows programming a linear sequence of constant current and constant voltage phases. In each phase, the user/student can configure three things:

- The physical dimensions the user wants to set/force, e.g., current, voltage, or temperature.
- The physical dimensions the user wants to record in the log files and plot.
- The condition which needs to be fulfilled to end this phase and continue to the next one.

The GUI supports enhanced control structures like conditional jumps, loops and other functions. These functions were not necessary to execute the laboratory experiments.

Figure F.16 shows the user interface. In the left column (Tool Box), the available devices are shown. By adding nodes in the centre column (Sequence) the user adds phases for the test sequence. Drag and drop operations from the Tool Box allow intuitive usage. In the right column (Edit) the user can configure the properties of the current selected node, e.g. the three above-mentioned properties.

After pressing the "Start Sequence" button, the sequence is executed, and a real-time graph is shown (see Figure F.17).

Support for remote experimentation

The software employed in the study at hand used a client-server structure to control the devices and simulations, capable to communicate through the network. This would also support remote experimentation.

[2]The GUI was developed and created by Alexander Nitsche during his studies, and was clearly not a part of this thesis, see the acknowledgements, page iii.

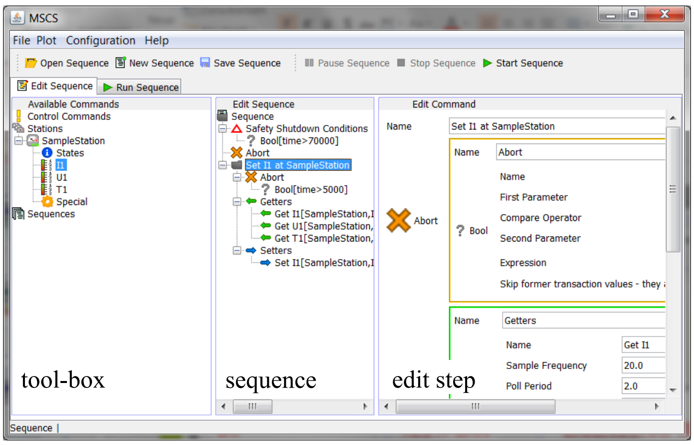

Figure F.16: GUI – Programming a sequence

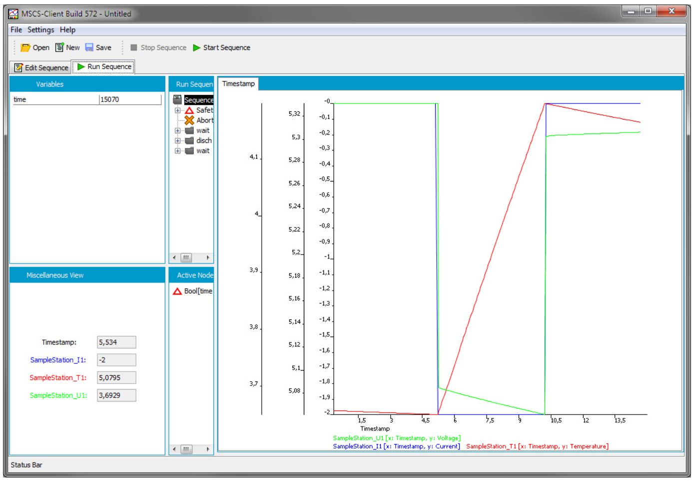

Figure F.17: GUI – Evaluation of measurements

The program was designed to be compatible with different laboratory devices. For each of these devices, a driver (server) which communicates with the GUI (acting as the client, and capable of controlling several servers at the same time) needs to be provided. The client communicates with the servers using a TCP/IP protocol.

For all working teams, communication was limited to the local computer, to avoid interfering sessions in the network and leading to a *local* experiment in all modes of the study.

Support for both study phases

In the first phase, the client controlled two different servers:

- First, the real hands-on devices server, which communicated to the hands-on devices via serial communication (in the laboratory).
- Secondly, the simulated devices and cells (in the computer pool).

In the "hidden simulations" mode of the second research phase, the client GUI controlled the manipulated hands-on devices server. This server was configured to show simulated results on the device displays while ignoring the actual measurements.

Outlook

The devices and software developed are not only suitable for training, but also for actual research on battery cells. The development of this software still continues: as of now (2020), it is possible to operate the hands-on devices in combination with other laboratory equipment (power supplies, electric load, precise thermometers) to allow complex laboratory experiments.

F.5 Quality checks/Reviews of the employed tools and simulations

It is crucial that the author built the employed tools, and tooling quality might influence the present research outcome. Thus, a review of the self-created tools, simulations and activities was performed by colleagues of the author.

The laboratory engineer regularly calibrated the developed hands-on tools as they were also used for professional battery research in the laboratory by various (senior) staff members when no teaching laboratory was running. Thus, the likelihood of some kind of error going undetected is very small. In fact, that these tools were used by staff in actual research, i.e. outside the realm of teaching, also serves as a testament to their reliability.

The simulation model used is based on the standard cell model of Prof. Dr. Hans-Georg Schweiger (associate supervisor of the thesis). He has several years of experience in energy storages. The model was further developed and parametrised by the author and Alexander Nitsche (four-eyes principle), both experienced in working with battery models and battery management algorithms.

The author of this thesis is regularly deployed for advanced company training to teach energy storage specialists in battery models and battery management. The simulation model was parametrised against determined cell parameters and the experimental results of the experiments performed in the hands-on mode. The laboratory engineer double-checked these results. While this does not rule out all possibilities for errors in the simulation, the fact that the results of the second study phase (eliminating bias towards simulations) showed no difference also strongly supports the view that the simulations were equivalent to the hands-on experiments.

Appendix G

Contents of the battery laboratory

Academic qualification is essential for the success of electric mobility, and the German government supports education in the field of electro-mobility [92, 153, 154].

It was difficult to find much research on teaching energy storage in practical lesson formats. Many scientific experiments regarding accumulators are described in research papers, but only a few experiments are intended for education.

In 2015, Norian [155] published a teachable experiment on a lead acid battery. Standard laboratory equipment such as power supply, load resistor and switch were employed. The experiment did not use any especially developed hardware for student training, making it easy to transfer to other laboratories. In 2018, Thanomsilp *et al.* [156] publicised a laboratory to build a simple battery with students, and Domínguez *et al.* [58] published another experiment in the field of energy storages (electrolyser for hydrogen production). Unfortunately, these authors did not examine the impact of educational benefits of their approaches.

Many universities provide instructions for student experiments, but these are only available to a limited audience. Thus, the laboratory sessions described in the following may give other educators a base to start their own battery laboratories.

This following chapter focuses on the pedagogical intentions, including the identified learning targets of the laboratory. The technical theory behind the lessons, the derived learning objectives, the lessons itself are described.

The laboratory experiments are primarily formulated in the context of the German Bachelor of Engineering program "Electrical Engineering and Electric Mobility" at Technische Hochschule Ingolstadt, as most of the data were sourced from these runs (R1, R2, R6, R9).

G.1 Information on the study program "B. Eng. Electrical Engineering and Electric Mobility"

The Technische Hochschule Ingolstadt (THI) offers a wide range of qualification programs in the field of electric mobility. Students may choose between various full-time and part-time bachelor and master degree programs, which include lectures on electrochemical energy storage systems.

The bachelor study program provides know-how in the fields of electro-mobility and energy storage. With the profile of a University of Applied Sciences, the consolidation of theoretical knowledge by means of practical experiments in the teaching laboratory is of particular importance [5]. Besides technical knowledge, skills for working independently as an engineer are taught through practice-oriented teaching. The study program is held in German. The required specialist knowledge is gained in an application-oriented manner by undergraduate students [157].

The bachelor's programme offers broad application-oriented education in the field of electro-mobility distributed over seven semesters. In many subjects, the lectures are accompanied by practical training, as laboratories promise to improve the outcome of student learning. To understand the context, it is necessary to understand which topics were taught in the semesters before the energy storages laboratory, which is grouped with a lecture given in the fourth semester.

The Bachelor's programme aims to train engineers who can develop components and systems in the broad field of electro-mobility. In the first semesters, students learn the engineering basics of mathematics and science. [158]

1st semester

- Introductory Project
- Applied Physics ^{l}
- Engineering Mathematics 1
- Electrical Engineering 1
- Basics of Programming ^{l}
- Electronic Components

2nd semester

- Engineering Mathematics 2
- Electrical Engineering 2
- Measurement Technology ^{l}
- Digital Technology
- Signals and Systems

3rd semester

- Modelling Dynamic Systems ^{l}
- Fields and Waves
- Circuit Technology ^{l}

- Digital Signal Processing [l]
- Physical Chemistry [l]

Following a background in electrical engineering subjects, the study program is more intensely concerned with the area of electro-mobility. The required specialised knowledge in the area of components such as electrical motors, power electronics and energy storage is presented in an application-oriented manner. The necessary competencies at the application and vehicle level, as well as safety aspects, are also taught. [158]

4th semester

- Control Engineering [l]
- Microcomputer Technology [l]
- Energy Storage [l], **
- Vehicle Electronics
- Power Electronics [l]

[l] That subject was grouped with a laboratory.

** This laboratory was part of this study.

Semester 5 to 7

In the following semesters, a one semester industry internship, interdisciplinary subjects (project management, business administration), as well as mobility specific subjects (electric and hybrid vehicles, driving dynamics, mechatronic components, and a technical project) are scheduled. After the joint study, students have to choose between several elective subjects and can deepen their studies according to their preferences [158].

G.2 Information on the energy storage study module

The energy storages module consisted of a theoretical lecture (66% of presence time) and an energy storages laboratory (33% of presence time). In sum, approximately 70 h of presence time in lectures and 80 h of self study time were planned for the subject.

The main learning targets of the theoretical reading were to enable the students to [158, p. 50f]

- describe electro-chemical energy storage systems and their characteristics.
- name and classify the basic characteristics of the different types of storage devices and transducers.

- name and describe the reactions and side reactions of the individual battery types as well as the essential ageing mechanisms.
- apply methods for the simulation of energy storage devices and develop battery models.
- assess the use of different electro-chemical energy storage technologies and select the optimal storage type for the respective application purpose.

The theoretical reading dealt with: [158, p. 51]

- Function and layout of energy storages
- Parameters of batteries, influencing variables and measuring methods
- Primary cells, secondary cells (lithium-ion, lead-acid, nickel-metal hydride), tertiary cells (fuel cells), double-layer capacitors
- Mechanical structures, physical/chemical processes, thermodynamics, current and voltage characteristics, and cell type selection
- New cell technologies and development trends
- Modelling of batteries (current and voltage behaviour, and ageing)
- Algorithms for battery condition determination (SoC, SoH ...)
- Dimensioning of battery systems
- Charging technology

A 90 minute test needed to be passed to finalise the module. The laboratory part had to be completed successfully before being admitted for this test [158, p. 50].

Both parts of the module ran parallel in the same semester. The laboratory began approximately two weeks after the lecture series to give the students the opportunity to gain the theoretical knowledge necessary to understand the background and significance of the first experiments.

G.3 The energy storages laboratory

This section describes the content taught in the laboratory – including the technical background, the derived learning objectives, and the procedure of the lessons.

The main targets of the laboratory were:

- To build up universal knowledge for the students future careers'.
  - To transfer a basic understanding of (electro-chemical) storage systems.
  - To teach the practical behaviour and most important parameters of battery cells.

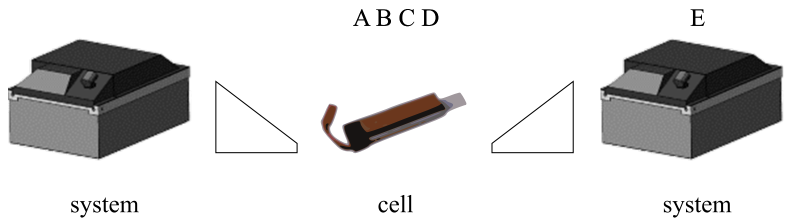

Figure G.1: Concept of the laboratory. The system level is scaled down as of safety and effort reasons to the cell level. The experiments are performed on a cell level. The experiments lead to a Matlab/Simulink workshop on system level.

- To enable students to determine the parameters of battery cells self-sufficiently by means of an appropriate experimental setups.
- To prepare students to work with industrial test benches.

The learning objectives of the individual lessons were included in the instructions, meaning the students were aware of them [10, 12, 35]. Electro-mobility students who are employed in the energy storage field in later professional life may work with industrial test benches. To prepare for this work, devices used in the laboratory are controlled very similarly to those industrial devices. The computer program controlling the devices allows for programming all the same sequences as industrial test benches. The laboratory instructions were tailored to support theoretical knowledge acquired from the accompanying theoretical class. Parts of the laboratory were piloted before the first run in summer 2016.

G.3.1 General structure of the energy storage laboratory

The general concept of the laboratory is shown in Figure G.1. For reasons of safety and resources, the electric vehicle battery system was scaled down to one cell. The laboratory experiments of the four content areas A to D are performed at the cell level. The main content areas were: A Four-conductor measurement/Contact Resistance/Insulating resistance; B Open-circuit voltage curve; C Internal resistance, Power and D Energy & Capacity. The learning objectives were taught during five discrete three-hour laboratory sessions. These laboratory sessions lead to a workshop in which the individual students' measurements were used to parametrise a simulation model and to design an energy storage system in Matlab/Simulink.

The simulation workshop after the experiments was also helpful for educational research, as it could be used for the test on the preceding lesson (see Figure 4.1, Content E was the Matlab/Simulink Workshop, within this session the Test on Content D was performed). The biggest challenge in creating the experiments was the limited time for a single student's laboratory, as some of the relevant processes (e.g. cooling down, discharging or charging a cell, waiting for stable states) are extremely time consuming. It was often necessary to provide differently parameterised experiments to subgroups (e.g. discharge rate) and to ask the students to consider later data from all groups for evaluation.

G.3.2 General course of action during each lesson

The instructions for the experiments were available to all participants on Moodle (the university intranet [127]). The laboratory instructions were written in a way that allowed learning by trial and error, as recent literature consists of hints that overly structured laboratories have reduced educational benefit, e.g. [66, 159]. The instructions were specifically tailored to support knowledge acquired by students in the accompanying theoretical class. The content was designed to deepen knowledge and understanding. The data sheets of the (simulated) measurement equipment and cells were also published on Moodle beforehand. Where possible, the students were free to choose parameters, depending on their recent results, reflecting findings by other researchers that reduced guidance supports student learning [64, 66].

At the beginning of each lesson, a short discussion of the lesson target, including a discussion on preparative questions the students had to prepare, was conducted on the blackboard. The students were required to fill out a laboratory journal ("measurement protocol") which they had to present to the lecturer before leaving the lab. The respective journal pages were attached to the experimental report, which had to be submitted via Moodle within one week after the experiment. For the report, a structured list of items to address was provided. The report was evaluated on a pass/fail basis. In case of a fail, the experiment had to be repeated. Only one retry was allowed.

All journals and protocols were generated in the working teams.

G.4 Introductory meeting

In each German Run, the first meeting was dedicated to introducing the laboratory rules. Also, data collection for the research, such as creating groups and collecting background info from the students, was performed.

Laboratory rules

The students were introduced to the following rules regarding the passing of the laboratory:

- failing the laboratory leads to exclusion from the test of the theoretical lecture (the laboratory is a pass/fail subject)
- if a session is missed, it has to be made up at a later date.
- proper preparation according to the experimental instructions is required of all students. Arriving unprepared leads to exclusion from the session, which has to be made up at a later date.
- only one date for make-up sessions was available, so missing two sessions meant failing the module

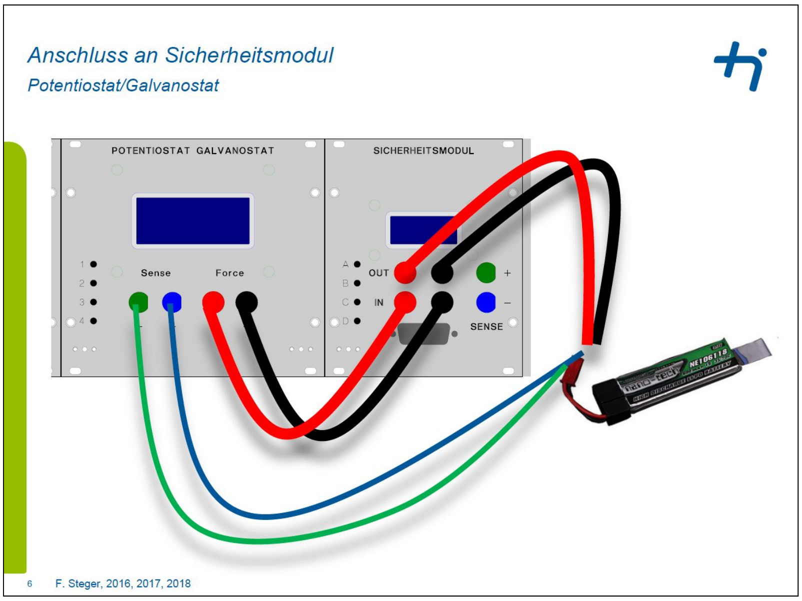

Figure G.2: Screenshot of the device presentation, wiring in Kelvin arrangement

Safety instructions

Lithium-ion cells can be dangerous items when treated wrongly. Students were provided a short introduction in handling these cells. The main points were avoiding over- and undercharge, abnormal temperature, mechanical defects (like bending, penetration, or pressure), and charging at low temperatures (see also subsection F.2.4).

Introduction to the battery test devices and the control software

The battery testers were presented to the students. First, they were instructed in how to wire the thermostat and presented the function of this device. After that, the function of the galvanostat, including how to wire the device to a battery in Kelvin-Arrangement, was demonstrated. Students were also shown the wiring (Figure G.2) and usage of the safety-module.

The control software was presented (Figure G.3), using a CC-CV discharge as an example, as it includes all necessary parameters of the laboratory. The students received information on where to find the program's logged data.

Participant information/Short introduction of the educational research project

The educational research project was introduced, and the participant information sheet according to the Royal Melbourne Institute of Technology ethics approval was handed over. A copy of that document was available in the university intranet.

Amount of Practical Experience questionnaire

The participants filled out the questionnaire regarding the "Amount of Practical Experience" (section 4.3), which allowed a quick formation of the two-semester groups and publication of this information on the university intranet the same day. The type and size of the questionnaire collecting the participants' background information on

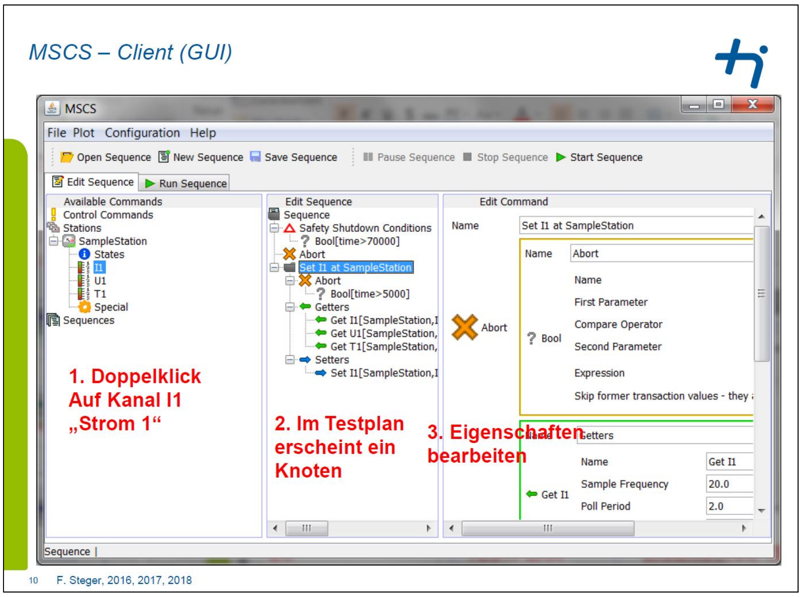

Figure G.3: Screenshot of the device presentation, use of the control software

the participants differed between runs, as the questionnaire was enhanced and modified throughout the study, see section 5.1.

G.5 Lesson A – Four-terminal sensing, contact resistance, insulation resistance, and flash-over voltage

G.5.1 Theoretical technical background

G.5.1.1 A1: Accurate measurements of low ohm resistances

Four-terminal sensing method

For the efficiency of battery systems, it is of particular importance to reduce the internal resistance of the system as much as possible. Thus, future engineers need to know how to determine these resistances accurately. These low resistance measurements can be applied to many components. Lithium-ion battery cells have very low internal resistance. While selecting the cells, conductors and passive electromechanical parts need to be connected with the lowest resistance possible. Also, fuses and contactors create less heat if types with lower resistances are selected.

Measuring a low ohmic resistance located away from an ohmmeter creates a measurement difficulty, as a standard ohmmeter measures all resistances in the circuit loop, as shown in Figure G.4. [160]

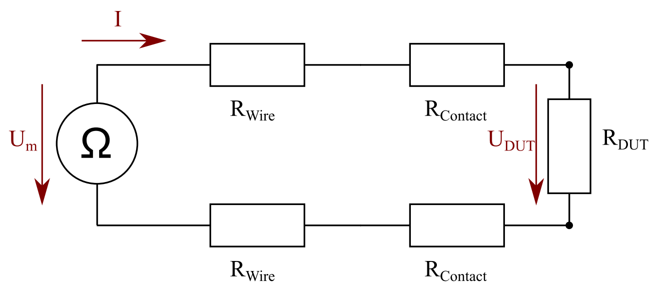

Figure G.4: Resistance determination using an ohmmeter, which uses voltage U and injected current I to display the resistance. Contact resistances are shown as one symbol in the schematic. $U_{\text{DUT}} \neq U_{\text{m}}$

This circuit loop includes the resistance of the wires (R_{Wire}) and all contact resistances which connect the ohmmeter to the device under test (R_{DUT}). With standard objects ($R_{\text{DUT}} >> R_{\text{Wire}} + R_{\text{Contact}}$) this does not cause a problem, as the additional resistances caused by the measurement arrangement are considerably small, but if the resistance of the device under test is tiny compared to the resistance of the measurement system, the error becomes substantial.

The standard method of measuring such small resistances is to use an amperemeter and a volt-meter separately connected to the device under test, as shown in Figure G.5 [161]. Additionally, a voltage- or current-source is applied in series connection to the ampere-meter. Ready-made devices combining all three functions are available, but the necessary operations and measurements can also be conducted on devices which are available in all electronics laboratories (power supply, multimeter).

Measuring low resistances accurately is a daily laboratory problem when working with batteries. To learn this method can thus be beneficial for students in their future careers.

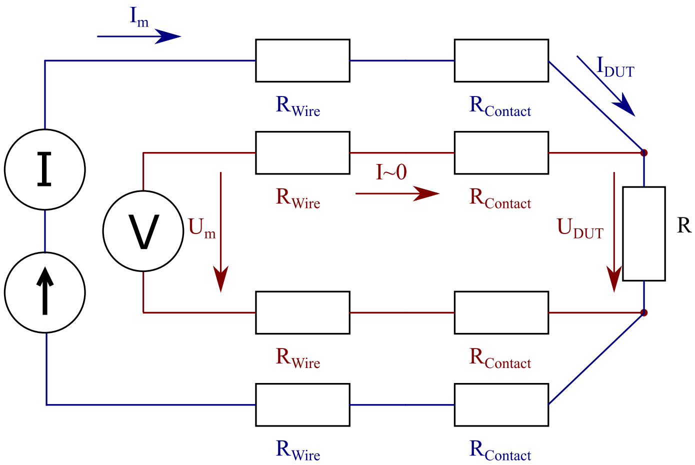

Figure G.5: Four-terminal sensing method using a volt-meter and an ampere-meter separately. $U_{\mathrm{DUT}} = U_{\mathrm{m}}$; $I_{\mathrm{DUT}} = I_{\mathrm{m}}$

As the outer circuit (blue in Figure G.5) is a series loop, the current is identical in all components and can thus be measured very accurately.

The volt-meter measures the voltage drop at the device under test separately connected and thus ignores the voltage drops of the connection resistances in the outer loop. Due to the usage of a high-ohmic or potentiometric volt-meter, nearly no current flows in that inner loop. Thus, the voltage drops on these inner resistances are nearly zero, allowing for very accurate voltage determination of U_{DUT}. The wire- and contact-resistances in the red (inner) loop in Figure G.5 can be ignored.

Using this arrangement, the calculated resistance is indicative for the device under test's resistance on its own ($R_{\mathrm{DUT}} = U_{\mathrm{DUT}}/I_{\mathrm{DUT}}$) [162, Section 4.3].

Resistance of fuses

Not all conductors in battery systems have resistances which are nearly independent of the current the object carries (see Equation G.1) [163]. Bus bars can be seen as nearly constant conductors, as their conductivity does not vary much with temperature (Equation G.1, $\alpha_{\mathrm{Copper}} \approx 0.004\,1/\mathrm{K}$) and in a well-designed system the temperature of these metal parts does not change a lot. In contrast, the behaviour of fuses is based on temperature and conductivity changes.

$$R = R_{\mathrm{ref}} \cdot (1 + \alpha(T - T_{\mathrm{ref}})) \tag{G.1}$$

The fuse is placed in series with the load it should protect, it is designed to be the weakest part in the circuit. Thus, it carries the same current as the whole circuit. The fuse should break the circuit (“burn”/ “blow”) if a fault causes too much current flow. In a real use case, this protects all components and the wiring of the system. The fuse contains a metal conductor which melts quickly. Current leads to losses at resistances and thus heats the material. The metallic conductor is sized in a way that it melts when the power dissipation is higher than the dissipation ($P = I^2 \cdot R_{\mathrm{fuse}}$)

caused by the rated current.

The fuse needs to be designed in a way that it breaks rapidly during over-currents (or withstands them for a certain time, depending on the amount of current). Thus, fuse designers have to precisely determine shape, material, and diameter to make the rated current of the fuse well defined and reliable. Even the support and fuse holder influences a fuse's heat management [164].

According to metal behaviour, the material resistance increases with its temperature, causing more losses. This mechanism defines the fuse characteristics more "sharply". Thus, fuses are usually made of materials with higher temperature coefficients or out of a special arrangement of two materials, which mix and reduce the melting point in case of a failure to define the blow characteristics more accurately (M-effect) [165, 166], while in normal operation, the fuse simply passes the smaller currents (to its rated current) with nearly no resistance. [167]

This demonstrates an essential aspect: in metrology, the device under test is sometimes influenced by the test system during the measurement. In the case of fuses (and other current dependent resistances), the relevant resistance at a specific current can be measured only when applying the current of interest.

Implementation of the four-terminal sensing method in Kelvin arrangement with standard tools

When working as an electrical engineer, the equipment required for a certain job (like ready-made four-terminal sensing devices) may not always be available. Nevertheless, four-terminal arrangements are possible even with inexpensive and widely available standard tools. Combining a multimeter in voltage mode with a power supply in constant current mode delivers (depending on the quality of both tools) acceptable results. The arrangement shown in Figure G.5 can be reproduced as in Figure G.6.

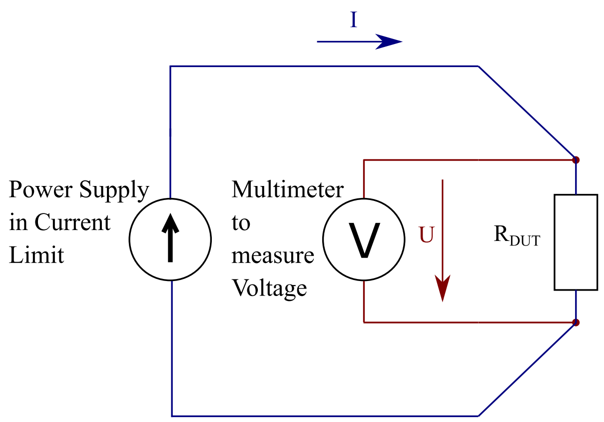

Figure G.6: Four-terminal sensing method using a power-supply in constant current mode and a multimeter determining the voltage

To allow for easy measurements, measurement tips which already contact twice are available on the market (e.g. Figure G.9). Two of the double tip probes are needed to create a Kelvin arrangement.

Calculation of measurement uncertainty

Usually, the injection/measurement uncertainty of these devices is described in the data sheets by a combination of an offset and a gain error. The final measurement uncertainty of the arrangement can be derived by:

$$\begin{aligned}
R &= {}^{U}\!/\!_{I} \\
\Delta U &= |U| \cdot E_{\text{Ugain}} + E_{\text{Uoffset}} \\
\Delta I &= |I| \cdot E_{\text{Igain}} + E_{\text{Ioffset}} \\
\Delta R &= \Delta U \cdot \left|\frac{\delta U/I}{\delta U}\right| + \Delta I \cdot \left|\frac{\delta U/I}{\delta I}\right| \\
&= \Delta U \cdot |1/I| + \Delta I \cdot |-U/I^2|
\end{aligned} \tag{G.2}$$

where

U	= measured voltage drop at DUT (V)
E_{Ugain}	= gain error of voltage measurement (V/V)
E_{Igain}	= gain error of current injection/measurement (A/A)
E_{Uoffset}	= offset error of voltage measurement (V)
E_{Ioffset}	= offset error of current injection/measurement (A)
ΔU	= measurement uncertainty of the voltage drop at DUT
ΔI	= uncertainty of the injected current
ΔR	= measurement uncertainty of the DUT's resistance

In case of positive currents, i.e. currents which produce positive voltage responses according to Equation G.2, the measurement uncertainty can be derived as in Equation G.3.

$$\Delta R = U/I \cdot \left(E_{\text{Ugain}} + E_{\text{Igain}} + E_{\text{Ioffset}}/I\right) + E_{\text{Uoffset}}/I \tag{G.3}$$

Thus, besides the aspect of influencing the value of the DUT itself, it makes sense to choose currents dependant on the measurement setup. As shown in Equation G.3, e.g. tiny test currents I cause high measurement inaccuracy, because of offsets in the devices.

G.5.1.2 A2: Contact resistances in battery systems

Every resistance in the main current path decreases efficiency of the battery system (see section G.9). When connecting cables to bus bars, fuses, or relays, contact between metal parts can be of different quality, which may lead to high losses.

Every material surface has a certain roughness – at least in the microscopic scale. Thus, real contact can occur only on the laces/summits of two metal contacting surfaces. The current passes only through these parts, which are a small portion of the nominally contacting surfaces. Contact resistance is highly random. Besides roughness, it depends on the hardness of the materials and the applied pressure. Most crucial are the residing oxides and contaminants at the surfaces [168, 169]. In 1967, Holm modelled the limiting behaviour by implementing two semi-infinite cylinders of radius b which are arranged together [170]. In his model, current can pass through a circular disk "a-spot" with zero thickness of radius $a << b$ [171].

G.5.1.3 A3: Insulation resistance / flash-over voltage

High voltages are utilised for a variety of applications, e.g. transmission of electric power over long distances.

In a battery system, the energy content depends on the capacity and number of employed cells. Increasing the voltage by placing more cells in a serial connection (and thus, having less cells in parallel), increases the efficiency of the battery system and power train. The same power can be distributed using lower currents by a thinner harness and pin contacts (power modules and connectors) [172, p. 20].

The trend to increase the voltage in high power, efficient automotive power trains is limited more by the semiconductors, i.e. the power electronics, than by the battery [173, 174].

While battery voltages of up to 400 V are currently used in electro-mobility (Table G.1), the German Electrical and Electronic Manufacturers' Association plans voltages of up to 1500 V_{DC}, especially for high performance cars and busses [172]. Future developments will increase the voltage even further [175].

Table G.1: Voltage ranges of electric cars

Car	Voltage Range / nom. Voltage	Source
BMW i3	259 V – 403 V	[176, p. 1]
BMW i8	269 V – 394 V	[177, p. 1]
Smart fortwo electric drive	230 V	[178, p. 47]
Tesla Model X	300 V – 350 V	[179, p. 193]
Tesla Model S	300 V – 350 V	[180, p. 180]
Tesla Model 3	300 V – 350 V	[181, p. 163]
Volkswagen e-UP	323 V	[182, p. 8]
NextEV Nio EP9	777 V	[183, p. 1]
Porsche Mission E concept study	>800V	[175]

The environment in which a high voltage device is used, influences the combination and arrangement of conductors and insulators. The insulating media used in electric drive trains are solids and gases (air), while conductors generally consist of a suitable conducting material such as aluminium or copper.

The voltage gradient (the electric field intensity) produces electric stress on the insulating material. [184]

$$E = -\nabla \cdot \phi \tag{G.4}$$

where

E = electric field intensity (V/m)
∇ = nabla or del operator
ϕ = applied voltage /potential (V)

Behaviour of insulation material

The dielectric strength of an insulator is defined as the voltage at which the current rises to high values when not constrained by the external resistance of the circuit. Pressure, temperature, humidity, nature of applied voltage, imperfections of the material, and surface conditions of electrodes influence the electric breakdown strength of insulators. [184]

Insulation failures are in most cases caused by discharges within the voids in the insulation or over the surface of the insulation. The probability of failure is reduced if such discharges can be excluded at the normal voltage/working conditions. Nevertheless, errors may still happen as a result of thermal or electrochemical deterioration of the insulation material [184].

In battery systems, isolation monitoring devices are integrated, avoiding dangers in case of failure [185].

To prove the effectiveness of insulation, different tests are performed with dielectrics or insulators:

- Dielectric Breakdown tests: The voltage is increased until the insulation breaks down and the material begins to conduct. This voltage is called the break-down voltage.

- Insulation Resistance tests: Determines the material's resistance by measuring the current under a given high voltage. Ohm's Law is applied. Insulators typically reach resistance values of millions of ohms.

Air gap isolation

In battery systems, next to solids such as polymers as insulators, air at atmospheric pressure is employed for isolation. The breakdown of air is of practical importance to the designers of electric drive trains.

Collisional ionisation is the dominant process which causes the breakdown in gases. Free electrons get multiplied exponentially, and, when the applied electric field intensity is sufficient, a breakdown occurs. [184, p. 5ff]

The discharges are either non-sustaining discharges or self-sustaining types. In case the voltage is low, the gas insulation retains its properties [184, p. 26]. At normal pressures and temperatures, air is an excellent insulator (air conduction in low field 1×10^{-16} A/cm² to 1×10^{-17} A/cm² [186, p. 294]). The conductivity is mainly caused by radioactive substances and cosmic radiation [186, p. 294].

In case the voltage is high, an electrical breakdown occurs, and the current increases distinctly. The gaseous spark breakdown marks the change of a non-sustaining discharge into a self-sustaining discharge. The conducting spark which appeared during the breakdown produces a short circuit linking both electrodes. The highest voltage applied to the insulation before the breakdown occurs is defined as the breakdown voltage [184, p. 26].

Air always contains some free electrons and ions. These charge carriers are accelerated if an electric field is applied. The electrons react much faster due to their lower mass. However, an avalanche effect can be created if the acceleration (energy) is high enough for an accelerated electron to ionise an atom and for the new free electrons to ionise other atoms. At this point, the breakdown voltage is reached. It depends on air density, pressure and temperature (and the type of gas between the wires).

Ionisation, in which neutral atoms and molecules alter to ions and electrons by inelastic collision processes, leads to high currents. Between those strikes, ions and electrons accelerate driven by the electric field. They gain energy between collisions and lose energy during collisions [184, p. 26].

Electrons lose little energy in elastic collisions while collecting kinetic energy from the applied field. During inelastic collisions, a significant amount of this energy is converted into potential energy, causing the ionisation of the hit molecule. The most crucial process for the breakdown of gases in strong fields is the ionisation by electron impact. The strength of the result depends on the kinetic energy that an electron can obtain on the free path between two collisions in the direction of the electric field. [186, p. 294]

Ionisation by collision is a probability phenomenon. There is an optimal electron energy range for each gas, which gives the maximum ionisation probability. [186, p. 295]

Availability of initial conduction particles, electric field configuration, pressure, temperature, and electrode surfaces are known to influence the ionisation process [184, p. 26].

Uniform and non-uniform electric fields

In a uniform field gap, the electric field E is the same everywhere in the field region. In case of a non-uniform field gap, E is dependent on the location. Approximately uniform fields exist between two infinite parallel plates. Electric fields between parallel plates of finite size can be considered uniform when the plate sizes are much larger than the gap distance. Generally, in non-uniform conditions, the field E is minimal at conductors' curvature of large radiuses, while being maximal at small radiuses. [184, p. 7]

In a given arrangement, the peak electric field value $E(x,y,z)$ is of interest. The mean electric field between two conductors is

$$E = {}^{U}/_{l} \tag{G.5}$$

where

E = mean electric field intensity (V/m)
U = potential difference between conductors (V)
l = distance between conductors (m)

[184, p. 7].

In non-uniform fields, the peak field value is above that mean value [184, p. 7f].

Ideally, in design, each part of the insulators is uniformly stressed with the value it safely withstands. Due to practical limitations of construction, this ideal configuration is impossible to achieve in engineering [184].

The exact calculation of dielectric strength provides information on the ratios of maximum local voltage gradients to the mean value in the regions of almost uniform stress (stress concentration factors). According to Naidu [184], in typical power apparatus, design factors ranging from two to five may be used. In the case of high factors, designs can be optimised in three ways: by reshaping the conductors to reduce field concentrations, by using insulation material which withstands higher fields at the stress points, and by choosing materials of proper permittivities to achieve more consistent voltage gradients [184].

G.5.2 Derived learning objectives

Based on the module targets, the dedicated theoretical lectures, and the aforementioned theory, the learning objectives for the experiment were selected. It is important to mention that these learning objectives were clearly mentioned at the beginning of the laboratory instructions, to guide the students in building the right skill set, following Sanders *et al.*'s proposition [78].

G.5.2.1 A1: Accurate measurements of low ohm resistances

- The student has experience with measurements of low resistance. The students are aware that a standard multimeter is the wrong tool for measuring very low resistance values. Accuracy is limited by the measurement arrangement.

- The student understands the principle of a Kelvin-Arrangement (four-terminal sensing method) for resistance determination.
  Addressed by task: Qu-A-5

- The student can determine low resistance values with easily available laboratory equipment (e.g. power supply and multimeter).
  Addressed by task: Qu-A-7, 8

- The student is aware of current dependent resistances (e.g. in fuses).
  Addressed by task: Qu-A-4

- The student is able to use AC methods and professional equipment to take measurements of low resistances.
  Addressed by task: Qu-A-8

G.5.2.2 A2: Contact resistances in battery systems

- The student understands the term "contact resistance" and knows typical values of contact resistances for specific electrical connections.

- The student takes contact resistances into account when designing a high power battery system.
  Addressed by task: Qu-A-2

- The student understands the difficulties in the set-up of a screwed cable lug connection in various situations.
  Addressed by tasks: Qu-A-1, 3, 10

- The student understands the negative effects of corrosion on contact resistances.
  Addressed by task: Qu-A-1

- The student understands that not all measurements can be performed directly and that some may have to be measured indirectly.

G.5.2.3 A3: Insulation resistance / flash-over voltage

- The student is able to handle the appropriate measuring equipment to determine insulation resistance and breakdown voltage.
  Addressed by task: Qu-A-9

- The student can estimate the maximum voltage between two wires isolated by air before a breakdown voltage occurs.

- The student recognises that, in some cases, the resistance depends on the applied voltage.

- The student is aware of the different parameters which determine the flash-over voltage.
  Addressed by task: Qu-A-6

G.5.3 Derived lesson

G.5.3.1 A1: Accurate measurements of low ohm resistances

In this experiment, the students measure resistances, which are very low. They use different measuring methods.

Demonstrating the disadvantages of the two-wire method

First, the students experience the disadvantages of using a two-wire method with a multimeter in the ohm measurement range. The students receive a standard 20 mm $\times$ 5 mm fuse and are requested to determine its resistance. All groups receive random values between 250 mA and 4 A. The students are requested to determine the resistance value and the accuracy of their measurement. Usually, the students perform a measurement as shown in Figure G.7. They receive a value which is very low, approximately one or two digits of the measurement device, and determine from the data

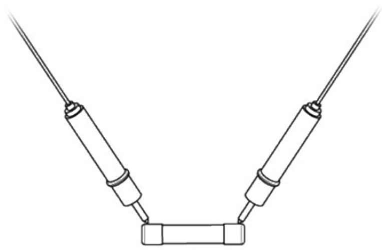

Figure G.7: Lesson A1: Two wire measurement (produces big measurement error due to the resistances in series connection)

sheet a tolerance range bigger than the value obtained. This low-quality result is discussed in the individual groups.

Still, a student could misunderstand the problem, and assume that the use of a more accurate multimeter with higher resolution would solve it. Thus, the students are now asked which additional type of error occurs with that measurement. They are requested to make a sketch, including device, wiring resistances and contact resistances, to explain the error. In the best case, the sketch should be similar to Figure G.4, showing students' understanding that it impossible to accurately determine small resistances with this method.

Performing the measurement with four-terminal sensing

The students are asked to limit the current of a power supply to 90% of the fuse's nominal value. They are requested to measure the voltage drop across the fuse with a multimeter (Figure G.8) and calculate the resistance by $R = {}^{R}/_{I}$, including error calculation (see Equation G.3). The students should now compare the quality of the results with the first method.

The measurement is also repeated for 75%, 50%, and 5% of the nominal value. They are requested to reflect on the different results on the same fuse. Where are the highest measurement tolerances? 5% of nominal current was included to show that the measurement error significantly increases with smaller currents. Aiming to answer the question "Was the resistance dependent on the current?", measurements with 90%, 75%, and 50% of nominal current were performed. The results should demonstrate the current dependency of the resistance of a fuse, and that in most cases, using this measurement method yields values with tolerances which do not overlap.

Exchanging the values to show the basic trend of fuses

All groups exchange their 50% nominal current values, and all working teams draw a diagram showing the relation between nominal value and resistance. Usually it becomes clear that fuses made for carrying higher current have smaller resistance.

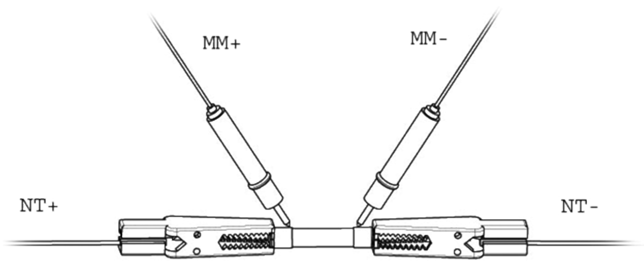

Figure G.8: Lesson A1: Kelvin or four-terminal method to determine a resistance; MM = multi-meter, NT = power supply in constant current mode

Professional equipment: alternating current measurements with HIOKI 3554

The students repeated a measurement with a battery tester, which determines the absolute value of impedance at 1000 Hz. They compared the result with the first method and got almost the same value, as the fuse behaves nearly fully resistive. The used device employs special probes, similar to the ones shown in Figure G.9 and calculates the result automatically. The student were requested to determine the applied measurement conditions of the device from the data sheet (frequency and current).

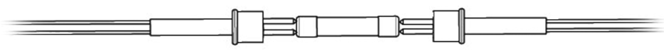

Figure G.9: Lesson A1: Four-terminal method with special double tip probes

G.5.3.2 A2: Contact resistances in battery systems

The students measured the contact resistance from a cable lug to a copper bus bar and compared the values. The students were given a copper bus bar with one side polished and one heavily oxidised. A table with torque values and screw diameters was provided.

1. Unpolished side: M5 and M8 screw (and fitting cable lug, with different contact area), each with (shown in Figure G.11)
    (a) Steel screw without washer
    (b) Steel screw with washer between screw and cable lug (cable connector)
    (c) Steel screw with washer between copper and cable lug
2. Oxides dependency: Polished side: same measurements
3. Screw material dependency: M8 copper screw with lower torque, compared to the steel screw used before with the same reduced torque
4. Torque dependency: M10 screw, 50% and 25% of the original torque.

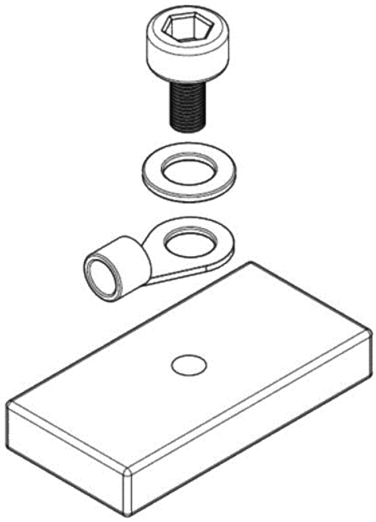

Figure G.10: Lesson A2: Different metal parts for cable lug connections used in the experiment

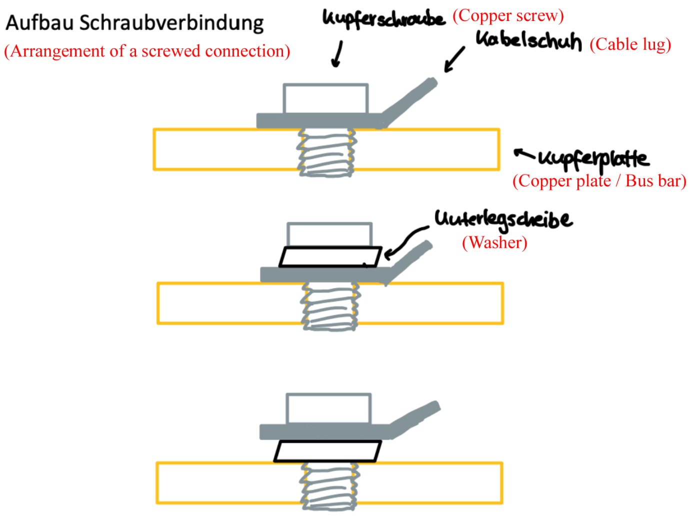

Figure G.11: Lesson A2: Student's protocol. Different styles of a cable lug connection. The best way is to place the washer between screw and cable lug.

The students found out that the washer needs to be between the screw and cable lug. Here, it distributes force and efficiently increases contact area, without adding a second contact resistance. Besides, the washer is usually made from steel, a poor conductive in comparison to copper.

Following that, the students were asked to draw diagrams on the rest of the above-mentioned dependencies. They find out that this type of connection depends on many factors and is also highly random. Trends which were usually observed are: High contact area, high pressure/torque and less oxides improve the connection.

The students were requested to further investigate the main current path in such an arrangement (outcome: nearly all current is going directly from cable lug to bus bar, nearly no current is using the screw as conductor, as shown in Figure G.12). The students understand that it makes no sense to replace the steel screw by other materials. It is a division of purpose in such a connection. The steel screw creates the pressure, while the copper cable lug and bus bar carry the current.

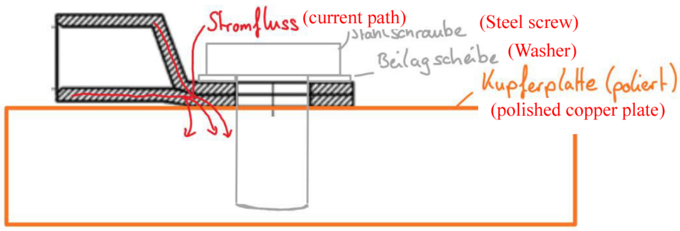

Figure G.12: Lesson A2: Student's protocol. The students correctly found the main current path of a cable lug connection.

Indirect Measurements

Sliding contacts: A short piece of a slot car track and a car, where both contacts on the motor have been bridged (short-circuited, assumption 0 Ohm), was handed over to the students. It is shown in Figure G.13. The students are asked to determine the contact resistance of the sliding contacts of the vehicle using the sense lines.

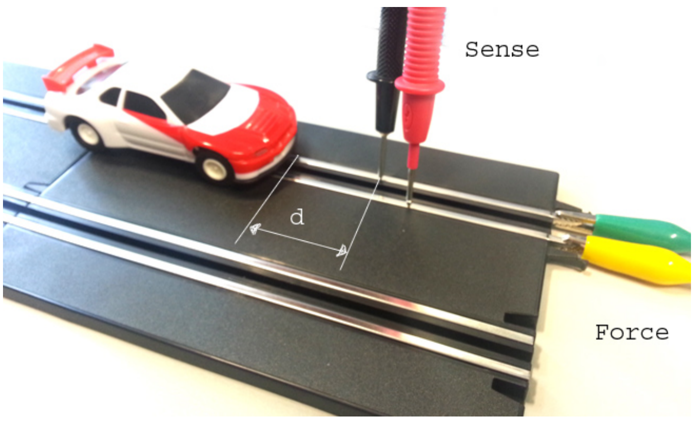

Figure G.13: Lesson A2: Slot car track

The students were allowed to select which devices they use for the measurement,

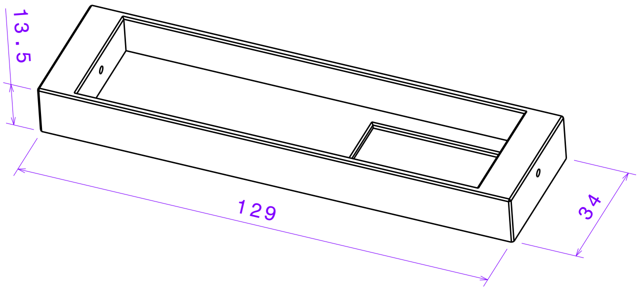

Figure G.14: Lesson A3: Measurement set-up for flash-over-voltage and dielectric breakdown experiment, all dimensions in mm

the power-supply with multimeter or the automatic AC four-wire measurement with the Hioki HiTESTER 3554. The contact resistance of the brushes on the underside of a car cannot be measured directly, as they cannot be reached with the measuring probes. The instructions included a hint to help the students solve this problem: “Use several measurements and a small calculation to solve the problem. The lengthwise resistance of the race track can be assumed to be constant.” The most natural solution is to measure at two different distances in front of the car and extrapolate the contact resistance (distance = 0 cm).

G.5.3.3 A3: Insulation resistance / flash-over voltage

First, in this experiment, the student learn the relationship between breakdown voltage and distance between conductors.

Secondly, the participants investigate the isolation capability of two-layered standard paper at high voltage. While running this experiment, the participants calculated the insulation resistance and noticed that this insulation resistance is not constant but depends on the applied voltage.

The experiments were conducted using a “Schleich Motortester MA1”, which was employed as a high-voltage generator. At the beginning of the instructions, the participants are instructed not to touch the set-up, as voltages of up to 5200 V may be applied.

The safety of the participants is steadily guaranteed by the MA1 device, which limits the current to a maximum of 1 mA. The students receive instructions on how to adjust the voltage and read the resulting current.

Flash-over voltage

The students received a 3D-printed arrangement (see Figure G.14, and Figure G.15), which allowed for adjusting the distance between two parallel copper bus bars by plastic screws (1 turn $\widehat{=}$0.5 mm) accurately. They directly configured the voltage created by the MA1 through a potentiometer on the device.

The students were instructed to slowly increase the voltage until the flash-over occurred (ionisation, limited to 1 mA). This measurement was repeated for 0.5 mm,

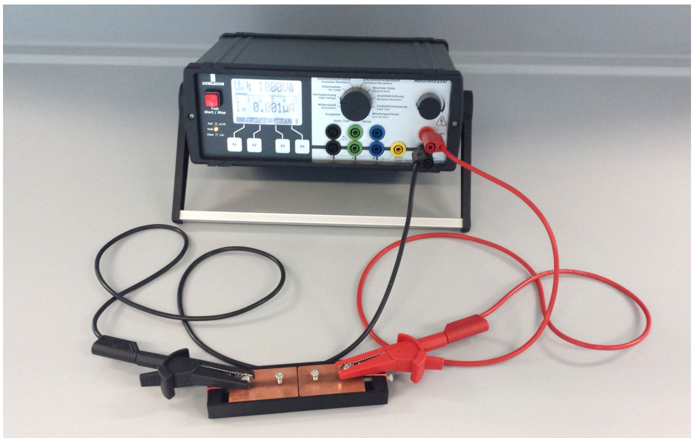

Figure G.15: Lesson A3: Set-up for flash-over voltage experiment

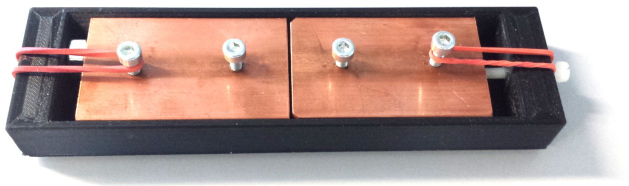

Figure G.16: Lesson A3: Set-up for flash-over voltage experiment

1.0 mm,1.5 mm, and 2.0 mm. The participants were asked to create a table with the measurement results.

Furthermore, they were requested to reflect on the current measured while the flash-over occurred, and the possibility of determining a single resistance, to avoid misunderstanding of the physical behaviour.

Insulation resistance and dielectric breakdown

For this experiment, participants used the same arrangement and devices as for the previous one. They were requested to put two layers of standard paper between the copper bars, and minimally adjust the distance without exerting pressure on the paper (see Figure G.17).

Then, they took current measurements for the following voltages: 250 V, 500 V, 750 V, 1000 V, 1250 V, 1500 V, 1600 V, 1700 V, and 1800 V. The results were plotted in a diagram (similar to Figure G.18).

The students were asked to determine the insulation resistance and the dielectric breakdown voltage. Furthermore, they were instructed to examine the paper, which exhibited charred parts caused by the breakdown.

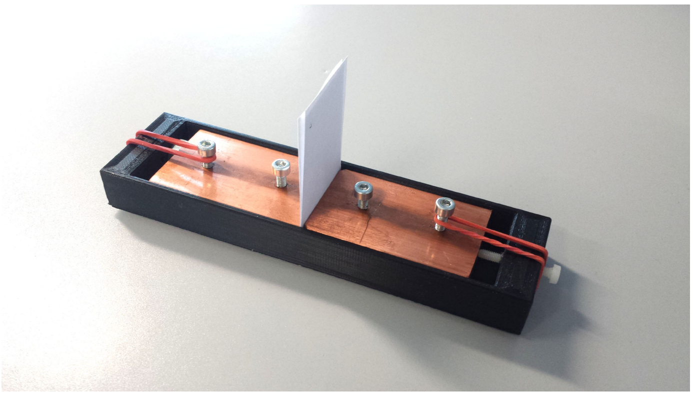

Figure G.17: Lesson A3: Set-up for insulation resistance and dielectric breakdown experiment

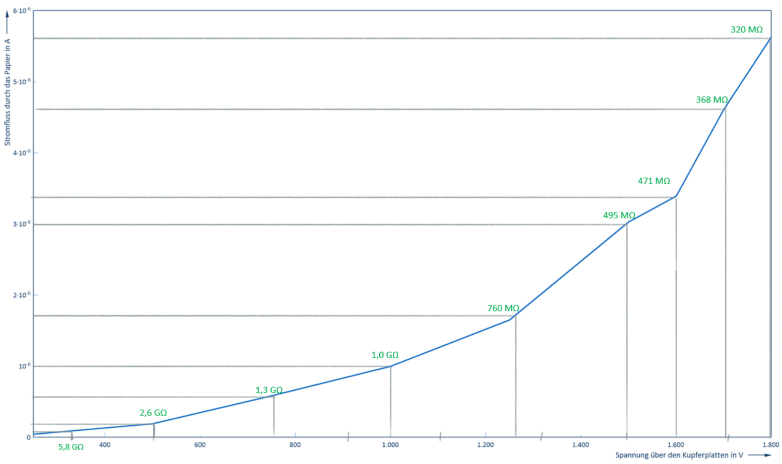

Figure G.18: Lesson A3: Student’s diagram: insulation resistance below breakdown voltage

G.6 Lesson B – Open-circuit voltage curve (OCV)

G.6.1 Theoretical technical background

Open-circuit voltage

The open-circuit voltage is the voltage between both poles of a battery where no current flows (also called the no-load voltage) and the cells are in the equilibrium state.

State of Charge (SoC)

The state of charge is defined as the actual charge (As) of a battery relative to the capacity (As) and stated as a percentage. It is 0% for an empty cell at a minimum voltage, and 100% for a full cell at a maximum voltage.

Determining SoC is usually performed by the integration of current over the time, based on a known start SoC (see Equation G.6).

$$SoC(T) = SoC(t=0) + \frac{\int_{t=0}^{T} I(t)dt}{C_{\text{cell}}} \tag{G.6}$$

where

$I(t)$ = current (A)
C_{cell} = capacity of the cell (As)
t = time (s)

Current measurement errors contribute to errors when calculating the SoC: First, gain errors of current measurement contribute directly to the error of SoC with the same factor. Secondly, offset errors are integrated and contribute depending on the measurement time. This applies in battery management and also in the laboratory experiments.

C-rate

The C-rate (M) defines a current relative to the capacity of a cell. 1 C is equal to a full CC discharge in 1 h. When, for example, doubling the current to 2 C, the discharge is completed after 30 min (Equation G.7).

$$\begin{aligned} M &= {}^{I}\!/_{C} \\ t_{\text{CC,discharge}} &\approx {}^{1}\!/_{M} \end{aligned} \tag{G.7}$$

where

M = C-rate ($^{1}/_{\text{h}}$)
I = current (A)
C = capacity (A h)
$t_{\text{CC,discharge}}$ = discharge time (h)

Usage of C-rates is handy when comparing cells of different capacity. For example, a 1 C load for determining the actual capacity of cells helps to formulate the standard test procedure, which can be used with cells of different size. The test procedure exposes all cells with the same relative load, and all tests last similarly long, independent of the capacity of the device under test.

Voltage (change) at accumulator cells

Note: For the German runs, the reasons for the potential change at the battery clamps were already taught in detail in the joint theoretical lecture and partly in the chemistry lecture one semester before. Thus, the experiments focused on the methods to determine the voltage curves.

The clamp voltage depends on the chemical potentials of the anode and cathode (see Equation G.8 and Figure G.19).

$$U = \Delta E = E_2 - E_1 \tag{G.8}$$

where

U = clamp voltage (V)
E_1 = anode potential (V) (Electrode with more negative std. electr. potential)
E_2 = cathode potential (V) (Electrode with more positive std. electr. potential)

Assuming standard conditions of the employed electrochemical voltage series (e.g. 298.15 K, c of electrolytes 1 $^{\mathrm{mol}}/_{\mathrm{L}}$), one can derive the values for the potentials of the individual electrodes (for metals: redox pairs) from this electrochemical voltage series.

In all other cases, both half cell potentials need to be calculated using the Nernst-Equation G.9.

$$E = E_0 + \frac{R \cdot T}{z_e \cdot F} \cdot \ln \frac{a_{\mathrm{Ox.}}}{a_{\mathrm{Red.}}} \tag{G.9}$$

where

E = half-cell potential (V)
E_0 = standard half-cell potential (V)
R = ideal gas constant ($^{\mathrm{J}}/_{\mathrm{K\,mol}}$) $\approx 8.314{,}47\ ^{\mathrm{J}}/_{\mathrm{K\,mol}}$
T = temperature (K)
z_e = number of transferred electrons (1)
F = Faraday constant ($^{\mathrm{C}}/_{\mathrm{mol}}$) $\approx 96{,}485.332{,}12\ ^{\mathrm{C}}/_{\mathrm{mol}}$
$a_{\mathrm{Ox.}}$ = activity oxidised molecules (1) *
$a_{\mathrm{Red.}}$ = activity reduced molecules (1) *

* activities are often replaced by concentrations

During charging/discharging, electrons travel around the cell through the sink (e.g. a resistor, bulb, or motor) or source (e.g. a charger or a solar panel) of electrical energy, while the lithium-ions travel inside the cell (see Figure G.20). As concentration changes while charging/discharging the accumulator, so does the clamp voltage (see Figure G.19).

The cells voltage range is limited, as leaving the allowed voltage range would allow electrical potentials (or concentrations) to occur which create other educts by irreversible processes (see Figure G.21).

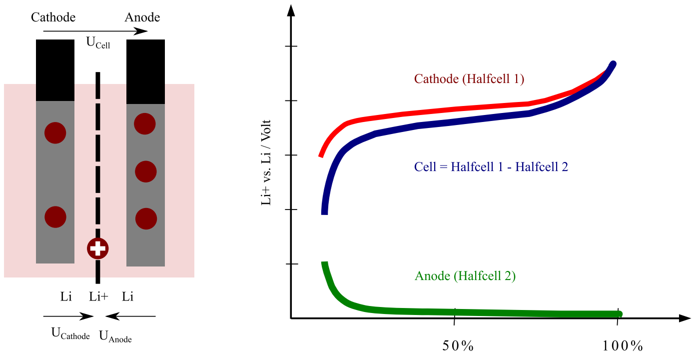

Figure G.19: Lesson B: Clamp voltage depends on the potentials of both half cells.

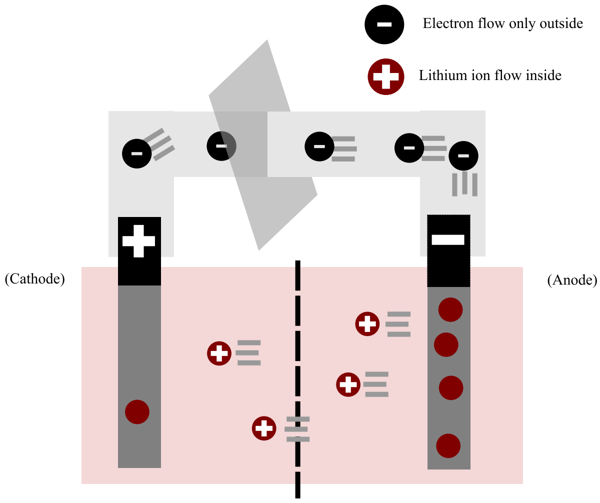

Figure G.20: Lesson B: Internal movements while discharging a lithium-ion cell

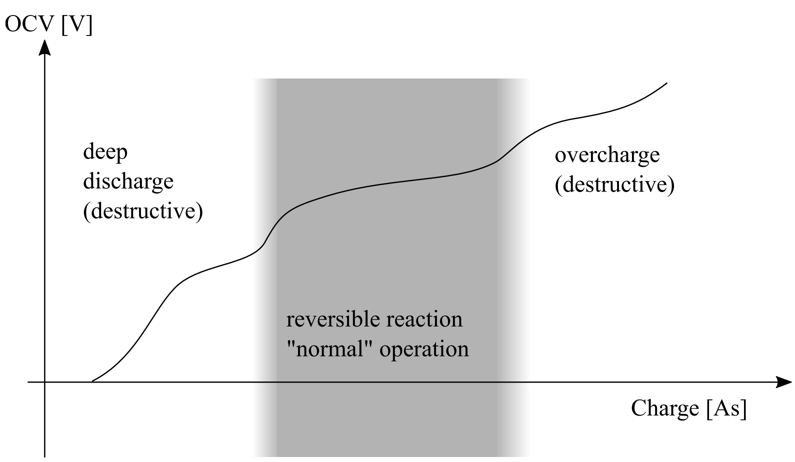

Figure G.21: Lesson B: Range of reversible reactions (accumulator)

Methods to determine the voltage/open-circuit voltage of cells dependent on SoC

The open-circuit voltage is defined as the voltage at cell clamps when no load is attached, in the equilibrium state of the cell. Thus all internal chemical and physical processes need to be completed and in a stable state. With battery cells, this may take several days, depending on the requested accuracy. This means a high effort is required to determine that information for new cell types.

Purpose of the open-circuit voltage
The open-circuit voltage is often the central point for battery models: in case of battery simulation, to design or calculate battery systems in the development phase. In battery models employed in battery management systems, for example to support SoC calculation. Inaccuracies in this information about the cell cause considerable problems in these algorithms.

Differences between open-circuit voltage and clamp voltage
As mentioned above, the open-circuit voltage is difficult to determine – if time is limited. To determine the open circuit voltage (OCV) dependent on the SoC for the above-mentioned purposes, it would be necessary to charge the cell to a known SoC, e.g., 100%. Then one has to wait long enough, record the voltage, change the SoC, wait, record the voltage, and so forth. In the case of 24h waiting time to reach equilibrium, and 5% SoC steps, the full measurement would take more than 20 days, which is far too much for the experiments in the teaching labs.

Thus, other methods need to be considered, which are also used in practical laboratory work to allow to finish the experiments in the given time. A steady discharge was chosen to determine a diagram similar to the OCV(SoC) diagram. Besides the cell not being in equilibrium, the big disadvantage here is that the recorded voltage is not determined in open circuit condition. Current is permanently flowing – and considering the internal resistance (see Figure G.22) at discharges, a lower voltage is recorded. To avoid misunderstanding and confusion about the differences between clamp voltage during current flow and OCV, a charge (opposite current direction) was also performed. Herein, both curves do not fit, which makes misunderstanding impossible. The discharge voltage curve is below the open-circuit voltage curve, while the charge voltage curve is above. The OCV(SoC) curve is between the discharge and charge curve. Assuming the discharge internal resistance is approximately the same as the charge internal resistance, the voltage drops are proportional to the employed currents and the measurements can also be used to locate the open-circuit curve.

G.6.2 Derived learning objectives

Based on the module targets, the dedicated theoretical lectures, and the aforementioned theory, the learning objectives for the experiment were selected.

- The student is able to handle the appropriate measuring equipment to determine cell voltages (different current, different temperatures).

Addressed by tasks: Qu-B-1, 3, 13

- The students knows the typical behaviour of lithium-ion cells regarding voltage, dependent on the state-of-charge. They know the standard voltage ranges of lithium-ion cell types.
  Addressed by tasks: Qu-B-2, 6, 7

- The student is aware of the effects of measurement errors (e.g. current error integration). The student knows about the consequences when (voltage/current/temperature) transients apply on a device under test for the measurement quality.
  Addressed by tasks: Qu-B-5, 11

- The student is able to deal with basic battery parameters such as capacity, SoC.
  Addressed by task: Qu-B-8

- The student knows about CC-CV charging and discharging procedures with laboratory equipment and can derive parameters for battery models.
  Addressed by tasks: Qu-B-9, 10, 12, 14

G.6.3 Derived lesson

At the beginning of each lesson, tasks are distributed to the different groups. Each working team performs a discharge and charge of a lithium-ion cell. The teams are given 600 mAh $LiMn_2O_4$ lithium-ion accumulators. To show that cells may have different voltage ranges 600 mAh $LiFePO_4$ were additionally employed in some runs. The discharge and charge rates varied from .75 C to 1.5 C. In the German runs, combinations with different temperatures were also selected. The list of proposed tasks was optimised for the available time of the lesson. The smaller the selected currents, the more accurate the results. The applied currents were calculated depending on the requested C-rate (see Equation G.7) and the actual capacity, which was stated on a sticker attached to the cells.

After programming a CC-CV discharge (using the GUI) and performing it, the students were requested to analyse the data. The program was a very simple one, which made this a good fit for the first experiment to use the battery test station/simulation control GUI. To correctly program the procedure, the students needed to check the data sheets of the cells for minimum and maximum voltage. Students had to configure the thermostat for the desired temperature, and had to wait a certain time before the test.

With the derived data, the students were requested to plot two diagrams, one for voltage over time, and one for voltage over charge/SoC. These diagrams differed, as the current in CV phase varies.

In a second step, the groups exchanged the data, and plotted all combinations in the second graph. Several facts were now visible: The discharging curve did not fit the charging curve (voltage drops on the internal resistance). Thus, the OCV-curve is between both of these curves. The students were able to see the typical shape of the

OCV(SoC) behaviour, and that it does not strongly depend on temperature (below measurement error). The students were able to see the differences in voltage range and shape between iron-phosphate and manganese cells. Besides, by comparing different groups' results, students were able to recognise that the internal resistance depends on temperature.

G.7 Lesson C1 – Internal resistance

G.7.1 Theoretical technical background

A lithium-ion cell can be considered a (stable) voltage source in series with a resistance of small value (Figure G.22). This resistance is called the “internal resistance” of a cell (also called, when observing a system from outside, “equivalent series resistance”. In a cell, this value is determined by the conductivities of the electrolyte, the current collectors, and the active materials. In the case of cell stacks or packs, the internal resistance is the sum of all internal cell resistances and the resistances of the other parts (fuses, relay, contact resistances) used to connect the cells.

The internal resistance is – besides the voltage level – the essential parameter of battery cells as it determines the stability of the output voltage, efficiency, and maximum power the cell can deliver. Also, battery efficiency is higher with low internal resistance [187]. For battery systems, lower heat generation means that cooling efforts can be reduced, which brings additional benefits in system volume and weight reduction [187]. Internal resistance may also have an influence on battery life: The higher the internal resistance, the higher the power losses ($P = U_R \cdot I = I^2 \cdot R$) heating the cell, which results in shorter cell life. When the average operating temperature is raised by 10 K, a system’s service life decreases by about 50% [188].

The easiest way to model a battery cell (Figure G.22) is a combination of an ideal voltage source (to provide the open circuit voltage U_{OCV}) and a resistor (to model the internal resistance R_I, which presents the voltage drop and losses in conditions when current flows) [149]. In battery management, awareness of internal resistance is crucial, as the value influences other algorithms, such as SOH, SoC, and maximum power prediction.

The internal resistances of battery cells are difficult to determine: Firstly, the value is low (usually in the 1 mΩ range for lithium cells with more than 20 A h). For measurement of small resistances refer subsubsection G.5.2.1. Secondly, because of the live power/voltage source (the voltage source in the equivalent circuit diagram Figure G.22), which does not allow the use of Ohm’s law directly at the cell clamps.

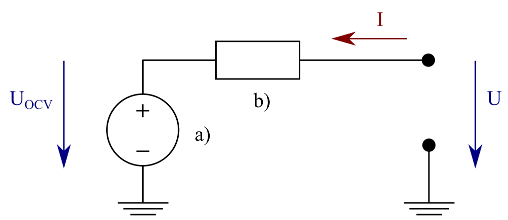

Figure G.22: Simple (static) equivalent circuit diagram of a battery cell to model the electrical characteristics of the battery cell. a) Open circuit voltage (ideal voltage source dependent on the state of charge) b) Internal resistance

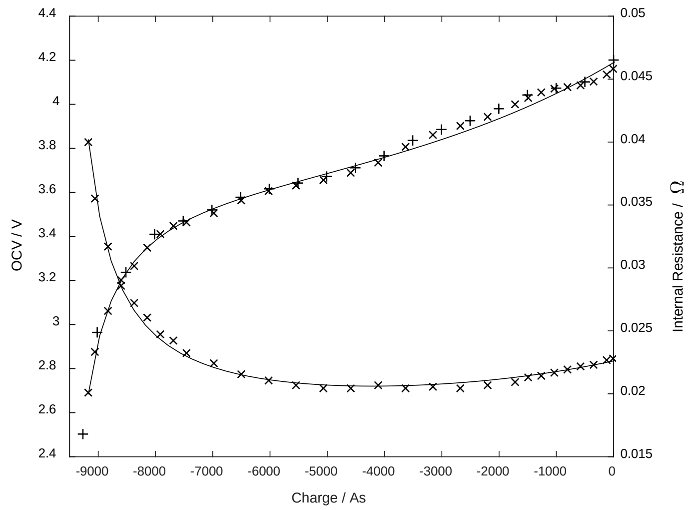

Figure G.23: Open circuit voltage and internal DC resistance at 25 °C dependent on the extracted charge, own measurements at a 2.5 Ah NMC-cell INR18650-25R from Samsung SDI [149]

G.7.1.1 Parameters influencing the internal resistance of lithium-ion cells

The influence of the state of charge

In general, the internal resistance over the SoC follows a parabolic U shape. The maximum values of the internal resistance are at minimum and maximum SoC. The highest values are reached when the cell is fully discharged. [149, 187], [189, e.g. Fig. 10, Fig 13].

The internal resistance in the middle SoC range (approx. 15% to 75% SoC) remains almost constant at low values, which are about a third of its maximum value at low SoC (see Figure G.23).

This is a consequence of kinetics and mass transport effects. The reversible process runs easier when the concentrations of products are identical [190]. Thus, the efficiency is reduced when a cell is at low or high SoC, and the cell's power capability is limited [189].

The influence of temperature

The optimal temperature range for a lithium-ion battery in operation is between 20 °C and 40 °C [188].

The internal resistance has a strong, negative temperature coefficient (higher temperatures reduce the resistance) [191]. Thus, when comparing different cells, it is most important to test all of them at identical temperature.

At temperatures below 20 °C, the internal resistance rises more strongly when the temperature decreases further [192]. This reduces cell performance and also results in lower maximum pulse power [188].

Exemplary measurement results are presented in Figure G.24.

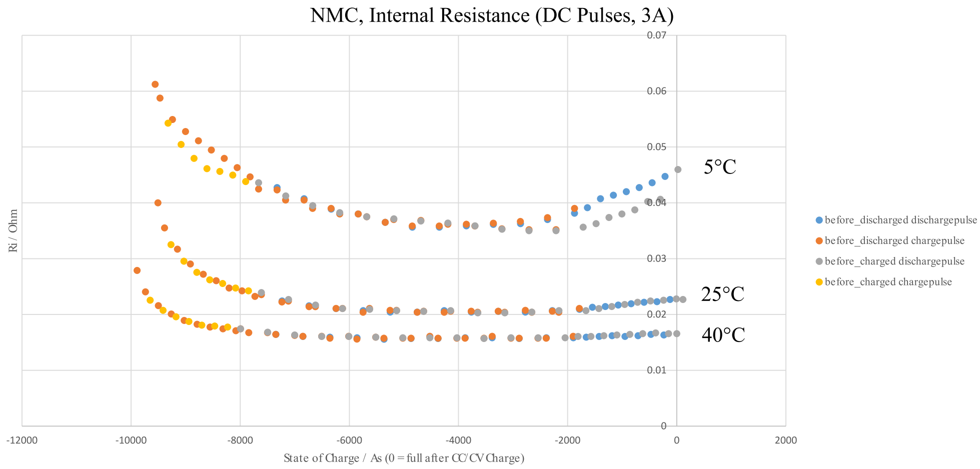

Figure G.24: Lesson C1: SoC and temperature dependency of the internal resistance, own measurements at a 2.5 Ah NMC-cell INR18650-25R from Samsung SDI [149]

The influence of ageing

Specific ageing mechanisms occur at temperatures below zero while charging that can lead to irreversible cell damage. This mechanism is called “lithium plating” – pure lithium is deposited on the anode while charging the cell. The mechanism reduces cell capacity as active material is reduced and can in the worst case cause an internal short when the plated lithium forms dendrites penetrating the separator. [188, p. 156]

The influence of capacity

Usually, the smaller the capacity, the higher the internal resistance. A cell with big capacity can be understood as many parallelly connected smaller cells. Thus, if capacity is doubled, one can expect halved internal resistance (assuming same cell types with different capacity).

G.7.1.2 The methods to determine the internal resistance

Calorimetric – losses

Internal resistance means losses, which heat the cell and the environment. To determine internal resistance it is possible to measure the generated heat using a calorimeter. By comparing the detected heat with the expected heat, assuming a constant internal resistance, the average internal resistance can be derived (Equation G.10). Besides the disadvantage that the method averages the losses of all working points during the measurement (like SoC), the method has drawbacks like the complex and expensive equipment (calorimeter). Because of exothermal and endothermal chemical reactions in a battery cell, the cell needs to be cycled symmetrically to derive accurate results. [193]

$$\bar{R_I} = \frac{Q_{\text{loss}}}{\int_{SoC_A}^{SoC_A} I(t)^2 dt} \tag{G.10}$$

where

$\bar{R}_I$ = mean internal resistance (Ω)
SoC_{A} = starting and end SoC (%)
$I(t)$ = test current profile (A) where $\int I(t)dt = 0\,\mathrm{A\,s}$

AC – impedance

AC impedance measurements are widely used to characterise batteries. Using that method, in most cases a sine-wave current with amplitude small enough to consider the battery cell as linear time-invariant (LTI) system is injected. The voltage response is recorded and the impedance (complex resistance) is computed, for example by using Fourier transformation. The impedance can express information about the absolute value of resistance, and the phase shift between peak voltage and current. When this measurement is repeated for several frequencies, it leads to an impedance spectroscopy graph of the regarding cell in the actual environmental situation. [193]

$$Z(\omega) = \frac{\hat{U} \cdot \sin(\omega \cdot t + \Delta\phi(\omega))}{\hat{I} \cdot \sin(\omega \cdot t)} \tag{G.11}$$

where

$Z(\omega)$ = impedance (Ω)
$\hat{U}$ = voltage amplitude (V)
$\hat{I}$ = current amplitude (A)
$\Delta\phi(\omega)$ = phase shift between voltage and current

A typical result of an example cell is presented in Figure F.10. The frequency dependent absolute values and phase shift of lithium-ion cell impedance is caused by capacitive and inductive effects, which can be modelled as shown in Figure G.30.

DC – current step method

Another option to determine internal resistance is the current step method. One needs to apply two different currents I_1 and I_2. The system under test's voltage is determined with both currents. By comparing both voltages using Equation G.12, the internal resistance can derived [193]. In most cases I_1 is chosen to be 0 A.

$$R_I = \frac{U_1 - U_2}{I_1 - I_2} = \frac{\Delta U}{\Delta I} \tag{G.12}$$

where

R_I = DC internal resistance (Ω)
U_1 = cell clamp voltage when applying current 1 (V)
U_2 = cell clamp voltage when applying current 2 (V)
I_1 = current 1 (A)
I_2 = current 2 (A)

As shown in Figure G.25, the time span between applying current and determining voltage is crucial when testing batteries. When currents are applied, the SoC

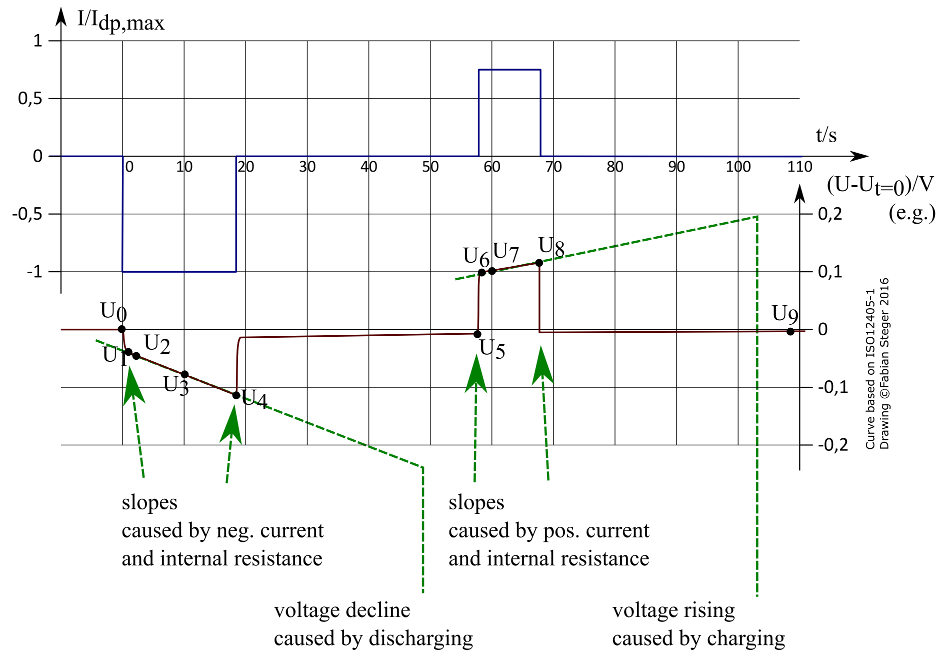

Figure G.25: Lesson C1: Current step method to determine DC internal resistance

changes, leading to a voltage drop, which is not caused by the internal resistance, but the reasons demonstrated by experiment B. The internal resistance as one value cannot describe the real system (e.g. shown in the more complex model in Figure G.30) with capacitive and inductive behaviour. Thus, the time span also influences the interpretation of the value internal resistance.

The ISO12405-1 standard "Electrically propelled road vehicles – Test specification for lithium-ion traction battery packs and systems - Part 1: High-power applications"

Figure G.25 presents the current profile proposed by the ISO 12405-1 standard "Electrically propelled road vehicles – Test specification for lithium-ion traction battery packs and systems - Part 1: High-power applications". The students receive a shortened version of the standard in English language before the lesson takes place. The current profile of the standard is useful to determine different characterising internal resistance values of energy storage systems. If repetitive measurements on one specimen are planned, the SoC changes, as the current profile according to the standard removes in the first phase more charge than it refills in the second phase. To avoid a moving SoC in a series measurement, it is recommended to increase the length of the charge pulse.

The ISO12405-1 standard defines the internal resistance based on particular pulse durations [194, p. 20]. The internal resistance is then calculated based on the voltage after a certain time span to characterise the specimen [194, p. 20] (see Table G.2).

These standard internal resistance values are determined at different SoC (80%, 65%, 50%, 35% and 20%) and at different temperatures (40 °C, T_{room}, 0 °C, -10 °C, and -18 °C). These values allow automotive engineers the comparison of different

Table G.2: Calculation of internal resistance according to the ISO12405-1 standard

Value	Equation	$\Delta t/s$
0.1 s discharge internal resistance	$Ri_{0.1\text{s,dch}} = {U_0 - U_1}/{I_{\text{dp,max}}}$	0.1
2.0 s discharge internal resistance	$Ri_{2\text{s,dch}} == {U_0 - U_2}/{I_{\text{dp,max}}}$	2
10.0 s discharge internal resistance	$Ri_{10\text{s,dch}} = {U_0 - U_3}/{I_{\text{dp,max}}}$	10
18.0 s discharge internal resistance	$Ri_{\text{dch}} = {U_5 - U_4}/{I_{\text{dp,max}}}$	18
overall discharge resistance	$Ri_{18\text{s,dch}} = {U_0 - U_4}/{I_{\text{dp,max}}}$	40
0.1 s regenerative internal resistance	$Ri_{0.1\text{s,cha}} = {U_5 - U_6}/{0.75 \cdot I_{\text{dp,max}}}$	0.1
2.0 s regenerative internal resistance	$Ri_{2\text{s,cha}} = {U_5 - U_7}/{0.75 \cdot I_{\text{dp,max}}}$	2
10.0 s regenerative internal resistance	$Ri_{10\text{s,cha}} = {U_5 - U_8}/{0.75 \cdot I_{\text{dp,max}}}$	10
overall charge resistance	$Ri_{\text{cha}} = {U_9 - U_8}/{0.75 \cdot I_{\text{dp,max}}}$	40

cell-types.

G.7.2 Derived learning objectives

Based on the module targets, the dedicated theoretical lectures, and the aforementioned theory, the learning objectives for the experiment were selected.

- The student understands the importance of the internal resistance for the overall system efficiency.
  Addressed by task: Qu-C-1

- The student knows methods to determine the internal resistance of cells. The student is able to use the DC determination method for internal resistance (discharge pulses following the ISO12405-1 standard “Electrically propelled road vehicles – Test specification for lithium-ion traction battery packs and systems - Part 1: High-power applications” [194]).
  Addressed by tasks: Qu-C-2, 13

- The student understands battery behaviour based on a simple equivalent circuit.
  Addressed by tasks: Qu-C-3, 14

- The student is able to estimate the behaviour of lithium cells regarding their internal resistance and temperature dependency. The student knows the strong dependency of internal resistance on temperature and current, and is aware of the general trends.
  Addressed by tasks: Qu-C-4, 5, 7, 15

- The student knows typical values regarding the internal resistance of lithium-ion cells.
  Addressed by task: Qu-C-6

- The student knows how to change the SoC to desired values.

G.7.3 Derived lesson

The students received a short extract of the ISO12405-1 standard “Electrically propelled road vehicles – Test specification for lithium-ion traction battery packs and systems - Part 1: High-power applications” [194]. Figure G.25 presents the current profile proposed by the ISO 12405-1 standard.

Preparation

For preparation, the students were requested to draw the simplest equivalent circuit of an accumulator cell, which describes the internal resistance effect (Figure G.22). They were then asked to check the definition of the internal resistance according to the ISO12405-1 standard [194] (and focus on the short time value of 0.1 s).

Grouping

The students perform the experiments in working teams. For this experiment the 600 mAh $LiMn_2O_4$ cells (see Table F.1) were employed. The data sheets were available to the students. To reduce the number of measurements, and depending on the number of working teams, each team was assigned an SoC (80%, 60%, 40%).

Changing SoC to the desired value

All cells were handed over in full condition (100% SoC) and accompanied by a note with the actual capacity. The first task for the students was to discharge the cell to the SoC corresponding to the assigned SoC of the group. Herein the students had some degree of freedom with regards to the discharging current. Based on the current, the discharge time was calculated (Equation G.6).

Modifying the profile of the standard and performing the first current pulse profile at room temperature

The students had to program the profile shown in Figure G.25 for a C-rate (see Equation G.7) of $M = 1\,{}^{1}\!/\mathrm{h}$. First, the students were requested to sketch the desired current profile on paper. Disadvantages were discussed, such as the overall change of SoC when performing several pulses in series. Finally, the students were requested to modify the current profile in a way that meant the charge pulse exactly recharged the removed charge of the discharge pulse (to extend the time of the second pulse). After the sketch met the expectations, students started to program the profile in the GUI. Once the program was correct, students started the profile with room temperature and recorded the data for evaluation.

Determining the DC internal resistance with different C-rates

Now, the students were asked to perform the same measurement with different C-rates ($I_{\mathrm{dp,max}}$ = 3 C, 1 C, and .5 C).

Measurement at different temperatures

Now, the students repeated the measurement at different temperatures (T = 40 °C, room temperature, 0 °C, −10 °C, and −18 °C) and 3 C. Here, the students learn to wait for an equilibrium. Since the used cells were of small physical dimensions, temperature reached the core of the cells fast. Some minutes of wait time were accepted in the laboratory. The students are told that, for fully scientifically viable results, they would have to wait longer. Because of high internal resistance, it may be possible that the safety module detects under-voltage. In this case, the students are asked to reduce the C-rate.

Plotting the diagrams

Three graphs are create by all working teams:

- Graph A: Internal resistances at room temperature as a function of the C–rate (R_I y-axis/ordinate, C-rate x-axis/abscissa).
- Graph B: Internal resistances at a constant C-rate as a function of the temperature (R_I ordinate, T abscissa).
- Graph C: Internal resistances at a constant C-rate as a function of the SoC (R_I ordinate, SoC abscissa).

For all three graphs, the working teams were asked to describe the behaviour of the graphs.

G.8 Lesson C2 – Power

G.8.1 Theoretical technical background

It is essential to have data on the available power that an energy storage system can deliver or absorb. In automotive engineering, it is crucial to be aware of the power the motor inverter can draw from the energy storage system. If the driver demands high torque when the battery is approaching empty, the system may draw too much power, which will cause an emergency switch-off of the energy storage system due to under-voltage. [195, p. 166]

Determining pulse power according to the automotive standard ISO12405-1

The ISO12405-1 standard "Electrically propelled road vehicles – Test specification for lithium-ion traction battery packs and systems - Part 1: High-power applications" defines the pulse power based on particular pulse duration. The pulse power is then calculated based on the voltage at the end of the pulse [194, p. 20] (see Table G.3). The pulse profile is identical to the profile for internal resistance determination (see Figure G.25).

Table G.3: Calculation of power according to the ISO12405-1 standard

Value	Equation	$\Delta t/s$
0.1 s discharge power	$P_{0.1\text{s,dch}} = U_1 \cdot I_{\text{dp,max}}$	0.1
2.0 s discharge power	$P_{2\text{s,dch}} = U_2 \cdot I_{\text{dp,max}}$	2
10.0 s discharge power	$P_{10\text{s,dch}} = U_3 \cdot I_{\text{dp,max}}$	10
18.0 s discharge power	$P_{18\text{s,dch}} = U_4 \cdot I_{\text{dp,max}}$	18
0.1 s regenerative power	$P_{0.1\text{s,cha}} = U_6 \cdot 0.75 \cdot I_{\text{dp,max}}$	0.1
2.0 s regenerative power	$P_{2\text{s,cha}} = U_7 \cdot 0.75 \cdot I_{\text{dp,max}}$	2
10.0 s regenerative power	$P_{10\text{s,cha}} = U_8 \cdot 0.75 \cdot I_{\text{dp,max}}$	10

These standard pulse power values are determined at different SoC (80%, 65%, 50%, 35% and 20%) and at different temperatures (40 °C, T_{room}, 0 °C, -10 °C, and -18 °C). These values allow automotive engineers the comparison of different cell-types.

Cells charge and discharge power is limited

The maximum power limit is safety-relevant. In electric cars, engineers aim to recuperate as much breaking energy as possible. Thus, the (disk) brakes perform only a part of braking. The primary energy should flow back to the energy storage system, using the motor(s) as a generator. In case of a full battery (high SoC means high OCV) in combination with low temperatures (high internal resistance), over-voltage may occur, which forces the battery system to switch off the main contactors [195, p. 166]. If this happens, braking power is lost immediately, which may cause a dangerous situation. Thus, peak power prediction is an important safety feature, which allows the car, as a system, to avoid those situations and control the distribution of the braking effort optimally between the two approaches.

Figure G.22 explains the behaviour in a simple model. On the left side, the "virtual" OCV, on the right side a resistor (internal resistance) presenting all losses. On the right side, the cells clamp voltage needs to stay in a particular range to allow for safe operation, eventually presented to the final user as the range 0% to 100% SOC. In the case of current flow, the additional voltage drop on the internal resistance may lead to a clamp voltage above or below the allowed range. In working conditions besides the 0% or 100% state, the maximum power (current) is limited as the allowed voltage drop on the internal resistance needs to be smaller than the voltage difference, see Equation G.13.

$$U < U_{\mathrm{OCV}}(100\%) = U_{\max} \wedge U > U_{\mathrm{OCV}}(0\%) = U_{\min} \tag{G.13}$$

where

U = cell clamp voltage (V)
$U_{\max}$ = maximum allowed voltage (V)
$U_{\min}$ = minimum allowed voltage (V)

Based on that and Equation G.26, the current needs to stay in certain limits, see Equation G.14.

$$\frac{U_{\min} - U_{\mathrm{OCV}}(SoC)}{R_I} < I < \frac{U_{\max} - U_{\mathrm{OCV}}(SoC)}{R_I} \tag{G.14}$$

where

I = cell current, pos. for charging (A)
$U_{\max}$ = maximum allowed voltage (V)
$U_{\min}$ = minimum allowed voltage (V)
R_I = internal resistance (Ω)
U_{OCV} = open circuit voltage, depending on the state of charge (V)

Those conditions need to be fulfilled, even at the end of the constant current pulse.

First, on the one hand, it needs to be considered that R_{I} is firmly temperature-dependent [187] (neg. temperatures limit the pulse power dramatically, ref. Figure G.24). On the other hand, R_{I} will shrink during the pulse (temperature increase due to power losses).

Secondly, R_{I} is SoC dependent [187]; in particular, R_{I} increases at low SoCs, and the OCV(SoC) voltage slope is higher compared to medium SoCs (see Figure G.23, and Figure G.29). Ageing influences R_{I} and capacity further, which leads to a changing maximum power over a cells life. All these effects should be considered during development, for example for the optimal operational strategy [187].

In working conditions with medium SoC, away from $U_{\min}$ and $U_{\max}$, other additional aspects, like maximum current limit for the conducting parts in the cell, need to be considered. In cells with protection devices, e.g. for over-current (attached PCB, PTC, CID), these may limit maximum power further. These maximum currents are stated in the cell data sheet, often in combination with the allowed period for certain currents.

Determining max. available power in a laboratory based on pulse duration

When determining the maximum available power of a cell, several strategies are feasible. First, one needs to select whether constant power or constant current pulses are realistic for the load of the desired system. Secondly, one needs to determine the maximum possible pulse for the defined pulse length. Due to a lack of (exact) information at the beginning of the procedure, an iterative method is recommended. It needs to be considered that the pulse according to the ISO standard discharges more charge than it refills during the charging phase. Thus, for the iterative method, the pulse length or current of the charging pulse should be modified to avoid a changing SoC for the next loop. In case of a constant power pulse, the amount of removed charge needs to be determined, and the recharge needs to be adjusted to avoid a moving SoC. As mentioned above, temperature influences the maximum pulse power. This effect needs to be considered for the waiting time between two trials.

Simple iterative method for maximum power determination
The first proposed method is to start with small currents for the pulses. The current is increased in small steps, until it reaches the desired end voltage (in case of discharge pulses U_{min}) precisely at the end of the pulse. The method needs to be repeated for other SoC, temperatures and desired pulse durations.

To improve on that method, one can determine the distance to the minimal/maximal voltage and derive bigger steps to speed up the pulse determination.

Using a model (R_I) to work more efficiently
Considering the model in Figure G.22, the current for the next iteration can be roughly derived by determining R_{I} and the voltage drop caused by ΔSOC from the last measurement and extrapolate it for the next step. For example, if a cell has a minimum voltage of 3 V, the OCV at start SoC was 3.2 V, and the final voltage of the last tested pulse was at 3.2 V, one can double the last current to approximate the desired maximum pulse. This will not yield perfect results immediately, but the procedure converges faster than other methods.

Power prognosis in real systems

In working systems, knowledge about maximum available power/current (examples above) is important, so that safety devices do not activate in aimed/controlled working conditions. To fulfil a power prognosis, a battery model (like subsection G.10.3) needs to be implemented which is able to forecast the maximum power for different pulse lengths. These results are submitted, e.g. to the acceleration control, which avoids pushing or pulling more current.

Determining temperature rise

Temperature changes with pulses as current flow causes losses [196–198]. Depending on the internal resistance one can calculate heat power (see subsection G.7.1). Ignoring heat exchange with the environment, the temperature rise can be determined by considering the heat capacity of the cell following Equation G.15. A more complex temperature model is given in subsection G.10.1. For short test pulses, heat exchange is of minor importance.

$$\Delta T = \frac{\int P_{\text{loss}}(t)dt}{m_{\text{Cell}} \cdot c_{\text{Cell}}} \tag{G.15}$$

where

$\int P_{\text{loss}}(t)dt$ = heat generated (J)
m_{Cell} = cell mass (g)
c_{Cell} = specific cell heat capacity ($\mathrm{J/g{\cdot}K}$)

Specific heat capacity of lithium-ion cells
According to different sources (0.83 J/gK [199], 1.01 J/gK [200], 0.80 J/gK [200], 0.85 J/gK [201], 1.02 J/gK [202], 0.95 J/gK [203], 0.80 J/gK [204], 0.92 J/gK [8]), the specific heat capacity of lithium-ion cells is approximately 1 J/gK. This value is easy to remember. In order to allow engineers to make rough calculations, it makes sense to remember this value.

G.8.2 Derived learning objectives

Based on the module targets, the dedicated theoretical lectures, and the aforementioned theory, the learning objectives for the experiment were selected.

- The student is aware that the maximum power lithium-ion cells can deliver is limited and he/she knows the reasons (temperature, internal resistance, Voltage $\leftarrow$ SoC) for this limitation.

- The student is able to estimate the behaviour of Lithium cells regarding their maximum power output dependent on the temperature.
  Addressed by tasks: Qu-C-17, 19

- The student understands that a voltage drop below U_{min} limits the maximum pulse power.
  Addressed by tasks: Qu-C-9, 18

- The student understands the connection between SoC and the possible maximum pulse power.
  Addressed by task: Qu-C-11

- The student can approximately derive the maximum power a cell can deliver (from capacity and/or internal resistance values).
  Addressed by task: Qu-C-10

- The student is able to estimate the relation between maximum power and pulse duration. The student is aware that requested pulse duration influences the maximum power.
  Addressed by task: Qu-C-12

- The student is able to determine the conditions at which the maximum power is extracted using a series of measurements. The student also can use the information in the data sheet correctly with respect to the battery cell. The student is able to determine the maximum discharge power by using an experimental setup.
  Addressed by task: Qu-C-16
- The student knows typical values regarding the heat capacity of lithium-ion cells.
  Addressed by task: Qu-C-8
- The student is able to estimate temperature developments which are caused by a power loss during pulse load on a cell and is able to calculate them approximately.
- The student is aware that standards need to be adjusted to fit to the necessary situation. (E.g. holding SoC constant during iterative procedures)

G.8.3 Derived lesson

As introduction, the students are requested to revisit the topic of internal resistance and refresh their knowledge about electrical power. They should look also through the ISO12405-1 standard they already used in experiment C1.

Determining the specific heat capacity of lithium-ion cells

The students received a table with different sources for specific heat capacities (see subsection G.8.1) and determined a value that was easy to remember.

Preparative questions

For preparation, the students had to prepare answers to the following questions. The expected answers are printed italic.

1. Which unit is used for power? *Watt*
2. How can one define the maximum pulse power? *It is the maximum pulse, defined by its length, where the cell clamp voltage is at the minimum allowed cell voltage at the end of the pulse (and during the pulse).*
3. Read the ISO12405-1 standard again. Consider the references to the pulse power. How are the pulse powers defined according to the ISO? *Minimal voltage during the pulse duration times the (constant) current.*
4. During that experiment the pulse power will be determined according the ISO12405-1 standard. Describe the basic approach after ISO12405-1. *Draw constant currents. The minimal voltage during the pulse (voltage at end of the duration) is multiplied with that current.*

5. Consult the data sheet of the cells to determine if it contains information about the maximum pulse power. *The data sheets do not have any information about that, a maximum current for short time is stated.*

6. What are the conditions under which the information of the data sheet was obtained? How would you grade the quality of the data sheet? *No additional information about boundary conditions, e.g. temperature, were given. The data sheet is of bad quality.*

7. Which unit is used for the specific heat capacity? ($\mathrm{J/gK}$)

8. Check the data sheets for the specific heat capacity. If there is no information, what could be used instead? *The average value we determined from research papers (*$1\,\mathrm{J/gK}$*)*

9. What is the value of the internal resistance of the cell from experiment C1? How is it related to the possible pulse duration, assuming a constant current? *xx* mΩ. *With increasing pulse duration, the cell gets heated up, which reduces the internal resistance.*

10. What is the maximum pulse power you expect from the cell? Will it increase or decrease with the pulse duration? *It was expected that the student uses a simple battery model to approximately determine the pulse power. It will decrease with pulse duration.*

Working teams

The working teams are assigned to perform the experiment at different SoCs: 80%, 65%, 50%, 35%, and 20%.

Determination of the heat capacity of the cells

The students were requested to determine the actual heat capacity of their device under test (cell). For that they used the mass stated in the data sheet and the value they gained from the sources stated in the instructions for specific heat capacity.

Changing the SoC to the desired value

All groups had to change the SoC to the desired group test condition. The cells were handed over with 100% and the actual capacity was stated on the cell. Herein the students had some degree of freedom regarding the discharging current. Based on the current, the discharge time was calculated (Equation G.6).

Designing a test plan

The students were asked to create an iterative testing plan to examine the correlation between the maximum discharge pulse power and the maximum pulse duration for 2 s and 10 s pulses. They were instructed to derive the values for 0 °C, 25 °C, 40 °C. They were then required to show that plan to the teaching staff and discuss it. The students were asked to discuss whether or not a recharge between iterations was necessary.

Performing the first test plan at room temperature

The students were asked to perform their test plan the first time for validation. They were reminded to stay in the voltage margin of the cell data sheet. The participants were tasked with reflecting upon whether their iterative plan was working to determine the right value or not. They were reminded to try to achieve a static SoC, and to include countermeasures against a moving SoC.

Delta T calculation

The working teams calculated the temperature change caused by a 10 s pulse with the determined maximum pulse power. A solution is to follow Equation G.30.

Data exchange

The working teams exchanged the final maximum pulse power results of their experiment (working teams' individual SoC) with all other groups.

Evaluation

The participants were requested to create graphs demonstrating the correlation between SoC and temperature against the maximum available pulse power (both for 2s, and 10s pulses). Additionally, they had to consider the measurement accuracy of the used devices when drawing the graphs.

G.9 Lesson D – Capacity and energy

G.9.1 Theoretical technical background

Discharge capacity with constant currents

The discharge capacity of an energy storage is the amount of charge that can be withdrawn when discharging it from the maximum allowed operating voltage to its minimum operating voltage (Equation G.16). It can be computed by integrating the current during the discharge procedure.

$$Q = \int_{t_{\text{start}}=0\text{ s}}^{t_{\text{end}}} I(t)dt \tag{G.16}$$

where

Q = discharge capacity (As)
t_{start} = starting time, fully charged battery cell, at maximum allowed voltage (s)
t_{end} = end time, cell reached minimum allowed voltage (s)
$I(t)$ = test discharge current profile (A)

Figure G.26 presents the standard method to determine the capacity of an accumulator cell. Herein, the fully CC-CV charged cell is discharged until the cell clamp voltage decreases down to its minimum, present on the data sheet. One can assume (due to $Q = t \cdot I$) that the discharge capacity is constant when varying the current and $t = {}^{1}/_{I}$. Nevertheless, the available charge Q varies for different currents. In 1897, Wilhelm Peukert observed this effect when discharging lead-acid batteries with a constant current. He found that the (empirical) Equation G.17 could represent the correlation between the discharge capacity of the battery and the applied discharge current [136]. The higher the discharge current, the lower the available capacity.

$$I^k \cdot t(I) = Q_{1\text{ A}} \tag{G.17}$$

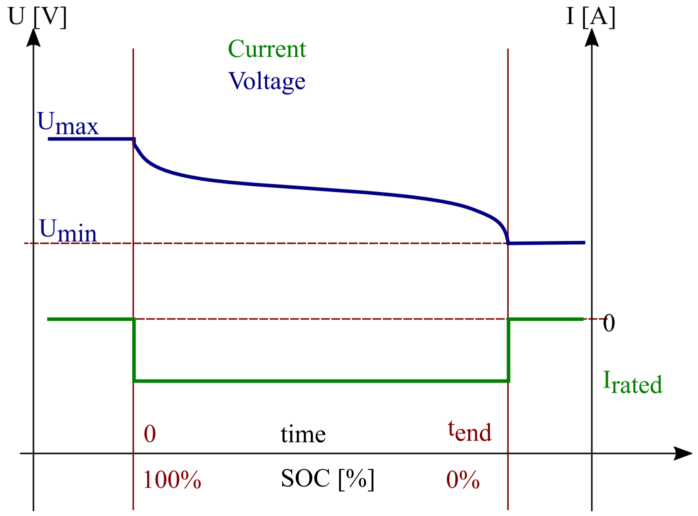

Figure G.26: Lesson D: Method to determine the capacity of a cell

where

$Q_{1\,\mathrm{A}}$ = discharge capacity at 1 A discharge current (As)
I = test discharge current (A)
k = dimensionless Peukert exponent, or Peukert's number (1)
$t(I)$ = time the battery needs to be completely discharged with I (s)

The Peukert exponent k equals 1.0 for an ideal accumulator, approximately 1.3 for a lead accumulator, and usually ranges below 1.05 for current lithium-ion cells when applying currents allowed by the data sheet. A significant disadvantage of the aforementioned (original Peukert) Equation G.17 is the unit ambiguity, in case the Peukert exponent k is not 1.0. Nevertheless, the Peukert exponent is always greater than 1.0 for real-world battery cells (due to the unavoidable presence of internal resistance).

Peukert normalised his equation with standard discharge currents of 1 A. Thus, all discharge capacities are calculated relative to that standard measurement. Nevertheless, modern cells provide very high capacities. The following Equation G.18 solves the unit ambiguity and allows for the usage relative to a differing standard discharge current (e.g. stated in the cell data sheet). In most cases a C-rate (see Equation G.7) of 1 C is used for that purpose.

$$Q(I) = t(I) \cdot I = t_{\text{rated}} \cdot \left(\frac{I_{\text{rated}}}{I}\right)^{k} \cdot I \tag{G.18}$$

where

$Q(I)$ = discharge capacity at test current (As)
$t(I)$ = time the battery needs to be completely discharged (s)
I = constant test discharge current (A)
t_{rated} = time the battery needs to be completely discharged at I_{rated} (s)
I_{rated} = rated discharge current (A)
k = dimensionless Peukert exponent, or Peukert number (1)

Given data of two discharge cycles, the Peukert exponent can be determined using Equation G.19 [149].

$$k = \frac{\log\left(t(I_2)/t(I_1)\right)}{\log\left(I_1/I_2\right)} \tag{G.19}$$

where

k = dimensionless Peukert exponent, or Peukert number (1)
$t(I)_{1,2}$ = time the battery needs to be completely discharged at $I_{1,2}$ (s)
$I_{1,2}$ = constant test discharge current (A)

Limits of Peukert's law

The limits of the Peukert Equation are well known [135, 149]. This equation is empirical and thus, not able to describe the effect based on its nature. Looking at the simple equivalent circuit (Figure G.22), the main reasons for the decrease of usable capacity when applying higher currents become clear:

- The internal resistance does not allow a discharge to the minimum OCV, as the clamp voltage reaches the minimum voltage before the "internal" voltage (voltage drop on internal resistance). The higher the current, the more of the charge becomes unavailable, as SoC stays above 0%. $\rightarrow Q\downarrow$
- The internal resistance increases when lithium-ion cells get empty (see Figure G.23). $\rightarrow Q\downarrow$
- The slope of the OCV(SoC) dependency is not constant. $\rightarrow Q\downarrow$
- Depending on the current (and actual internal resistance) the cell heats up more, which reduces the internal resistance (see Figure G.24) [188, p. 156]. Because of this effect, exponents below one can derived from real world experiments (see Figure G.27, real results at approximately 10 A) [149]. $\rightarrow Q\uparrow$
- That effect is also observed when cooling the cell using a thermostat to a constant temperature as a temperature gradient exists within the cell. $\rightarrow Q\uparrow$
- The temperature at which the experiment is conducted tends to have a high impact on the internal resistance, and thus also on the losses. $\rightarrow Q\updownarrow$
- Ageing effects depend on temperatures and currents [188, p. 156]. Thus, with high currents, more irreversible processes occur during discharge, reducing the capacity. $\rightarrow Q\downarrow$

Thus, with cells which are compatible for use with high currents (C-rates) or used outside the specified range, the existence of a single Peukert constant/exponent/number cannot be assumed, and researchers need to select the range of validity depending on the employed current and use a "locally" valid constant (Figure G.27) [149]. A mathematical method to determine the "local" Peukert constant was proposed in [149].

For future engineers, it is essential to know the general trend described by the Peukert equation and how to apply the equation correctly. It is also essential to know the limits of the empirical equation, to avoid that Equation G.19 and Equation G.18 are used outside the validity range.

Energy

Generally, the electrical power delivered by a cell can be derived from Equation G.20.

$$P(t) = U(t) \cdot I(t) \tag{G.20}$$

where

$P(t)$ = power (W)
$U(t)$ = cell clamp voltage (V)
$I(t)$ = current (A)

Thus, delivered energy can be determined by integration as in Equation G.21.

$$E = \int P(t)dt = \int U(t) \cdot I(t)dt \tag{G.21}$$

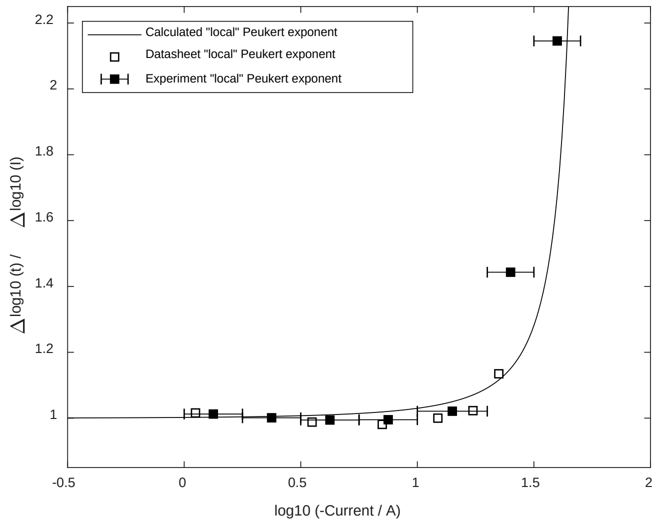

Figure G.27: Lesson D: "local" mean Peukert exponents, which are valid for a certain span of discharge current as presented by the bars. Own experimental Results for a 18650 2.5 Ah NMC lithium-ion cell INR18650-25R from Samsung SDI [149]

where

E = stored/delivered energy (J) at the cell clamp

In the case of a full constant current discharge procedure as mentioned above, the available energy content of a cell can be calculated using Equation G.22. In the case of time discrete equidistant data recording, the average voltage can be easily derived by calculating the average of all recorded voltages. The higher the discharge current, the lower the electrical energy, which can be extracted from the battery.

$$E = \int_{t_{\text{start}}=0\text{ s}}^{t_{\text{end}}} P(t)dt = I \cdot \int_{t_{\text{start}}=0\text{ s}}^{t_{\text{end}}} U(t)dt = I \cdot \bar{U} \cdot t_{\text{end}} \tag{G.22}$$

where

$\bar{U}$ = average discharge clamp voltage (V)
I = constant discharge current (A)
t_{end} = discharge duration (s)

When applying higher currents, $\bar{U}$ is lower (internal resistance), meaning E shrinks stronger compared to Q.

Rated capacity and nominal voltage derived from the data sheet
Usually, the rated energy storage capacity of a lithium-ion cell can be derived by multiplying the rated (nominal) voltage with the rated capacity of a cell [205, p. 104].

Energy density
Energy density of energy storages can be defined relative to volume (specific/volumetric energy density, J/m^3) and weight (gravimetric energy density, J/kg) [206, p. 18]. Furthermore, an engineer always needs to distinguish between the energy density at cell

and system level. Due to additional necessary components (like BMS, relay, housing, plugs, etc.) the system energy density ($\gtrapprox 80\,\mathrm{Wh/kg}$, [207, p. 189]) is lower than the cell energy density ($\gtrapprox 160\,\mathrm{Wh/kg}$ for $LiFePO_4$ or NMC cells, [207, p. 188], [208, p. 39]).

Efficiency

Charge efficiency
Charge efficiency of secondary cells (accumulators) is derived by experimenting during a constant current cycle within a fixed SoC range. For several full cycles, the charge flow is recorded during charging and discharging separately. Now, the ratio $Q_{\mathrm{discharge}}/Q_{\mathrm{charge}}$ is calculated. In an ideal battery, no charges will be missing, and thus the ratio will be one. In a real battery however, charges are lost due to self-discharge and side reactions/irreversible chemical processes.

Energy efficiency
Energy efficiency is the one of the most critical parameters in the context of lithium-ion energy storage.

Losses in the internal resistance are inevitable when charging or discharging battery cells. Thus, the battery dissipates heat during the cycles when charging and discharging. These losses are finally the major impact on energy efficiency. In other words, the necessary overpotential to cause current flow is the reason for energy efficiency below 100%: The cell clamp voltage at a specific SoC during charge and discharge processes differ; during charging it is above OCV(SoC), during discharging below.

Determining the efficiency of an energy storage system means comparing the energy necessary to charge it with the energy that can be withdrawn. As explained above, both numbers depend strongly on the current/power profile.

$$\eta = \frac{E_{\mathrm{discharging}}}{E_{\mathrm{charging}}} = \frac{-\int_{\mathrm{B}}^{\mathrm{A}} P(t)dt}{\int_{\mathrm{A}}^{\mathrm{B}} P(t)dt} = \frac{-\int_{\mathrm{B}}^{\mathrm{A}} U(t)\cdot I(t)dt}{\int_{\mathrm{A}}^{\mathrm{B}} U(t)\cdot I(t)dt} \tag{G.23}$$

where

η	= energy efficiency (1)
E_{charging}	= energy necessary to charge the battery from SoC A to SoC B (J)
$E_{\mathrm{discharging}}$	= energy drawn from battery while discharging from B to A (J)
$P(t)$	= power (W)
$U(t)$	= cell clamp voltage (V)
$I(t)$	= current (A)

To derive η, both the charging and discharging conditions need to selected realistically, as many factors (such as SoC range, chosen charging current, discharge load profile) influence the outcome. In practice, temperature, cell age (including cell storage and handling conditions), and recovery effects also influence the result.

In a laboratory environment, η can be derived by choosing standard conditions (like a specific SoC-range, temperature, constant currents for discharging and charging).

G.9.2 Derived learning objectives

Based on the module targets, the dedicated theoretical lectures, and the aforementioned theory, the learning objectives for the experiment were selected.

- The student is able to determine the capacity of a lithium-ion cell. The student knows the empirical equation by Wilhelm Peukert to describe the dependency of available capacity on current.
  Addressed by tasks: Qu-D-1, 3, 4

- The student is able to determine the energy efficiency of a cell during charge and discharge cycles. The student understands the reasons for the change in efficiency of an energy storage system.
  Addressed by tasks: Qu-D-8, 9

- The student understands the dependency of usable energy on a cell's voltage level, knows the influence of the cell type on nominal voltage and can estimate the influence on efficiency.
  Addressed by task: Qu-D-6

- The student knows the influencing factors on the available capacity of a lithium-ion cell. The student understands a simple equivalent circuit to explain the reasons for change of available capacity (current, temperature).
  Addressed by tasks: Qu-D-2, 5, 7

G.9.3 Derived lesson

The ultimate target of this experiment is to determine the capacity and energy content of a lithium-ion cell. Both values change depending on the load profile used. As time is limited, different constant current load profiles were distributed to the working teams and the overall result is gained by exchanging data between groups. Depending on the Run (number of working teams) different cell chemistry was also compared.

Grouping

To cover different conditions, the working teams received different parameters for their experiments. While the temperature for the experiment was always 25 °C, the discharge rates varied (1 C, 1.5 C, 2 C, 2.5 C, 3 C). In the German runs, all working teams were given a $LiMn_2O_4$ and a $LiFePO_4$ cell, while in the short runs only one cell type was employed. All cells were handed over fully charged by the laboratory engineer or the simulation was started accordingly (SoC, cell type) for the students. In case of two cell types, the discharge rates were combined in a way that the time span for discharging the first cell and the second cell were approximately the same.

Performing the discharge

All working teams programmed their test plan in the GUI and recorded voltage and current during the discharge procedures of both cell types. While the discharge was running, the conditions for discharge stop (reaching clamp under-voltage) were discussed based on the equivalent circuit diagram. All measured data were exchanged between the working teams. Thus, all teams were able to access time-discrete discharge data for different discharge rates .

Evaluation

All working teams created a diagram using the different discharge rates regarding each cell type (abscissa (x-axis) = time, ordinate (y-axis) = current). In the diagrams, they also marked the measurement uncertainty accordingly.

Then the teams determined the energy efficiency of the cycles with different C-rates. The time discrete data $(U(t), I(t))$ of the charging process of each cell type was uploaded to the university intranet and made available to all groups. They created a diagram of all energy efficiencies with all available C-rates for both cell types. For that, the participants were allowed to use their preferred computer program to perform the calculation based on the recorded data (Excel, Matlab, Flexpro).

After that, the teams evaluated the cell data sheet with regards to nominal voltage and capacity. They compared the product of both values ($U_{\text{rated}} \cdot Q_{\text{rated}} = E$) with the energy content derived in their own experiment. They compared the energy content of both employed cell types (both 600 mAh) and found that the energy content of the $LiFePO_4$ was smaller – as the voltage level was below the $LiMn_2O_4$ type.

The participants were instructed to calculate how high the temperature rise would be if the thermostat were inactive, assuming a good thermal isolation around the cell. They compared the theoretical heat development based on the internal resistance determined in Lesson C1 (section G.7) with a calculation based on the determined energy efficiency.

Now the working teams were requested to determine the Peukert exponent of both cell types based on the data collected by all working teams. This meant they had to use the different results (for different C-rates) per cell type. Thus, a graphical solution with double-logarithmic depiction (X-axis lg(I/1 A), Y-axis lg(t/1 s)) was useful, see Figure G.28. Alternatively, they were allowed to use mathematical methods like least square error fitting the computer program of their choice to calculate the value of k.

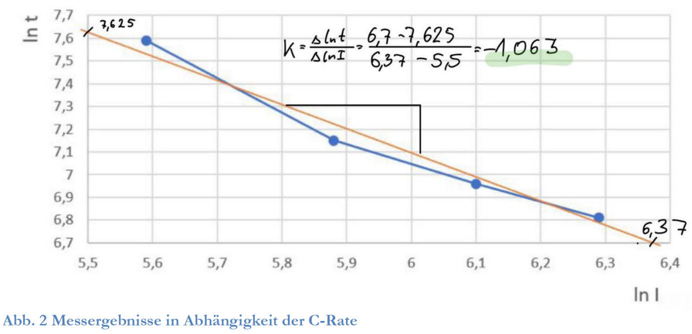

Figure G.28: Lesson D: Example solution of a working team, graphical determination of the Peukert exponent based on four measurements.

G.10 Simulation Workshop E – Equivalent circuit with Matlab/Simulink and fitting

Design and operation of safe high-performance electric vehicles depend on precise knowledge of the state of their electrochemical energy storage systems [209, 210]. Batteries are the key energy storage in full electric power-trains. Hybrid and fuel-cell power-trains are used as secondary energy storage. The efficiency and dynamic behaviour of the electric energy storage have a significant impact. Optimisation via simulation and proper functioning of battery management systems often rely on battery models that accurately represent the battery characteristics.

Thus, for electric mobility engineers, it is essential to understand not only the cell behaviour itself (as investigated in the laboratory experiments), but also the system behaviour (serial and parallel connection of cells) and the impact of battery models with different complexities.

G.10.1 Theoretical technical background

Simulation of batteries (cell packs/stacks)

To simulate battery packs (which consist of a certain number of cells), it is necessary to emulate the cell as the base model. Then, the model of the whole system can be easily derived.

If several cells are in series connection, the voltage multiplies.

$$U_{pack} = U_{\mathrm{Cell}} \cdot N_{\mathrm{ser}} \tag{G.24}$$

If several cells are in parallel connection, the current distributes on these cells.

$$I_{\mathrm{Cell}} = \frac{I_{pack}}{N_{\mathrm{par}}} \tag{G.25}$$

Additionally, depending on the required accuracy, one may have to consider the voltage drops on the connectors between the cells and the behaviour of the conducting electromechanical components (e.g. fuses, contactors, bus bars, and plugs).

Simulation of battery cells

Battery cell models can be divided into three main types: experimental, electrochemical (first-principle), and electric circuit-based (empirical) models [149].

Experimental models
Experimental models interpolate the required data mathematically based on measurement results, requiring lengthy laboratory tests [149]. An example for such an empirical model is [135]. The proposed formula describes the available discharge time of a specific battery cell dependent on the used discharge current. The empirical formula was parameterised by the measurement of a selection of discharge currents.

These models have the disadvantage of delivering only an interpolation of the directly measured attribute. They cannot serve any predictive function.

First-principle models (bottom-up approach models)
First-principle models predict the behaviour by a description of the fundamental processes, such as thermodynamics, mass transport, and electrochemistry. These models need exhaustive information on the materials and cell structure. Thus, they are very detailed and specific for a given cell type, which makes them complex to configure and calibrate. Electrochemical models require an in-depth understanding of underlying mechanisms – in many cases, it may be challenging to realise correct model parameters for batteries [211].

Empirical models (top-down approach models)
Empirical models describe the input and output behaviour based on experimental data, regardless of the chemistry and physics involved. They are not predictive, but intuitive and straightforward to build and use. When the general behaviour of a battery is to be represented without an understanding of the underlying processes, equivalent electric-circuit models proved to be very useful. Depending on the complexity of the chosen equivalent circuit, they can have good accuracy for real world use-cases when describing dynamic battery behaviour, without the necessity of tedious testing of all operating conditions. Moreover, they can easily be adapted and implemented in battery management systems [149, 212]. Such a model was selected for the lesson.

Equivalent electric circuit models simulate the behaviour of the real cell by combining ideal electrical elements like capacitance, resistance and voltage sources, which are known to undergraduate electrical engineering students. The parameters for the ideal elements are taken from parameterised formulas or look-up tables. These tables describe the most important conditions. In batteries, these are generally temperature and state of charge dependencies [149]. For increasing accuracy, some parameters – such as internal resistance – should be modelled including a current dependency (direction and absolute value) [213].

Static Model
The simplest way to model a battery cell (Figure G.22) is a combination of an ideal voltage source (to provide the open circuit voltage U_{OCV}) and a resistor (to model the internal resistance R_{I}, which presents the voltage drop and losses in conditions when current flows) [149]. This model can be parameterised by static measurements, but does not include cell dynamics [214].

Kirchhoff's second rule allows for the calculation of the battery terminal voltage U.

$$U = U_{\mathrm{OCV}} + I \cdot R_I \tag{G.26}$$

where

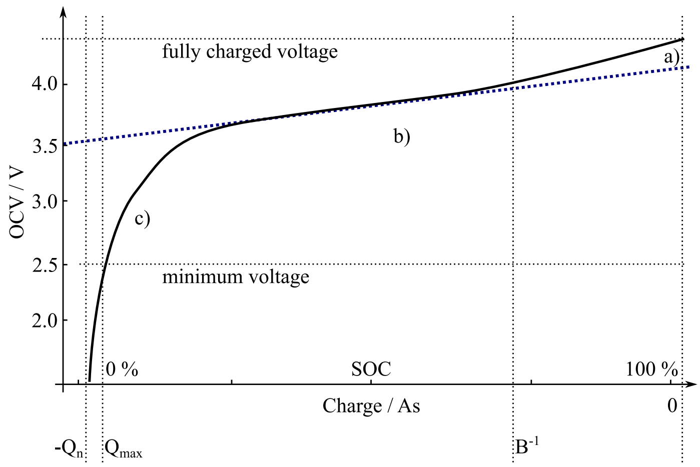

Figure G.29: Workshop E: Typical behaviour of the open-circuit-voltage dependent on SoC

R_I = internal resistance (Ω)
U = calculated battery terminal voltage (V)
U_{OCV} = open circuit voltage, depending on the state of charge (V)
I = cell current (A)

To calculate the open circuit voltage, the state of charge of the battery cell must be known. The zero point for capacity (integral of the current) according to Equation G.27 can be set freely. For the lessons, a zero point after a full constant current/constant voltage charge to the maximum allowed cell voltage was chosen. In combination with the definition of the direction of the current in Figure G.22, the selected zero point leads to negative values for Q under normal operating conditions as shown on the abscissa in Figure G.29. [149]

$$Q(t) = \int_{t_{CCCV charge}}^{t} I dt \tag{G.27}$$

The state of charge (defined as 100% for a full and 0% for an empty cell) can be derived using Equation G.28.

$$SoC(t) = 1 + {}^{Q(t)}\!/_{Q_{max}} \tag{G.28}$$

where

$SoC(t)$ = state of charge (1)
$Q(t)$ = charge (As)
Q_{max} = maximum (dis)charge capacity (As)

The correlation between SoC and open circuit voltage U_{OCV} may be given by a look-up table or a parameterised equation (e.g. [149, 215]). To derive all parameters many measurements (for a method, see [149]) or an in-depth analysis of the available cell data sheets (for a method, see [215]) are necessary.

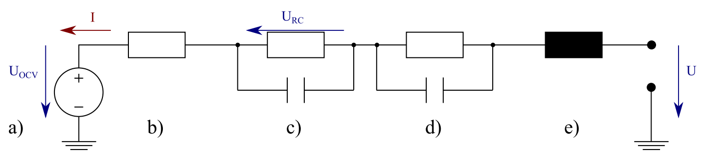

Figure G.30: Simulation model of the electrical characteristics of the battery cell. a) Open circuit voltage b) Internal resistance c) & d) Cathode and anode double layer capacitance and charge transfer resistance e) Inductance. Such models are widely used in battery simulation/characterisation to meet the frequency dependent behaviour of cells (e.g. [151]).

Dynamic model

The driven load decides the current when discharging. In the case of electric vehicles, the current can be very dynamic (e.g. in case of acceleration, braking). For such circumstances, it is necessary to enhance the simple static model (Figure G.22) with one or more RC branches (Figure G.30) to emulate dynamic behaviour. Each of those RC branches can handle responses for different time constants. [214]

The voltage drop of these RC elements can be expressed by Equation G.29.

$$U_{RC}(t) = U_{RC@t=0} + \frac{1}{C} \cdot \int_{t=0}^{t} (I - \frac{U_{RC}}{R}) dt \tag{G.29}$$

where

U_{RC} = voltage drop at the RC branch (V)
C = capacitance as part of the RC branch (F)
R = resistance as part of the RC branch (Ω)

The number of RC branches defines the order of the model. Figure G.30 represents a second order model enhanced by the wire inductance. On one hand, the higher the order of the model, the higher the accuracy. On the other hand, the effort to parametrise these models increases with the order of the model. All parameters are additionally dependent on the operating conditions (e.g. temperature), which increases the effort for experiments and curve fitting. Thus, a compromise has to be found.

Thermal modelling

Temperature influences internal resistance the most, which is the dominating factor for the efficiency of the battery system and also limits the maximum peak power (subsection G.7.1, subsection G.8.1). While the internal resistance of a typical NMC lithium-ion cell is slightly dependent on the state of charge (deviation < factor 2 in the full range), the temperature dependency of the internal resistance is much stronger (> factor 4 at 0°C compared to room temperature) [216, p. 93, Fig. 7–18] [191].

A simple thermal model is shown in Figure G.31. It is based on direct contact cooling (conductive heat transfer). Assuming the temperature of the cooling system is held constant, it is easy to understand when compared to models combining conductive cooling with convection (air, cooling fluid) or radiation. Here the temperature of the cell mass is modelled as a single volume of homogeneous temperature, thus

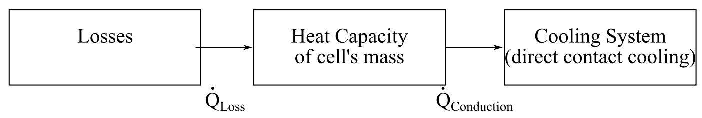

Figure G.31: Simplified thermal model of a battery system, demonstrates the main flows of heat, which are the internal battery losses and the heat transferred to the cooling system

this model ignores temperature gradients inside cells or between cells in a cell stack. These temperature gradients are of particular importance for more detailed models [196–198].

Losses that generate heat in the battery cell and system can be calculated from the equivalent circuit model shown in Figure G.30. Generally, electrical losses are calculated from $P = \Delta U \cdot I$. As the model is based on a serial connection of the different parts, the sum of all losses (compare Figure G.30) can be described by Equation G.30.

$$\dot{Q}_{\text{Loss}} = P_{\text{Loss}} = (U - U_{\text{OCV}}) \cdot I \tag{G.30}$$

These losses lead to an increased temperature in the battery cells, which leads to heat flow into the environment of the cells. The differential form of Fourier's law (law of heat conduction, Equation G.31), states that the amount of heat transferred through a material is negatively proportional to the temperature gradient and to the area through which the heat flows. [217, p. 38]

$$\mathbf{q} = -k \cdot \nabla T \tag{G.31}$$

where

$\mathbf{q}$ = local heat flux density ($^{\text{W}}/_{\text{m}^2}$)
k = material's conductivity ($^{\text{W}}/_{\text{m}\cdot\text{K}}$)
∇T = temperature gradient ($^{\text{K}}/_{\text{m}}$)

Fourier's law can be reformulated and simplified in terms of intensive properties and one-dimensional form as shown in Equation G.32 .

$$\dot{Q}_{\text{Conduction}} = {}^{1}/_{R_{\text{thermal}}} \cdot \Delta T \tag{G.32}$$

where

$\dot{Q}_{\text{Conduction}}$ = heat flux (W)
R_{thermal} = thermal resistance ($^{\text{K}}/_{\text{W}}$)
ΔT = temperature difference (K)

Assuming conductive heat transfer in a static state, the heat flow is proportional to the temperature difference between the battery cell and cooling system. R_{thermal} depends on the arrangement parameters (geometry, materials), while the temperature difference is the delta between the cooling system and battery cells $\Delta T = T_{\text{Cell}} - T_{Cooler}$.

Using the simple thermal model given in Figure G.31, the actual temperature of the battery cell can be derived. The difference between in- and outgoing power flows leads to an increase or decrease in the temperature of the cell, as shown in Equation G.33.

$$T_{\text{Cell}} = T_{Cell@t=0} + \frac{1}{m_{\text{Cell}} \cdot c_{\text{Cell}}} \int_{t=0}^{t} (\dot{Q}_{\text{Loss}}(t) - \dot{Q}_{\text{Conduction}}(t))dt \tag{G.33}$$

where

m_{Cell} = cell mass (g)
c_{Cell} = specific cell heat capacity ($^{\text{J}}/_{\text{g}\cdot\text{K}}$)

The specific heat capacity of lithium-ion cells is approximately 1 $^{\text{J}}/_{\text{gK}}$ (see subsection G.8.1).

Fitting / optimising of parameters by real measurements of profiles

The students of the German runs visited a modelling lecture with training on the computer in the same semester, but it was focused technically on Simulink. Model parametrisation was not taught.

Model parameters can be selected and optimised automatically by comparing real measured data to the results of the simulation. In this way, one or more model parameter can be optimised to emulate the special conditions chosen for the real measurements. In the case of equidistant time-discrete data, smaller residuals ($\text{residual}_i = \text{measured}_i - \text{simulated}_i$) describe a better parameterised simulation.

For example, a given current profile is injected in the real cell and the voltage response is recorded. The same current profile is used as an input parameter to a simulation model of a cell, and the simulated voltage response is recorded. When both results are compared, the residuals in case of a good match would be small, indicating a working model with well chosen parameters.

By defining an error-function which describes the quality of the simulation parameters to produce a simulated output fitting to the real measured results, Matlab can iteratively minimise that function and thus calculate the best fitting model parameters. Equation G.34 describes an error-function based on least squares. Thus, a minimising algorithm will minimise the sum of squared residuals.

$$\text{error}(\text{parameters}) = \sum_{n}^{i=1} (\text{mea}_i - \text{sim}_i(\text{parameters}))^2 \tag{G.34}$$

G.10.2 Derived learning objectives

Based on the module targets, the dedicated theoretical lectures, and the aforementioned theory, the learning objectives for the experiment were selected.

- The student knows a simple model of a lithium-ion cell in Matlab/Simulink.

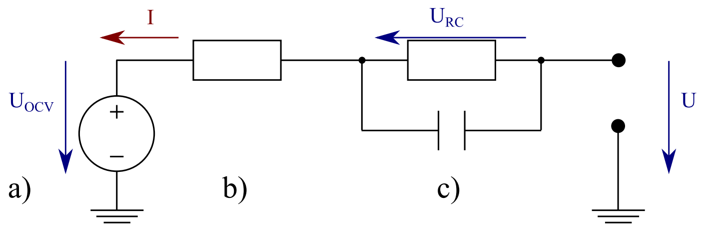

Figure G.32: Taught model a) Open circuit voltage b) Internal resistance c) & d) Cathode and anode double layer capacitance and charge transfer resistance

- The student is aware of the limits of the model and is able to enhance it for specific needs.
- The student is able to fit cell parameters for the simulation based on real measurements.
- The student knows how to simulate cell stacks and can use that knowledge to test a dimensioning of a battery system.
- The student is able to apply power and current profiles to test a simulated battery system.

G.10.3 Derived lesson

As lesson time was limited to 3 hours (minus the 10 minutes for the test regarding content area D) and 4th semester undergraduate students do not have a lot of experience in using Matlab/Simulink, learning objectives in the direction of the usage of Matlab/Simulink were avoided. Instead, they received a prepared model in the week before the lesson. It was available in the intranet of the university (Moodle). This model simulated a 20 A h lithium-ion cell ($LiFePO_4$) [218], a pouch type, with a bigger capacity than the cells used in the laboratory sessions. During the lesson, this model was tested to gain experience, and it was modified to describe the cell type used in the previous laboratory sessions.

Taught simulation model and presentations

The model in Figure G.30 was simplified by ignoring the wire inductance and combining the equivalents of cathode and anode in one RC-element and is shown in subsection G.10.3. The thermal model of the cell was a simple conductive cooling model. None of the parameters are state of charge or temperature dependent, allowing for an easy understanding.

One week before the workshop, the students were asked to prepare a short presentation on a subtopic of the model. For the presentation, they work in the same small working teams as in the laboratory sessions. The topics were given as a marked area of the handed Matlab/Simulink model presented in Figure G.33.

The first group investigated the open circuit voltage dependent on the state of charge (see experiment B section G.6). Here, the aspects regarding the starting SoC

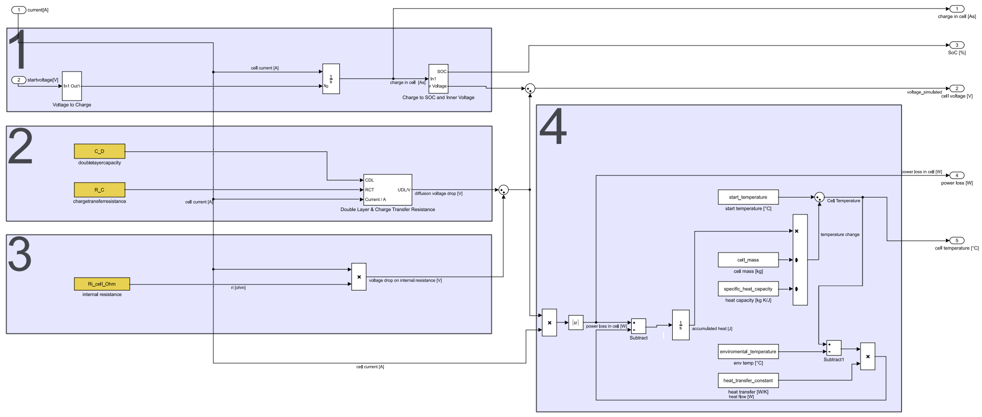

Figure G.33: Workshop E: Matlab/Simulink model of the cell – 1) Open circuit voltage, 2) RC 1. Order, 3) Internal Resistance, 4) Thermal Model

need to be mentioned. The open circuit voltage was derived after building the current integral through a look up table. The second group presented on the voltage drop on the first order RC branch, describing mainly the double layer capacity and the transfer resistance of cathode and anode of the cell. The focus was on the differential equation describing the behaviour of the RC branch. The third group's topic was internal resistance, which was investigated in experiment C1 (page 307 in the appendix). Here, the model was straightforward: a constant resistance. The voltage drop was given by Ohm's law. All groups were asked to gather up the results of the respective experiments. The groups were asked to reflect on the capabilities of the model to describe the experienced effects, and discuss if important effects were missing. The fourth group presented the thermal model. As mentioned above, the internal voltage drop of the model creates heat, while conductive cooling removes heat. The resulting temperature was calculated by the usage of the cell's specific heat capacity and mass. A fifth group presents the Matlab environment of the cell simulation, for example how the cell parameters are derived from the workspace and how the current profiles are given. As a result, all students know about the next step to improve the model: including a temperature dependent internal resistance, fed by the result of the thermal model.

Fitting model parameter to given experimental results

The students received a model with parameters of low quality, which approximately fit the parameters of the real cell. The students analysed a script which iteratively called on the Simulink simulation to optimise three parameters: internal resistance, double-layer capacity, and transfer resistance (yellow in Figure G.33). The students investigated the programming and the procedure to derive optimised model parameters based on the real recorded current profile and voltage response of a battery cell. Particularly, they learned how to create an error function out of time-discrete Simulink data and how the start parameters of the fitting procedure influenced the result. Finally, the students compared the results based on the 20 A h cell with the parameters gained in the laboratory experiments with the small (<1 A h) cells.

A visible script in Matlab was deliberately chosen over the standard simulation parameter fitting tool [219] included in the software. This way, the students learned parameter optimisation step by step, a general knowledge that can be used in several engineering contexts.

Cross-faculty project E-Falke (e-falcon) as an example to create an energy storage system

Technische Hochschule Ingolstadt initiated the cross-faculty project E-Falke (e-falcon). The overall goal of this project is the substitution of the combustion engine of a powered glider aeroplane with an electric motor. Over 130 students were involved in this project, reaching for the fulfilment of pre-defined goals, such as the feasibility evaluation of the project, ground bench testing, integration of the system into the airframe and the official flight certification of the German Federal Aviation Office. [146]

Simulating a battery system

The aeroplane-project is based on the above-mentioned 20 A h $LiFePO_4$ cells. In this phase of the workshop, the students derive a model of the whole battery pack (17s3p, an assembly of 51 cells, 3 in parallel, 17 in serial connection) on their own. Their task was to check if the battery pack fulfils the requirements regarding energy content to be awarded official flight certification. The minimal mechanical power profile was given (4 min 14.5 kW to climb the minimum height, followed by 6.7 kW to travel without loss of height). Additionally an efficiency factor of 90% (power inverter, motor, cables) was stated. The students simulate how many minutes the plane can be used after climbing.

Transferring the system to the battery cell of the laboratory experiments

In this phase the students were asked to transfer the 20 A h cell model to a model of the battery cell used in the laboratory experiments (e.g. data for open-circuit-voltage of B and internal resistance of C1). Then, the students were requested to use Matlab/Simulink to create an equivalent system out of these cells to meet the performance and voltage level of the original system of 20 A h $LiFePO_4$ (17s3p) cells.

Appendix H

Testing the gained knowledge

Consistency between desired learning outcomes and chosen assessment methods is necessary [2]. This chapter was split into two sections, to describe first the set of questions and tasks which were used to check the participants' knowledge, and secondly, which of these tasks were combined to form the tests used in the different study runs.

H.1 Questions and tasks

This section contains a description of all tasks which were used to test the students' knowledge in the short tests. Table H.3 shows an overview of all questions. The tasks were categorised according to the item format (Table H.1) and the main learning objective the question addresses (Table H.2). The accompanying laboratory lessons in which the tested knowledge was taught can be found in chapter G. In the present study, tasks were designed to test knowledge and understanding.

Table H.1: Item formats

Abbr.	Item format
DG	Draw/Graph
DS	Draw/Sketch to explain
MC	Multiple-choice
SC	Single-choice
TV	Text/state or calculate a value
TR	Text/state reason or equation

Table H.2: Main learning objective categories

Abbr.	Category
BB	Battery Behaviour
BP	Battery Parameters
BSD	Battery System Design
ES	Experimental Setup

Table H.3: Questions for testing the knowledge of the participants

Question ID	Points max.	Point res.	Item format	Learning objective category	specific
Qu-A-1	2	1	MC	BSD	Transfer-resistance of connections
Qu-A-2	1	1	SC	BSD	Priority while designing connect.
Qu-A-3	1	1	SC	BSD	Cable-lug connection
Qu-A-4	1	1	SC	BSD	Fuses – R dep. on current
Qu-A-5	2	1	MC	ES	General – Kelvin-Arrangement
Qu-A-6	2	1	MC	ES	General – Isolation resistance
Qu-A-7	2	.5	DS	ES	General – KA – simple equipment
Qu-A-8	2	1	DS	ES	General – Kelvin-Arrangement
Qu-A-9	2	1	MC	ES	General – Isolation resistance
Qu-A-10	2	1	TR	BSD	Cable-lug connection
Qu-B-1	1	1	MC	ES	General – Thermostat
Qu-B-2	2	.5	DS	BP	OCV Curve
Qu-B-3	1	1	MC	ES	General – Electric Sink/Load
Qu-B-5	2	.5	DS	ES	Batteries – Temperature change
Qu-B-6	2	1	TV	BB	Battery Voltage
Qu-B-7	2	1	TV	BB	Battery Voltage
Qu-B-8	1	1	SC	BP	State of Charge
Qu-B-9	2	1	MC	ES	Batteries – Charging Method
Qu-B-10	1	1	TV	ES	Batteries – Calc Charging Duration
Qu-B-11	2	1	TR	ES	General – Measurement methods
Qu-B-12	5	.5	DG	BB	Battery Voltage and Current
Qu-B-13	1	1	MC	ES	General – Galvanostat
Qu-B-14	1	1	TR	ES	Batteries – Open Circuit Voltage
Qu-C-1	1	1	SC	BB	Battery Efficiency R_{int}
Qu-C-2	2	.5	DG	ES	Batteries – Measurement methods
Qu-C-3	2	.5	DG	BB	Voltage
Qu-C-4	1	1	SC	BB	Temperature Dependency of R_{int}
Qu-C-5	1	1	SC	BB	Temperature Dependency of R_{int}
Qu-C-6	1	1	TV	BB	Internal Resistance
Qu-C-7	1	1	SC	BB	Current Dependency of the R_{int}
Qu-C-8	1	1	TV	BB	Specific heat capacity
Qu-C-9	2	1	MC	BB	Max. Power
Qu-C-10	1	.5	TV	BB	Max. Power
Qu-C-11	1	1	SC	BB	Max. Power
Qu-C-12	1	1	SC	BB	Max. Power
Qu-C-13	4	.5	DG	ES	Batteries – Measurement methods
Qu-C-14	4	.5	DG	BB	Voltage
Qu-C-15	3	.5	DG	BB	Temp. dependency of the R_{int}
Qu-C-16	2	.5	TR	ES	Batteries – Max. Power
Qu-C-17	1	1	SC	BB	Max. Power vs. Temp.
Qu-C-18	1	1	SC	BB	Max. Power vs. R_{int}
Qu-C-19	1	1	SC	BB	Max. Power vs. Temp.
Qu-D-1	2	1	TR	BP	Peukert
Qu-D-2	1	.5	MC	BB	Usable Capacity
Qu-D-3	1	1	SC	BP	Peukert
Qu-D-4	1	1	TR	BP	Peukert
Qu-D-5	1	1	MC	BB	Energy Efficiency
Qu-D-6	1	1	MC	BP	State of Charge
Qu-D-7	2	1	MC	BB	Usable Capacity
Qu-D-8	1	1	SC	BB	Efficiency
Qu-D-9	1	1	SC	BB	Transfer: Temp – R_{int} – Efficiency

H.1.1 Question types (item formats) & marking

The questions in the knowledge tests can be categorised into these six question types (item formats):

Draw/Graph

Students were asked to draw a graph or diagram to illustrate certain aspects of the laboratories. Points were awarded for correct graphical representation and analysis.

Draw/Sketch to explain

Students were asked to draw/sketch a certain arrangement in order to explain certain aspects of the laboratories. Answers were scored according to preset criteria.

Multiple-choice

Questions were given with a number of options. Multiple answers could be correct. Points were awarded for checked correct answers, and subtracted for checked incorrect answers. In the case of more incorrect than correct answers, the question was marked 0 points.

Single-choice

Questions were given with a number of options. A single option was correct. Full points were awarded for checking the correct answer, unless another, incorrect answer had also been checked.

Text/state or calculate a value

Students were asked to perform calculations or state empirical values based on the knowledge they acquired in the laboratories. Points were awarded for correct values, with partial credit in the case of imprecise or partially correct answers.

Text/state reason or equation

Students were asked to elaborate on their reasoning and could usually earn points for correct statements. These questions often asked for free-form answers, and were evaluated as such.

H.1.2 Questions regarding experiment A

The following tasks address the learning objectives from lesson A (section G.5). Each question is stated and identified with a question type. In the case of Multiple-choice or Single-choice questions, the correct answers are marked with a cross.

Question Qu-A-1

Multiple-choice – 2 points, increment 1
Learning objective: "Battery System Design", Subtopic "Resistance of connec-

tions".
Taught in experiment A2 (subsubsection G.5.3.2).

What has a considerably positive influence to the reduction of contact resistance of an electrical cable-lug connection?

☒ *Increased contact pressure*
☐ *Better conducting material of the screw*
(☒) Smoothed contact area
☒ *Removing of oxides on the contact area*

Question Qu-A-2

Single-choice – 1 point.
Learning objective: "Battery System Design", Subtopic "Design of connection".
Taught in experiment A2 (subsubsection G.5.3.2).

The higher the intended current on a connection, the . . .

⊗ *Lower the contact resistance has to be.*
◯ *Higher the contact resistance has to be.*

Question Qu-A-3

Single-choice – 1 point.
Learning objective: "Battery System Design", Subtopic "Cable-lug connections".
Taught in experiment A2 (subsubsection G.5.3.2).

When assembling a connection of a cable-lug with a conductor rail, the washer has to be ...

◯ *between the cable-lug and the conductor rail.*
⊗ *between the screw and the cable-lug.*

Question Qu-A-4

Single-choice – 1 point.
Learning objective: "Battery System Design", Subtopic "Resistance of Fuses / Resistance dependency on current".
Taught in experiment A1 (subsubsection G.5.2.1).

Which fuse has comparatively the higher resistance?

◯ *High rated current*
⊗ *Low rated current.*

Question Qu-A-5

Multiple-choice – 2 points, increment 1
Learning objective: "Battery System Design", Subtopic "General Experimental Setups – Kelvin-Arrangements".
Taught in experiment A1 (subsubsection G.5.2.1).

Which statements are suitable for the four wire measurement?

☐ *The current at the sense wire is high,*

☒ *The measurement at the sense wire is of high impedance*
☒ *There is no significant potential difference at the sense wire*
(☐) The current of the force wire is variable adjusted fitting to the resistance
☐ *The potential difference of the force wire is subtracted by the measuring instrument*
and ☐ *The measuring instrument knows about the resistance of each wire and subtracts the values later*

In case the fourth option ("The measuring instrument knows ...") was selected, no point was subtracted, as the answer might be right with special measurement equipment.

Question Qu-A-6

Multiple-choice – 2 points, increment 1
Learning objective: "Experimental Setup", Subtopic "General Experimental Setups – Isolation resistance".
Taught in experiment A3 (subsubsection G.5.3.3).

Which parameters are significant for the insulation resistance?
☒ *Distance between the conductors*
☒ *Humidity of the surrounding area*
☒ *Testing voltage*
☐ *Thickness of the wires*
☒ *Measuring time*
☐ *Length of the wires*

Question Qu-A-7

Draw/Sketch to explain – 2 points, increment .5
Learning objective: "Experimental Setup", Subtopic "General Experimental Setups – Kelvin-Arrangement with simple equipment".
Taught in experiment A1 (subsubsection G.5.2.1).

Make a sketch on the back of this sheet. The drawing should show a setup with a multi-meter and power-supply for measuring of a small resistance. Label the settings of all devices if applicable.

One point was awarded for the correct arrangement. One point was given for the selection of the right settings (e.g. the CC adjustment of the power supply and the usage of the multimeter to measure voltage).

Example Solution in Figure H.1.

Question Qu-A-8

Draw/Sketch to explain – 2 points, increment 1
Learning objective: "Experimental Setup", Subtopic "General Experimental Setups – Kelvin-Arrangement".
Taught in experiment A1 (subsubsection G.5.2.1).

Make a sketch on the back of this sheet. The drawing should show an appro-

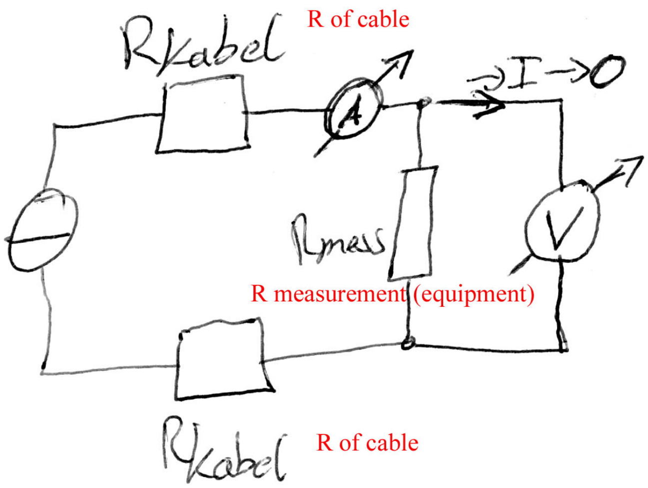

Figure H.1: Example solution of a participant for Qu-A-7

priate measuring probe tip which is able to carry current and voltage for four wire measuring without producing a measurement error. Explain its function.

Example solution in Figure H.2.

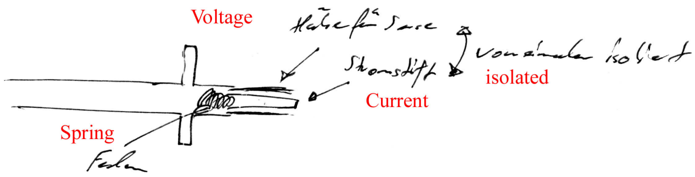

Figure H.2: Example solution of a participant for Qu-A-8

One point for a correct drawing (whether it showed the type of probe used in the laboratory experiment, or any other suitable type of probe), one point for a clear description of its function.

Question Qu-A-9

Multiple-choice – 2 points, increment 1
Learning objective: "Experimental Setup", Subtopic "General Experimental Setups – Isolation resistance".
Taught in experiment A3 (subsubsection G.5.3.3).

Why is it not possible to perform an insulation measurement with a standard multi-meter in ohm range?

☐ *Lack of strength of current*
☒ *Lack of high voltage*
☒ *Lack of time measurement*

Question Qu-A-10

Text/state reason or equation – 2 points, increment 1
Learning objective: "Experimental Setup", Subtopic "Cable-lug connection".

Taught in experiment A2 (subsubsection G.5.3.2).

Elaborate on your reasoning for your answers to Qu-A-3.

One point was awarded per correct statement. Possible solutions included:
there are less contact resistances
steel washer is a bad conductor due the low conductance of steel
reduced conducting area.

H.1.3 Questions regarding experiment B

The following tasks address the learning objectives from lesson B (section G.6).

Question Qu-B-1

Multiple-choice – 2 points, increment 1
Learning objective: "Experimental Setup", Subtopic "General Experimental Setups – Thermostat".

A thermostat ...

☐ *Discharges a battery cell notable mild*
☒ *Ensures a certain temperature*
☐ *Always operates with a cooling medium*

Question Qu-B-2

Draw/Sketch to explain – 2 points, increment .5
Learning objective: "Experimental Setup", Subtopic "OCV Curve".

Draw a chart of a typical OCV(SoC) curve of a lithium-ion cell on the back side of this sheet. Label both axis. Name the units.

Example solution in Figure H.3.

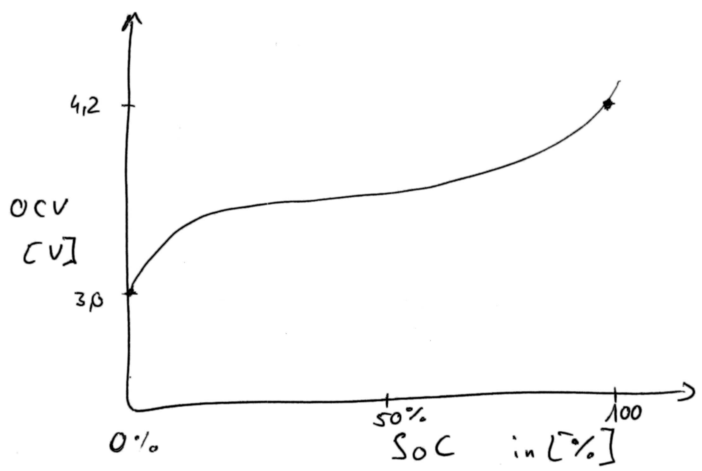

Figure H.3: Example solution of a participant for Qu-B-2

One point for the right axes and for the fact that voltage rises with an increase of SoC. One point for the right slope change at the ends. No subtraction if absolute scales were missing. Half point subtraction in case the x-axis was labeled time instead of SoC/charge.

Question Qu-B-3

Single-choice – 1 point
Learning objective: “Experimental Setup”, Subtopic “General Experimental Setups – Electric Sink/Load”.

An electronic load (sink) ...

☐ *Powers a circuit if necessary*
☐ *Allows for charging of a battery cell to a specific SoC*
☒ *Converts electrical energy mostly in heat or injects it into the power grid*
☐ *Is always operated with a cooling medium*

Question Qu-B-5

Draw/Sketch to explain – 2 points, increment .5
Learning objective: “Battery Parameters”, Subtopic “Battery Experimental Setup – Behaviour after Temperature change”.

Why is it necessary to wait with the measurement of voltage of a cell if you changed the temperature? Make a sketch of a battery cell on the back of this sheet to explain.

Full points were awarded for a sketch with text explaining the temperature gradient and the time required for an equal temperature. If only something like ‘transient effects’ was stated, one point was awarded.

Example solution in Figure H.4.

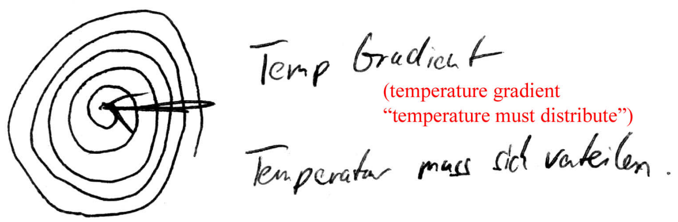

Figure H.4: Example solution of a participant for Qu-B-5

Question Qu-B-6

Text/state or calculate value – 2 points, increment 1
Learning objective: “Battery Behaviour”, Subtopic “Battery Voltage”.

State the approximate voltage range of a Lithium Manganese Dioxide cell. Empty fields with given unit volt for 0% SoC and 100% SoC were available.

Generally, voltage for this type of cells ranges from 2.5 V or 3.0 V to 4.2 V – depending on the manufacturer. One point each were awarded for minimum Voltage between 2.4 V and 3.1 V, and maximum voltage between 4.0 V and 4.3 V.

Question Qu-B-7

Text/state or calculate value – 2 points, increment 1
Learning objective: “Battery Behaviour”, Subtopic “Battery Voltage”.

State the approximate voltage range of a Lithium Iron Phosphate cell. Empty fields with given unit volt for 0% SoC and 100% SoC were available.

Generally, voltage for this type of cells ranges from 2.0 V to 3.7 V – depending on the manufacturer. One point each were awarded for minimum Voltage between 1.9 V and 2.2 V, and maximum voltage between 3.5 V and 3.8 V.

Question Qu-B-8

Single-choice – 1 point
Learning objective: "Experimental Setup", Subtopic "State of Charge".

The SoC (State of Charge) is stated in percent. Which physical dimension is this dependant on? (Which size is stated in percent?).

◯ *Voltage (V)*
◯ *Current(A)*
⊗ *Charge (As)*
◯ *Impedance (Ohm)*
◯ *Energy (Joule)*

Question Qu-B-9

Multiple-choice – 2 points, increment 1
Learning objective: "Experimental Setup", Subtopic "Battery Experimental Setup – Charging Method".

A cells data sheet states a charge end voltage of 4.2 V and a capacity of 500 mAh. Which (set) parameters of a laboratory power-supply or a battery test system have to be set to CC-CV charge this cell with?

☒ *4.2 Volt*	☐ *1 Ohm*	☒ *1.5 Ampere*
☐ *10 Volt*	☐ *4 Volt*	☐ *2 Ampere*
☐ *1500 mAh*		

Question Qu-B-10

Text/state or calculate value – 1 point
Learning objective: "Experimental Setup", Subtopic "Battery Experimental Setup – Calculate Charging Duration".

How long does the charging process described in Qu-B-9 approximately last?

The point was awarded for all answers between 20 minutes and 1 hour.

Question Qu-B-11

Text/state reason or equation – 2 points, increment 1
Learning objective: "Battery Parameters", Subtopic "General Experimental Setups – Measurement methods".

Why is the measurement of the capacity of a cell afflicted with such a great error? Which dimension has to be measured especially precisely?

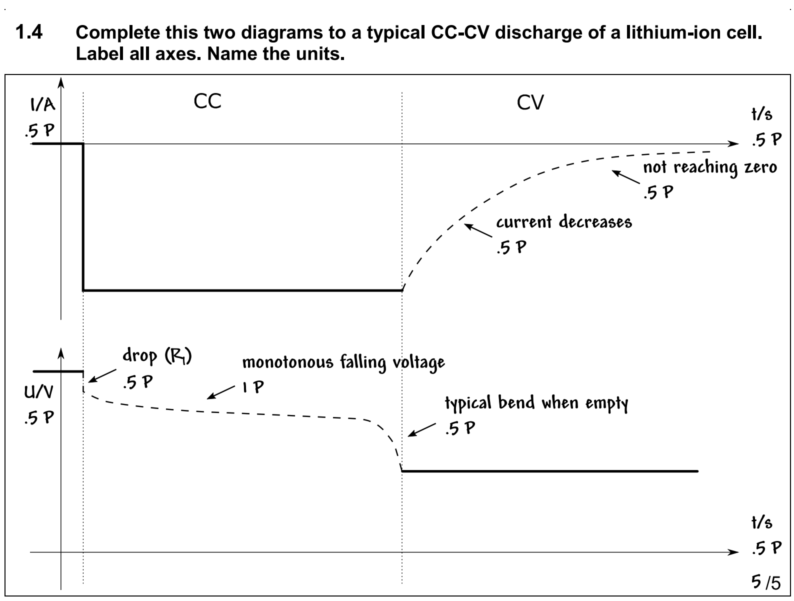

Figure H.5: Example solution for Question Qu-B-12

One point for any statement regarding current or charge measurement. Another point in case the participant had written something about the integration of the error.

Question Qu-B-12

Draw/Graph – 5 points, increment .5
Learning objective: “Battery Parameters”, Subtopic “Battery Voltage and Current”.

Complete these two diagrams (see Figure H.5) to a typical CC-CV discharge of a lithium-ion cell. Label all axes.

One point for a monotonous falling voltage in the CC phase. A half point each for the labels and units on the axes, the voltage drop caused by the activation of discharge current, the typical bend/higher slope of voltage when reaching low SoC, a decreasing current in the CV phase, and if the current never reaches zero ampere.

Question Qu-B-13

Multiple-choice – 1 point
Learning objective: “Experimental Setup”, Subtopic “General Experimental Setup – Galvanostat”. *A galvanostat ...*

☐ *supplies any circuit with constant voltage.*
☒ *injects a certain current into the test object*
☐ *always operates with a refrigerant*

Question Qu-B-14

Text/state reason or equation – 1 point
Learning objective: “Experimental Setup”, Subtopic “Battery Experimental Setup – Open Circuit Voltage”.

You have discharged a completely CC-CV charged cell with a c-rate of 1C, *recharged it using CC-CV, then discharged again with 3C. You have recorded voltage and current. Describe a procedure for approximately determining the OCV(SoC) characteristic curve from the measurement data.*

Expected answer along these lines: ‘the open-circuit-voltage is found in the range between the recorded voltages while discharging and charging, in case the current-integral (= charge) is represented on the x-axis.’

H.1.4 Questions regarding experiments C1 and C2

The following tasks address the learning objectives from lessons C1 and C2 (section G.7 and section G.8).

Question Qu-C-1

Single-choice – 1 point
Learning objective: “Battery Behaviour”, Subtopic “Battery Efficiency – Internal Resistance”.

The efficiency of a charge and discharge cycle with a constant current is mostly influenced by ...
◯ *Double layer capacitance*
◯ *Thermal capacitance*
⊗ *Internal resistance*
◯ *Power density*

Question Qu-C-2

Draw/Graph – 2 points, increment .5
Learning objective: “Experimental Setup”, Subtopic “Battery Experimental Setup – Measurement methods”.

This two point version was used in German runs, while Qu-C-13 was used in the international study runs’ tests. The tasks were (besides grading system) identical.

Question Qu-C-13
Draw/Graph – 4 points, increment .5
Learning objective: “Experimental Setup”, Subtopic “Battery Experimental Setup – Measurement methods”.

Draw a chart which shows a current curve which makes it possible to determine the DC internal resistance of a battery cell. Try to follow the ISO standard used in the laboratory. Label both axis. Name the units.

Figure H.6 represents the example solution.

Points were awarded for any current profile which allows for the determination of internal resistances. Using the taught ISO standard profile was optional. Half points were given for correct labels and units on the axes.

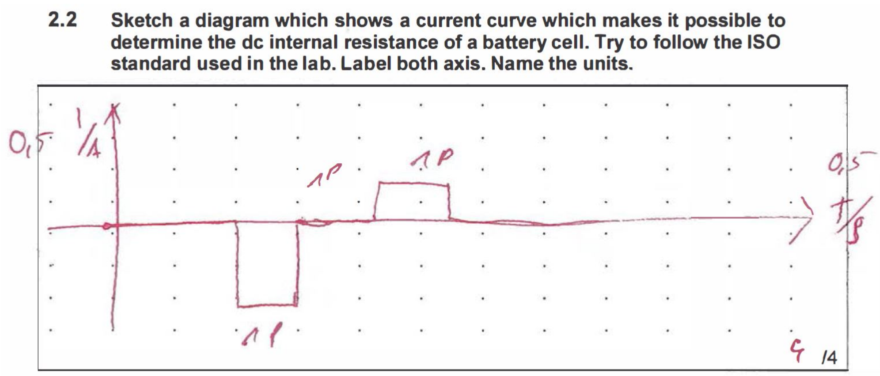

Figure H.6: Example solution for Qu-C-2/13

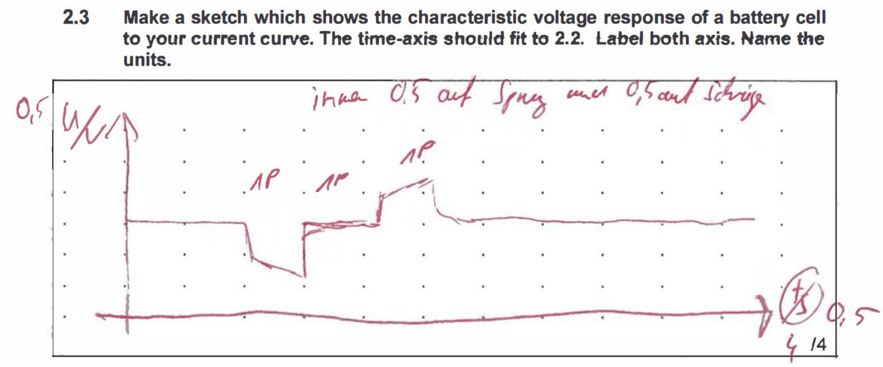

Figure H.7: Example solution for Qu-C-3/14

Question Qu-C-3

Draw/Graph – 2 points, increment .5

Learning objective: "Battery Behaviour", Subtopic "Voltage".

This two point version was used in German runs, while Qu-C-14 was used in the international study runs' tests. The tasks were (besides grading system) identical.

Question Qu-C-14

Draw/Graph – 4 points, increment .5

Learning objective: "Battery Behaviour", Subtopic "Voltage".

Draw a chart which shows the characteristic voltage response of a battery cell to your current curve. The time-axis should fit Question Qu-C-2/13. Label both axes. Name the units.

Points were given if the voltage response met the proposed current profile of Question Qu-C-2. Half points were provided for correct labels and units on the axes.

Example solution in Figure H.7.

Question Qu-C-4

Single-choice – 1 point

Learning objective: "Battery Behaviour", Subtopic "Temperature dependency of the internal resistance".

In case the temperature decreases below room temperature, ...

⊗ *the internal resistance increases*

◯ *the internal resistance decreases*

Question Qu-C-5

Single-choice – 1 point
Learning objective: "Battery Behaviour", Subtopic "Temperature dependency of the internal resistance".

In case the temperature increases above room temperature, ...

◯ *the internal resistance increases*

⊗ *the internal resistance decreases*

Question Qu-C-6

Text/state or calculate value – 1 point
Learning objective: "Battery Behaviour", Subtopic "Internal Resistance".

State the approximate internal resistance of a XYZ mAh lithium-ion cell at room temperature (including unit). XYZ was replaced by the actual capacity of the cell used in the experiments.

As internal resistance varies a lot with SoC and temperature, and students may used cells of different age, deviations from the ideal value were tolerated (33% to 300%).

Question Qu-C-7

Single-choice – 1 point
Learning objective: "Battery Behaviour", Subtopic "Current dependency of the internal resistance".

With negative temperatures (approx. -15 degree Celsius), ...

◯ *the internal resistance increases with higher current*

⊗ *the internal resistance decreases with higher current.*

Question Qu-C-8

Text/state or calculate value – 1 point
Learning objective: "Battery Behaviour", Subtopic "spec. heat capacity".

The specific thermal capacitance of lithium-ion cell is approximately:
______ *J/(g K)*

The right solution is approximately 1 J/(g K). A point was awarded for answers between .1 and 10 J/(g K).

Question Qu-C-9

Multiple-choice – 2 points, increment 1
Learning objective: "Battery Behaviour", Subtopic "Maximum available power".

The maximum output power of a lithium-ion cell is specified or restricted by ...

☐ *Exceeding the maximum cell voltage*
☐ *1 Ohm*
☒ *If clamp voltage falls below the minimal cell voltage*
☒ *The internal resistance*
☐ *2 Ampere*
☐ *The maximum output power is always available whenever the load's resistance is equal to the internal resistance of the cell.*

Question Qu-C-10

Text/state or calculate value – 1 points, increment .5
Learning objective: "Battery Behaviour", Subtopic "Maximum available power".

The maximum output power of a XYZ mAh Lithium-Ion cell at room temperature is approximately: __________ W
XYZ above was replaced by the capacity of the cell used in the specific experiment.

As maximal withdrawable power is influenced by SoC and the age of the cells, a span of a third and factor three of the desired value was awarded a point.

Question Qu-C-11

Single-choice – 1 point
Learning objective: "Battery Behaviour", Subtopic "Maximum available power".

With increasing SoC ...
⊗ *the possible maximum output power increases*
◯ *the possible maximum output power decreases.*

Question Qu-C-12

Single-choice – 1 point
Learning objective: "Battery Behaviour", Subtopic "Maximum available power".

With increasing requested pulse duration (pulse with constant current) ...
◯ *the possible maximum output power during this pulse increases*
⊗ *the possible maximum output power during this pulse decreases*

Question Qu-C-15

Draw/Graph – 3 points, increment .5
Learning objective: "Battery Behaviour", Subtopic "Temperature dependency of the internal Resistance".

Draw the characteristic tendency of the internal resistance as a function of temperature (Internal Resistance ordinate, T abscissa). Assume a constant C-rate and a constant SoC. Label both axes with the respective unit. Specify some points on the X axis.

Example solution in Figure H.8.

One point for a monotonous falling internal resistance with increasing temperature, the second point for a noticeable increase in the slope at negative temperatures,

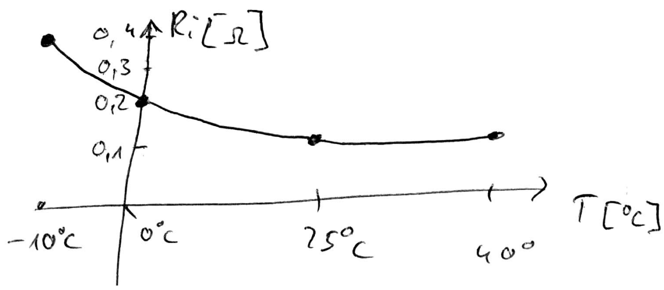

Figure H.8: Example solution for Qu-C-15

and the third point for correct names and units at the axis.

Question Qu-C-16

Text/state reason or equation – 2 points, increment .5
Learning objective: "Experimental Setup", Subtopic "Battery Experimental Setup – Maximum Power determination".

On the back of this paper, describe the procedure for determining the possible maximum discharge power of a battery cell. (Text or flow chart, sketch, etc.)

One point for demonstrating knowledge about the shrinking voltage during the discharging current pulse, and that voltage should not drop below the minimum voltage stated in the cell data sheet. The second point for planning an iterative test procedure to determine the maximum available power.

Question Qu-C-17

Single-choice – 1 point
Learning objective: "Battery Behaviour", Subtopic "Maximum Power vs. Temperature".

With decreasing the cell temperature below room temperature ...

◯ *the available maximum power increases*
⊗ *the available maximum power decreases*

Question Qu-C-18

Single-choice – 1 point
Learning objective: "Battery Behaviour", Subtopic "Maximum Power vs. Internal Resistance".

Cells with higher internal resistance ...

◯ *tend to deliver a higher possible maximum output power*
⊗ *tend to deliver a lower possible maximum output power*

Question Qu-C-19

Single-choice – 1 point
Learning objective: "Battery Behaviour", Subtopic "Maximum Power vs. Temperature".

When increasing the cell temperature above room temperature ...

⊗ *the available maximum power increases*

◯ *the available maximum power decreases*

H.1.5 Questions regarding Lesson D

The following tasks address the learning objectives from lesson D (section G.9).

Question Qu-D-1

Text/state reason or equation – 2 points, increment 1
Learning objective: "Battery Parameters", Subtopic "Peukert". *How can one calculate the available capacity of an accumulator in dependence of the discharge current? State the equation (Peukert's equation).*

The first point was awarded for all formulas which have the right form of exponent. The second point was awarded if the formula was fully correct.

Question Qu-D-2

Multiple-choice – 1 point, increment .5
Learning objective: "Battery Behaviour", Subtopic "Usable Capacity".

Why is it not possible to take the full amount of charge from a fully charged accumulator during practical operation? Four options were offered: ☐ *The accumulator gets too hot, so it is not suitable for practical applications*

☒ *The discharge current has to approach zero at the end of discharge, so it is not suitable for practical applications*

☒ *The cell clamp voltage drops to a critical state before SoC=0% when the accumulator is in practical application*

☐ *Current is not stable enough at the end of the discharge process, so it is not suitable for practical applications.*

In runs after R6, a fifth option was offered:

☐ *All answers do not apply, as the full charge can be withdrawn in practical operation.*

Question Qu-D-3

Single-choice – 1 point
Learning objective: "Battery Parameters", Subtopic "Peukert".

What is the value of the Peukert-Exponent of an ideal accumulator?

◯ *1.05*

⊗ *1*

◯ *1 As/A*

◯ *0.95 As*

◯ *It is exactly the value of the nominal capacity of the cell, units are Ah*

Question Qu-D-4

Text/state reason or equation – 1 point
Learning objective: "Battery Parameters", Subtopic "Peukert-Effect".

What is the definition for a fully discharged battery according to Peukert (name the turn-off condition)?

"The turn-off condition according to Peukert is reached when the battery (under constant discharge current) is going below the minimum voltage according to the data sheet", or similar statements regarding the cell clamp voltage.

Question Qu-D-5

Multiple-choice – 1 point
Learning objective: "Battery Behaviour", Subtopic "Energy Efficiency". *Why is it not possible that the energy efficiency ratio η (fully charge/discharge) will reach $\eta = 1$?*

☐ *You can never get the same amount of charge out of the accumulator that you have charged in.*
☒ *The cell's clamp voltage at a specific SoC during charge and discharge processes differs.*
☐ *It is possible to reach $\eta \geq 1$ if the charge process is slow enough, because η increases with a decreasing discharging current.*

Question Qu-D-6

Multiple-choice – 1 point
Learning objective: "Battery Parameters", Subtopic "SoC".

A SoC value of 50% of an energy storage states that...

☒ *The charge of the cell is at the half of total*
☐ *The energy of the cell is at the half of total*
☐ *The voltage of the cell is at the half of total*
☐ *The power of the cell is at the half of total*
☐ *It is possible to use the battery cell from SoC = 50% to 0% just as it was used from SoC=100% to 50% in a real use case*

Question Qu-D-7

Multiple-choice – 2 points, increment 1
Learning objective: "Battery Behaviour", Subtopic "Usable Capacity".

What are the influences on the usable total capacity of a cell (constant current, fully discharge following Peukert's definition?)

☒ *Temperature during discharge*
☒ *Requested power*
☒ *Internal resistance of the cell*
☐ *Inductivity of the wires connected with the measuring instrument*
☒ *Discharge duration*

☒ *Current during discharge*

Question Qu-D-8

Single-choice – 1 point
Learning objective: “Battery Behaviour”, Subtopic “Efficiency”.

Energy efficiency at a constant current is manly influenced by ...

◯ *Double-layer capacitance*
◯ *Temperature capacitance*
⊗ *Internal resistance*
◯ *Power density*

Question Qu-D-9

Single-choice – 1 point
Learning objective: “Battery Behaviour”, Subtopic “Transfer: Temp – Internal Resistance – Efficiency”.

Transfer: When decreasing cell temperature below room temperature ...

◯ *energy efficiency comparatively increases*
⊗ *energy efficiency comparatively decreases.*

H.1.6 Difficulties

The difficulties of the individual tasks were calculated, the result is shown in Table H.5. The task difficulties were also analysed for statistical differences dependent upon the learning objectives and item formats, please see subsubsection 5.2.1.4.

Overall difficulties of item formats
Differences between the difficulties in answering the item formats were found, the results are presented in Table H.4.

Table H.4: Item formats absolute difficulties, first phase

Item format	M
Single Choice	74%
Text/state or calculate value	56%
Draw/Chart	55%
Draw/Sketch to explain	45%
Multiple Choice	40%
Text/state reason or equation	31%

Note. M = Mean score,
these statistics are based on the percentage score of points achieved,
not on (normalised) student performance.

Table H.5: Difficulties of the tasks, Additionally grouped for both modes, first research phase, R1 and R2

				hands-on			simulations				
Task	M	SD	N	M	SD	N	M	SD	Mann-Whitney		
A-1	^{s}*20%*	*33%*	32	^{s}*19%*	*30%*	34	^{s}*22%*	*35%*	-.235		.814
A-2	^{s}*82%*	*39%*	32	^{s}*78%*	*42%*	34	^{s}*85%*	*36%*	-.055		.956
A-3	^{s}*80%*	*40%*	32	^{s}*84%*	*37%*	34	^{s}*76%*	*43%*	.749		.454
A-4	^{s}*55%*	*50%*	32	^{s}*63%*	*49%*	34	^{s}*47%*	*51%*	1.250		.211
A-5	^{s}*20%*	*30%*	32	^{s}*19%*	*33%*	34	^{s}*22%*	*28%*	-.836		.403
A-6	^{s}*21%*	*32%*	32	^{s}*33%*	*37%*	34	^{s}*10%*	*21%*	2.716	**	.007
A-7	^{s}*42%*	*40%*	32	^{s}*49%*	*37%*	34	^{s}*35%*	*41%*	1.555		.120
A-8	^{s}*52%*	*49%*	32	^{s}*55%*	*50%*	34	^{s}*50%*	*49%*	.392		.695
A-9	^{s} *9%*	*28%*	32	^{s}*14%*	*34%*	34	^{s} *4%*	*19%*	1.309		.190
A-10	^{s}*38%*	*38%*	32	^{s}*38%*	*38%*	34	^{s}*38%*	*39%*	-.801		.423
B-1	^{s}*97%*	*17%*	35	^{s}*100%*	*0%*	33	^{s}*94%*	*24%*	1.467		.142
B-2	^{s}*62%*	*36%*	35	^{s}*60%*	*41%*	33	^{s}*64%*	*30%*	-.154		.878
B-3	^{s}*77%*	*43%*	20	^{s}*70%*	*47%*	19	^{s}*84%*	*37%*	-1.039		.299
B-5	^{s}*25%*	*40%*	35	^{s}*26%*	*42%*	33	^{s}*23%*	*40%*	.368		.713
B-6	^{s}*59%*	*40%*	35	^{s}*64%*	*41%*	33	^{s}*53%*	*39%*	1.221		.222
B-7	^{s}*53%*	*43%*	35	^{s}*41%*	*45%*	33	^{s}*65%*	*38%*	-2.236	*	.025
B-8	^{s}*79%*	*41%*	35	^{s}*83%*	*38%*	33	^{s}*76%*	*44%*	.718		.473
B-9	^{s}*63%*	*43%*	35	^{s}*79%*	*33%*	33	^{s}*47%*	*47%*	2.865	**	.004
B-10	^{s}*75%*	*44%*	35	^{s}*77%*	*43%*	33	^{s}*73%*	*45%*	.417		.677
B-11	^{s}*36%*	*46%*	35	^{s}*43%*	*47%*	33	^{s}*29%*	*43%*	1.283		.200
B-12	^{s} *9%*	*23%*	15	^{s} *3%*	*13%*	14	^{s}*14%*	*31%*	-1.166		.244
C-1	^{s}*80%*	*41%*	30	^{s}*80%*	*41%*	34	^{s}*79%*	*41%*	.058		.954
C-2	^{s}*62%*	*41%*	30	^{s}*62%*	*39%*	34	^{s}*62%*	*44%*	-.138		.890
C-3	^{s}*45%*	*41%*	30	^{s}*38%*	*36%*	34	^{s}*52%*	*45%*	-1.324		.186
C-4	^{s}*64%*	*48%*	30	^{s}*70%*	*47%*	34	^{s}*59%*	*50%*	.923		.356
C-5	^{s}*56%*	*50%*	30	^{s}*63%*	*49%*	34	^{s}*50%*	*51%*	1.065		.287
C-6	^{s}*59%*	*50%*	30	^{s}*73%*	*45%*	34	^{s}*47%*	*51%*	2.119	*	.034
C-7	^{s}*41%*	*50%*	30	^{s}*57%*	*50%*	34	^{s}*26%*	*45%*	2.435	*	.015
C-8	^{s}*45%*	*50%*	30	^{s}*47%*	*51%*	34	^{s}*44%*	*50%*	.203		.839
C-9	^{s}*34%*	*42%*	30	^{s}*38%*	*41%*	34	^{s}*31%*	*43%*	.873		.383
C-10	^{s}*51%*	*49%*	30	^{s}*58%*	*47%*	34	^{s}*44%*	*50%*	1.148		.251
C-11	^{s}*97%*	*18%*	30	^{s}*97%*	*18%*	34	^{s}*97%*	*17%*	-.089		.929
C-12	^{s}*86%*	*35%*	30	^{s}*87%*	*35%*	34	^{s}*85%*	*36%*	.156		.876
D-1	^{s}*52%*	*39%*	34	^{s}*60%*	*38%*	31	^{s}*44%*	*38%*	1.735	†	.083
D-2	^{s}*32%*	*42%*	34	^{s}*26%*	*41%*	31	^{s}*37%*	*43%*	-1.140		.254
D-3	^{s}*91%*	*29%*	34	^{s}*94%*	*24%*	31	^{s}*87%*	*34%*	.969		.332
D-4	^{s}*25%*	*43%*	34	^{s}*26%*	*45%*	31	^{s}*23%*	*43%*	.361		.718
D-5	^{s}*14%*	*35%*	34	^{s}*12%*	*33%*	31	^{s}*16%*	*37%*	-.505		.614
D-6	^{s}*58%*	*50%*	34	^{s}*71%*	*46%*	31	^{s}*45%*	*51%*	2.062	*	.039
D-7	^{s}*25%*	*33%*	34	^{s}*29%*	*37%*	31	^{s}*21%*	*28%*	.765		.444
D-8	^{s}*66%*	*48%*	34	^{s}*68%*	*47%*	31	^{s}*65%*	*49%*	.264		.791
D-9	^{s}*65%*	*48%*	34	^{s}*62%*	*49%*	31	^{s}*68%*	*48%*	-.499		.617

Note. † = p = .10. Z pos. = more correct answers after hands-on experimenting.
The table reports difficulties of individual tasks, separated for both modes of the first study phase.
The percentages represent the share of possible points achieved by the participants.
^{s} = Shapiro-Wilk p < .05, no normal distribution.

H.2 Tests

All tests regarding taught content areas were combined by choosing tasks described before.

The selected tasks for the German runs are stated in Table H.7. The tests for the hidden simulation research phase, used in the German study runs R6 and R9, are presented in Table H.8. The tests of the shortened English runs can be found in Table H.6.

Table H.6: Tasks employed in R3 – R5 and R7 – R8

Nr. in Test	Question ID
1.1	Qu-B-8
1.2	Qu-B-6
1.3	Qu-B-7
1.4	Qu-B-12
1.5a	Qu-B-9
1.5b	Qu-B-10
2.1	Qu-C-1
2.2	Qu-C-13
2.3	Qu-C-14
2.4	Qu-C-5
2.5	Qu-C-7
2.6	Qu-C-6

The combinations of learning objective category and item formats are presented for study run R2 in Table H.9, for all international runs in Table H.10, and for the German runs in the second research phase (R6, R9) in Table H.11.

Table H.12 presents the number of questions and the average achievable number of points in the individual tests.

Table H.7: Tasks employed in R1 and R2

Test	Nr. in Test	Question ID
A	1	Qu-A-1
	2	Qu-A-2
	3	Qu-A-3
	3b	Qu-A-10
	4	Qu-A-4
	5	Qu-A-5
	6	Qu-A-6
	7	Qu-A-7
	8	Qu-A-8
	9	Qu-A-9
	1	Qu-B-1
B	2	Qu-B-2
	3	Qu-B-3 *
	4	Qu-B-4 **
	5	Qu-B-5
	6	Qu-B-6
	7	Qu-B-7
	8	Qu-B-8
	9	Qu-B-9
	10	Qu-B-10
	11	Qu-B-11
C	1	Qu-C-1
	2	Qu-C-2
	3	Qu-C-3
	4	Qu-C-4
	5	Qu-C-5
	6	Qu-C-6
	7	Qu-C-7
	8	Qu-C-8
	9	Qu-C-9
	10	Qu-C-10
	11	Qu-C-11
	12	Qu-C-12
D	1	Qu-D-1
	2	Qu-D-2
	3	Qu-D-3
	4	Qu-D-4
	5	Qu-D-5
	6	Qu-D-6
	7	Qu-D-7
	8	Qu-D-8
	9	Qu-D-9

* Not evaluated for mode comparison in R2
** Not evaluated for mode comparison

Table H.8: Tasks employed in R6 and R9

Test	Nr. in Test	Question ID
A	1	Qu-A-1 *
	2	Qu-A-2 *
	3	Qu-A-3 *
	3b	Qu-A-10 *
	4	Qu-A-4 *
	5	Qu-A-5 *
	6	Qu-A-6 *
	7	Qu-A-7 *
	8	Qu-A-8 *
	9	Qu-A-9 *
B	1	Qu-B-1
	2	Qu-B-2
	3	Qu-B-13
	4	Qu-B-14
	5	Qu-B-5
	6	Qu-B-6
	7	Qu-B-7 *
	8	Qu-B-8
	9	Qu-B-9
	10	Qu-B-10
	11	Qu-B-11
C1	1	Qu-C-1
	2	Qu-C-2
	3	Qu-C-3
	4	Qu-C-4
	5	Qu-C-5
	6	Qu-C-7
	7	Qu-C-15
	8	Qu-C-6
C2	1	Qu-C-9
	2	Qu-C-16
	3	Qu-C-10
	4	Qu-C-11
	5	Qu-C-17
	6	Qu-C-18
	7	Qu-C-19
	8	Qu-C-8
	9	Qu-C-12
D	1	Qu-D-1
	2	Qu-D-2 **
	3	Qu-D-3
	4	Qu-D-4
	5	Qu-D-5
	6	Qu-D-6
	7	Qu-D-7
	8	Qu-D-8
	9	Qu-D-9
	10	Qu-B-12

* not evaluated for mode comparison
** used with five answer options

Table H.9: Matrix learning objective category and item format, R2

	DG	DS	MC	SC	TV	TR	Sum
Battery Behaviour	1		4	8	5		18
Battery Parameters		1	1	2		2	6
Battery System Design			1	3		1	5
Experimental Setup	1	3	5		1	1	11
Sum	2	4	11	13	6	4	40

Table H.10: Matrix learning objective category and item format, international runs

	DG	MC	SC	TV	Sum
Battery Behaviour	2		3	3	8
Battery Parameters			1		1
Experimental Setup	1	1		1	3
Sum	3	1	4	4	12

Table H.11: Matrix learning objective category and item format, R6 and R9

	DG	DS	MC	SC	TV	TR	Sum
Battery Behaviour	3		3	11	4		21
Battery Parameters		1	1	2		2	6
Battery System Design			1	3		1	5
Experimental Setup	1	3	6		1	3	14
Sum	4	4	11	16	5	6	46

Table H.12: Tests: Points and number of questions, first research phase

Test	Questions N	Max. points per question	Points sum	achievable points avg. per question
A – R1, R2	10	2	17	1.70
B – R1	10	2	16	1.60
B – R2	9	2	15	1.66
C – R1, R2	12	2	15	1.25
D – R1, R2	9	2	11	1.22
B*, C* – R3-R5	12	5	25	2.08

Appendix I

Statistical methods

The following appendix describes the major statistical methods used to evaluate the collected data in short form.

I.1 Test for normal distribution

Statistics can be separated in parametric and non-parametric statistics. Parametric tests assume that a population can be modelled by a probability function based on parameters gained from a sample.

Using parametric statistical tests requires that certain model assumptions are not violated.

Important parametric tests (e.g., t-tests, and regressions) for this research assume normal distribution [220]. Normal distribution is modelled by two parameters: mean and standard deviation.

The importance of the assumption of normal distribution is under debate. There is evidence that regression models (including the t-test) are in many cases robust against a violation of the normal distribution [221] [222].

The Shapiro-Wilk test

The Shapiro-Wilk test is a test of normality. More precisely, the *null* hypothesis states that the sample derives from a normally distributed population.

In case the Shapiro-Wilk test's p-value (the probability of finding the respective sample if the null hypothesis is true) is smaller than the selected alpha level (e.g. $p < .05$), the null hypothesis is rejected, which means the population is likely *not* normally distributed.

In case the p-value is bigger than the selected alpha level (e.g. $p > .05$), the null hypotheses is accepted and the population can be assumed to be normally distributed. [129]

I.2 Statistical hypothesis testing / Comparing groups for statistically significant differences

Hypothesis testing can be grouped into:

- Independent samples tests to compare the means of two groups
- Paired sample tests to compare means from the same group (e.g, answers formulated identically but regarding different learning modes)
- One sample tests to compare the mean of a single group with a known mean.

I.2.1 Independent samples test (unpaired samples test)

The independent samples test can be used to compare two sets of independent samples drawn from two populations. In this study for example, comparing the effect of the learning modes on student learning, with 25 participants assigned to learning mode A and 25 assigned to learning mode B (control group). Here, two independent samples of test results are collected and their respective means are compared using the unpaired form of the statistical test.

Parametric: Independent-samples *t*-test (Student's *t*-test)

The independent samples *t*-test (Student's *t*-test) is used to compare two sets of independent *and identically distributed* samples (parametric test).

$$t = \frac{\bar{x_1} - \bar{x_2}}{SE_{\bar{x_1}-\bar{x_2}}} \quad [223, \text{p. } 38] \tag{I.1}$$

with

t = t-statistic for a independent-sample t-test
$\bar{x}_i$ = sample mean
$SE_{\bar{x_1}-\bar{x_2}}$ = standard error of the difference of means

and

$$SE_{\bar{x_1}-\bar{x_2}} = \hat{\sigma}_{\text{pooled}} \cdot \sqrt{1/n_1 + 1/n_2} \tag{I.2}$$

with

$\hat{\sigma}_{\text{pooled}}$ = pooled standard deviation of group data
n_i = sample size

The degrees of freedom are $df = n_1 + n_2 - 2$ [223, p. 39].

Influence of variances of the compared groups

While the Student's t-test assumes that the variances of the two compared samples are equal, Welch's t-test does not:

$$SE_{\bar{x_1}-\bar{x_2}} = \sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}} \quad [223, \text{p. } 37] \tag{I.3}$$

with

$\hat{\sigma}_i^2$ = variance of individual group data
n_i = sample size

However, if the two samples being compared are very big or have equal sizes, Student's original t-test is highly robust when the assumption of equal variances is violated [224].

Effect sizes
Statistical significance indicates whether an outcome is the result of mere chance or not, taking into account a residual risk. Nevertheless, not every statistically significant result is of practical relevance. [225]

Depending on the sample size and the statistical methods used, minor effects may be statistically significant, even though they are hardly noticeable in reality. In order to assess the practical relevance, there are various effect size measures that help to interpret the effects found. The best known is the effect size d by Cohen [226], which is a measure of the standardised mean difference between two groups. In case of different sample sizes and variances, it is recommended to use the pooled variance as denominator in Equation I.4.

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\hat{\sigma}_{x_1}^2 + \hat{\sigma}_{x_2}^2}{2}}} \quad [223, \text{p. } 48] \tag{I.4}$$

with

d = Cohen's d / effect size
$\bar{x}_i$ = sample mean
$\hat{\sigma}_{x_i}^2$ = sample variance

According to Cohen, a Cohen's d between 0.2 and 0.5 can be considered a small effect, between 0.5 and 0.8 a medium effect and a d greater than 0.8 a strong effect [226].

Non-Parametric: Wilcoxon-Mann-Whitney U-test / Wilcoxon rank-sum test

The WMW U-test is non-parametric and tests the null hypothesis that both samples have been drawn from the same population. If this hypothesis is rejected, it can be assumed that the values from one population tend to be larger or smaller than those from the other population. [220]

The WMW U-test is applied to independent samples.

As the U-test tests the general probability that a value randomly selected from one population is greater or smaller than a value randomly selected from another population, besides the mean, the shape of the distributions influences the result of the test. Thus, when the null hypothesis is rejected, proving a difference in median/location requires exclusion of other parameters of the distributions (e.g. kurtosis, skewness and spread comparison) as possible causes for the rejection. [227]

Instead of the independent samples t-test, the WMW U-test was used when samples were not distributed normally.

I.2.2 Paired sample tests

Paired sample tests compare a sample of matched pairs of variables, or values of a variable that has been measured twice (e.g. before/after a treatment). [223]

- Repeated measurements: The values measured are derived from/regarding the same person, for example when collecting data before and after a treatment.
- Couples: The values measured are from different sources belonging together, for example wife-husband, teacher-student, or siblings.
- Matching: The values measured come from different sources belonging together, for example, due to a comparable value on a third variable (this is not the focus of the study).

[223] [129]

In the study at hand, the same questions were used to evaluate the different learning modes and were answered by the same person. Therefore, answers were derived from the same participant and were compared using a paired samples tests.

Parametric: Paired-samples t-test

The tests' null hypothesis is that the mean value of one sample is equal to the mean value of the other sample ($H_0 : \mu_0 = \mu_1$). If the null hypothesis is rejected, the alternative hypothesis states that the mean values of both samples are unequal ($H_1 : \mu_0 \neq \mu_1$).

$$t = \frac{(\bar{x_1} - \bar{x_2}) - m_0}{\hat{\sigma_d}/\sqrt{n}} \quad [223, p. 62] \tag{I.5}$$

with

t = t-statistic for a paired sample t-test
$\bar{x_g}$ = sample mean of group/condition g
m_0 = hypothesised value, in most cases zero (test for equality)
$\hat{\sigma}_d$ = standard deviation of the differences
n = sample size

and

$$\hat{\sigma}_d = \sqrt{\frac{\sum_{i=1}^{n}((x_{i1} - x_{i2}) - (\bar{x_1} - \bar{x_2}))^2}{n-1}} \quad [223, p. 62] \tag{I.6}$$

with

$x_{i1} - x_{i2}$ = difference of sample i between groups/conditions
$\bar{x_1} - \bar{x_2}$ = difference of the mean values

The degrees of freedom are $df = n - 1$[223, p. 63].

Non-Parametric: Wilcoxon signed-rank test (WSR)

The Wilcoxon signed-rank test (WSR) for *dependent* samples tests whether the central trends of two *dependent* samples are different. The minimal requirement for using a WSR is ordinally scaled data and a symmetrical distribution of the differences of both data sets. [129]

The Wilcoxon test was used when the requirements for a paired sample t-test were not met.

I.2.3 One-sample t-test

One sample tests use the mean value of a sample to check whether the mean value of a population is equal to a specified value (or lower/higher). In the present research for example, one-sample tests were used to find out if the selected answers statistically significantly differed from the neutral values of Likert scales.

The test (see Equation I.7) checks the null hypothesis that the mean value of the sample is equal a set value ($H_0 : \mu = \mu_0$). If the null hypothesis is rejected, the alternative hypothesis is accepted which states that the mean value of the sample is unequal to a set value ($H_1 : \mu \neq \mu_0$).

$$t = \frac{\bar{x} - m_0}{\hat{\sigma}_x / \sqrt{n}} \quad [223, \text{p. } 67] \tag{I.7}$$

with

t $= t$-statistic for a one sample t-test
$\bar{x}$ $=$ sample mean
m_0 $=$ hypothesised value
$\hat{\sigma}_x$ $=$ samples' standard deviation
n $=$ sample size

The degrees of freedom are $df = n - 1$ [223, p. 67].

I.2.4 Cross tab or fourfold test

The chi-square cross tab test serves to investigate whether two dichotomous variables are statistically independent / whether the distribution of a dichotomous variable is identical in two groups. The test is based on a contingency table (2 × 2) that visualises the bi-variate frequency distribution of two characteristics, see Table I.1.

Table I.1: Cross tab or fourfold test

	Variable X		
	Characteristic X_1	Characteristic X_2	Sum
Variable Y			
Characteristic Y_1	a	b	$a+b$
Characteristic Y_2	c	d	$c+d$
Sum	$a+c$	$b+d$	$n = a+b+c+d$

The test variable $\widehat{\chi^2}$ is calculated to test the null hypothesis that both characteristics are statistically independent.

$$\widehat{\chi^2} = \frac{n \cdot (a \cdot d - c \cdot b)^2}{(a+c) \cdot (b+d) \cdot (a+b) \cdot (c+d)} \quad [129] \tag{I.8}$$

The test statistic is approximately $\widehat{\chi^2}$ distributed with one degree of freedom. It should only be used if there are at least six characteristic observations in each of the two samples.

If the test value obtained from the sample is less than the critical value associated with the selected significance level, the test can not prove that a significant difference exists. A test value greater than or equal to the critical value (Table I.2) shows a significant difference between the samples.

Table I.2: Chi-square distribution: (1-α) for one degree of freedom

Significance level	.900	.950	.975	.990	.995	.999
Critical value	2.71	3.84	5.02	6.63	7.88	10.83

I.3 Testing scale reliability

I.3.1 Principal Component Analysis

Principal component analysis (PCA) is used to detect underlying dimensions of a data set. The goal of the explorative method PCA is to extract linearly uncorrelated variables – so called principal components – from a complex data set with multiple variables. The amount of variance each principal component is able to explain is in descending order. The data should be reduced in a way that minimises information loss, while preserving the model's predictive quality. In addition, it is important to figure out how high each variable loads on each principal component. These loadings are a kind of weight: the greater its value, the more the variable has in common with the principal component. [129, 132]

In the presented research, PCA was applied to reduce the number of items needed to determine the Amount of Practical Experience in four sub-dimensions (see section 5.3).

I.3.2 Cronbach's alpha / scale reliability

Cronbach's α is an estimate of the reliability of a psychometric test which is based on n items. Its calculation assumes that all factor loadings are equal [228].

Coefficient α represents the average of all correlations that arise when the indicators of a construct are divided into two halves in any possible combination and the sums of the samples of the resulting halves are correlated [229, p. 8]. In other words it can be understood as the correlation of two tests that measure the same construct [230].

$$D = I_1 + I_2 + \cdots + I_n \tag{I.9}$$

$$\alpha = \frac{n}{n-1}\left(1 - \frac{\sum_{i=1}^{n}\sigma_{I_i}^2}{\sigma_D^2}\right) \quad [228, p. 299] \tag{I.10}$$

with

D = test score/dimension
(sum of all components/items with identical factor loadings)
I_i = items
α = Cronbach's alpha / scale reliability
n = number of items
$\sigma^2_{I_i}$ = variance of component/item/scale i for the observed sample
σ^2_D = variance of the observed total test scores

Cronbach's α can take on values from negative infinity (Coefficient alpha is negative in case the within-subject variability is greater than the between-subject variability.) to positive 1, but only positive values are of interest when it comes to the interpretation of a scales reliability. Acceptable values vary, values $\geqq .7$ are usually considered acceptable [133, p. 153], [231, p. 1279].

I.4 Correlation

I.4.1 Parametric: Pearson's *r*

The Pearson product-moment correlation coefficient (PPMCC) is a measure of the linear correlation between two variables (the covariance of the two variables divided by the product of their standard deviations, see Equation I.11).

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad \text{[129, p. 170], [223, p. 85]} \tag{I.11}$$

with

r_{xy} = Pearson's r
x_i = individual sample point
y_i = individual sample point
$\bar{x}$ = mean of x
$\bar{y}$ = mean of y
n = sample size

r has a value between +1 and -1, where

- 1 is total positive linear correlation (data-points in x-y plot exactly on a line),
- 0 equals no linear correlation,
- and -1 is total negative linear correlation (data-points in x-y plot exactly on a line). [223, p. 85]

I.4.2 Non-Parametric: Spearman's ρ

Spearman's rank correlation coefficient describes how well the relationship between two variables can be described by a (rising or falling) monotonic function. Unlike Pearson's r, Spearman's ρ is a non-parametric measure. As it is based on ranks, Spearman's ρ can describe the relationship between continuous as well as discrete ordinal variables. [223, p. 87]

Calculating the Spearman correlation requires the computation of a Pearson correlation between the ranked variables (see Equation I.11). In case the same value was recorded multiple times (which would lead to the same rank), the respective data sets are all attributed the mean of ranks of the group of identical values.

I.4.3 Testing correlations for statistically significant differences between groups

The correlation coefficient r can range from -1 to 1, which means it is strictly limited and not normally distributed (consider the small probability of $r = 1$). This fact makes statistical methods – such as calculating the confidence interval – difficult, particularly when the considered correlation coefficient approaches +1 or -1. The Fisher transformation (r to z') converts correlation coefficients into approximately asymptotically normally distributed values and therefore allows the application of test methods that require such a distribution.

$$z' = .5 \cdot \ln\left(\frac{1+r}{1-r}\right) \quad \text{[223, p. 87], [232, 233, p. 198]} \tag{I.12}$$

with

$z' =$ Z-score
$r =$ correlation coefficient

Calculating the z-score to determine the difference of both correlations leads to

$$Z_{\text{difference}} = \frac{z_1' - z_2'}{\sqrt{\frac{1}{n_1-3} - \frac{1}{n_2-3}}} \quad \text{[129, p. 191]} \tag{I.13}$$

with

$Z_{\text{difference}} =$ Z-score
$z_i =$ sample Z-score
$n_i =$ sample sizes

The Fisher transformation can also be used for comparing Spearman's rank correlation coefficients [234, 235].

I.4.4 Statistical comparison of the slopes of two regression lines

Spearman's ρ and Pearson's r standardise the scales (division by standard deviation (leads to β slopes) or by the used ranks). Comparing the slopes of two regression lines (B slopes) based on data of two groups in the same coordinate system makes it possible to detect statistical differences in the strength of the correlation (abscissa-ordinate) between the groups. [223, p. 103] [236] An independent t-test (Equation I.1, Equation I.3) can be used to compare the slopes for statistical difference.

$$t = \frac{B_1 - B_2}{\sqrt{{SE_{B_1}}^2 + {SE_{B_2}}^2}} \tag{I.14}$$

with

t = t-value
B_i = fitted slope of data group i
SE_{B_i} = standard error of the slopes

I.5 Studentisation / Z-variate

In statistics, Studentisation (introduced by the statistician William Sealy Gosset using the pseudonym “Student”) refers to data transformation of a randomly distributed variable in a way that results in values that have the mean = 0 and the empirical variance = 1 (therefore, after the Studentisation, also the standard deviation correspondents = 1).

Among other possibilities, Studentisation enables researchers to compare random variables that are distributed differently.

As shown in Equation I.15 for n realisations of a random variable x_i with arithmetic mean $\bar{x}$, the corresponding values through Studentisation are obtained by subtracting the mean and dividing by the standard deviation of that variable. A variable that was standardised accordingly is also called Z-variate. [118]

$$Z_\mathrm{i} = \frac{x_\mathrm{i} - \bar{x}}{\sqrt{1/n \sum_{k=1}^{n} (x_\mathrm{k} - \bar{x})^2}} \quad [223, p. 17] [237, p. 31] \tag{I.15}$$

with

x_i = (random) values with arithmetic mean value $\bar{x}$
Z_i = Z-variate (normalised values after Studentisation)

Using the normalised values, it is possible to determine whether an associated original value differs (above/below in units of standard deviations) from the overall mean.

This method was used to evaluate students’ test performance (calculated through Studentisation based on their test score percentage relative to their peers writing the same test in the same study run).