THE EFFECT OF USING SPACED REPETITION IN MOBILE LEARNING GAMES ON THE LEARNING SUCCESS

Florian Schimanke

The Effect of using Spaced Repetition in Mobile Learning Games on the Learning Success

Florian Schimanke

Submitted in partial fulfilment of the requirements for the degree of
Doctor of Natural Sciences (Dr. rer. nat.)
of the University of Bamberg,
Faculty Information Systems and Applied Computer Sciences

Advisor and Reviewer:
Prof. Dr. Ute Schmid

External Reviewer:
Prof. Dr. Robert Mertens, Berufliche Hochschule Hamburg

Further Committee Member:
Prof. Dr. Andreas Henrich

Submitted: March 01, 2022
Day of Defense: July 14, 2022

URN: urn:nbn:de:bvb:473-irb-553175
DOI: https://doi.org/10.20378/irb-55317

Abstract

The field of learning has been researched in multiple ways and directions. The different learning strategies of individual learners represent one of those fields. Especially in traditional learning environments, such as school or university, learners tend to result-driven learning strategies like massing and cramming, which foremost aim at using the short-term memory to have the necessary information available during a test or an exam. An approach that focuses more on the long-term is called "spacing" and uses sophistically determined intervals between learning sessions that are unique to each learner and are proven to have a positive effect on the sustainable retention of knowledge. However, since spaced repetition is still a repetition-based learning strategy, it is prone to bore the learners by requiring them to repeat the same action over and over again.

One approach to take the boredom out of repetition-based learning is gamification or game-based learning. Using learning games can lead to a higher motivation and thus to a higher probability that the learners continue to use the game and keep on learning through it. Current technologies like smartphones and tablets are capable of handling sophisticated learning games, offer the possibility to use them ubiquitously and therefore act as a facilitator for a learning on demand.

While spaced repetition is not a new field of research, the effect of this learning approach on the learning success has not yet been researched when being used in combination with a mobile learning game. Combining the motivating effect of ubiquitous game-based learning with spaced repetition seems to be a promising combination for an improved learning success. The aim of the presented research is to investigate if using spaced repetition in a mobile learning game leads indeed to the desired success.

To do so, several aspects were considered, including how the SM2 spaced repetition algorithm can be integrated into a mobile learning game, the negative effect that too much motivation can have, how to deal with early and late repetitions, how to conclude automated tests of the implementation, how to collect and filter real-world data to analyze the effectiveness, and how learning analytics can be applied to this analysis. For the data collection, an algorithmic framework, containing the SM2 and the self-developed FS algorithm was integrated into an established geography learning game. The collected data was then filtered and analyzed in accordance with an established definition for learning success. For this, a test group with players who played at least one category from the learning game strictly according to the spaced repetition approach and

a control group with players who played the game at more or less random times were built from the collected data. Their progress throughout the observation period was compared and a conclusion was drawn from this.

The two main indicators for the effect on the learning success were defined as the “Average Base Knowledge” (ABK) and the “Average Retention Score” (ARS). The ABK is the average number of correct answers during the first game in both groups, while the ARS is the average correctness during the so-called "leaerning-games", which are the games from the 3^{rd} to the n^{th} game. While both groups started with a comparable ABK of 0.5, the test group already achieved an almost perfect ARS for the fourth game (i.e. the third repetition) of 0.95. The control group on the other hand reached an ARS of 0.63 for the same repetition without the interval calculation by the SM2 algorithm. This effect continues during the following repetitions, during which the ARS for the test group stayed between 0.98 and 1.0, while the control group reaches a maximum ARS of 0.87 after 14 games, which marked the end of the observation period. Further structure exploratory analyses, such as a cluster analysis and a logistic regression confirmed the findings from the comparison of the development of the ABK and the ARS.

Summarized, the analyses showed that there is a considerable difference between the two groups. Foremost, while the retention improved in both groups, the test group achieved a quicker improvement and a lasting retention of knowledge over time, which is matches exactly the definition that was established for learning success in the presented research.

Zusammenfassung

Die Tätigkeit des Lernens wurde von der Wissenschaft bereits in den verschiedensten Richtungen untersucht. Einer der dabei betrachteten Bereiche ist der der Lernstrategien von verschiedenen Lernern. Vor allem in traditionellen Lernumgebungen wie der Schule oder der Universität setzen viele Lerner auf eine zielgetriebene Lernstrategie, bei der das vorrangige Ziel darin besteht, möglichst viele Informationen im Kurzzeitgedächtnis zu speichern, um es während einer Prüfung parat zu haben, diese zu bestehen und im Idealfall ein gutes Ergebnis zu erzielen. Eine Lernstrategie, die auf einen nachhaltigeren Ansatz abzielt, ist das sogenannte "spacing", bei dem Lerninhalte zu nach psychologischen Modellen und für jeden Lerner individuell berechneten Intervallen wiederholt werden, um einen Langzeiteffekt zu erzielen. Da jedoch auch dieser Ansatz auf dem Wiederholen von Inhalten basiert, ist er anfällig dafür, den Lerner mit der Zeit zu langweilen.

Ein Ansatz, dieser Langeweile entgegenzuwirken, ist die "Gamifizierung" oder das spielebasierte Lernen. Die Verwendung von Lernspielen kann zu einer höheren Motivation bei den Lernern führen, das Spiel dauerhaft zu nutzen und auf diese Weise zu lernen. Aktuelle Technologien wie Smartphones und Tablets sind in der Lage, selbst aufwändige Lernspiele zu jeder Zeit und an jedem Ort verfügbar zu machen. Hierdurch wird der Lerner in die Lage versetzt, Lernspiele in allen möglichen Situationen zu nutzen, was den aktuellen Trend zum „learning on demand“ unterstützt.

Während das Konzept der "Spaced Repetitions" kein neues Forschungsfeld darstellt, wurde bislang noch nicht der Effekt untersucht, den diese Lernstrategie auf den Lernerfolg hat, wenn sie in gemeinsam mit einem mobilen Lernspiel eingesetzt wird. Die Kombination aus der motivierenden Wirkung von überall verfügbaren mobilen Lernspielen und „Spaced Repetition“ stellt einen vielversprechenden Ansatz dar, um einen verbesserten und nachhaltigen Lernerfolg zu erzielen. Das Ziel der vorliegenden Forschungsarbeit ist es zu untersuchen, ob dieser Effekt tatsächlich zu beobachten ist.

Hierzu wurden verschiedene Aspekte berücksichtigt, wie beispielsweise die Frage, ob und wie der SM2 Spaced-Repetition-Algorithmus in ein mobiles Lernspiel integriert werden kann, wie sich zu viel Motivation negativ auswirken kann, wie mit verfrühten und verspäteten Wiederholungen umgegangen wird und wie automatisierte Tests der Implementierung durchgeführt werden können. Um eine Auswertung durchführen zu können, wurde ein Framework, bestehend aus dem SM2 Spaced-Repetition-Algorithmus und dem selbstentwickelten FS-

Algorithmus in ein etabliertes Geografie-Lernspiel integriert. Die auf diese Weise gesammelten Daten wurden anschließend im Hinblick auf eine aufgestellte Definition für den Lernerfolg gefiltert und analysiert. Zu diesem Zweck wurden eine Test- und eine Kontrollgruppe gebildet und ihre Ergebnisse während des Beobachtungszeitraums miteinander verglichen.

Als die beiden wesentlichen Indikatoren für die Auswirkungen auf den Lernerfolg wurden die "Average Base Knowledge" (ABK) und der "Average Retention Score" (ARS) definiert. Die ABK ist dabei die durchschnittliche Anzahl von richtigen Antworten während des ersten Spiels innerhalb der jeweiligen Gruppen, während der ARS die durchschnittliche Korrektheit während der "Lernspiele", also während des 3. und des n. Spiels wiedergibt. Während beide Gruppen mit einer vergleichbaren ABK von 0,5 starteten, erreichten die Mitglieder der Testgruppe einen beinahe perfekten ARS von 0,95 bereits beim vierten Spiel, also der dritten Wiederholung. Bei der Kontrollgruppe hingegen lag der ARS bei derselben Wiederholung ohne die Nutzung der Intervallberechnungen durch den SM2 Algorithmus lediglich bei 0,63. Dieser Effekt kann auch bei den weiteren Wiederholungen beobachtet werden, bei denen sich der ARS bei der Testgruppe zwischen 0,98 und 1,0 bewegte, während die Kontrollgruppe einen maximalen ARS von 0,87 nach 14 Spielen erreichte, was das Ende der Beobachtungszeit markierte. Eine Clusteranalyse und eine logistische Regression als weitere strukturentdeckende Analysen bestätigten die Erkenntnisse aus dem Vergleich der Entwicklungen von ABK und ARS.

Zusammengefasst zeigte die Auswertung der Spieldaten einen deutlichen Unterschied zwischen den beiden Gruppen. Während sich die Qualität der Antworten in beiden Gruppen verbesserte, trat dieser Effekt bei der Testgruppe deutlich schneller ein und blieb anschließend während des Beobachtungszeitraums auf einem hohen Niveau, was exakt der zuvor aufgestellten Definition von Lernerfolg für die vorliegende Forschungsarbeit entspricht.

Acknowledgements

A project like this doctoral thesis, especially when it is done in part-time and as an external doctoral candidate, needs the support from a lot of people and I am very grateful to have received this support throughout the last ten years. Therefore, I would like to thank my parents Bärbel and Helmut, my brother Patrick, and the rest of my family and friends for always keeping my motivation at a high level, supporting and understanding me and helping me through struggles. A big thank you goes also out to Prof. Robert Mertens for all his ideas, for the discussions, the input, the suggestions and the constant feedback.

I am very grateful that my advisor Prof. Ute Schmid accepted me as an external doctoral candidate, patiently answered my questions and always helped me with her advice. Furthermore, I thank Prof. Oliver Vornberger with whom I started this doctoral thesis at the University of Osnabrück before switching to the University of Bamberg.

A big thank you goes also to Prof. Sophie Huck for advising me regarding the data collection and interpretation, to Jonathan Hillebrand for integrating the spaced repetition framework in his successful iOS game, to Leonard Hill for helping to develop the simulation and testing frameworks and to Sophie Ribbers for contributing to the psychology aspects as well as to Stephanie Vollmer, Florian Hallay, and Arkadij Enders who also contributed to my work.

I would like to thank my current and former colleagues from the Hochschule Weserbergland, especially Prof. Hans Ludwig Meyer and Prof. André von Zobeltitz for all their support, feedback, suggestions and so much more.

Contents

List of Figures

List of Tables

Acronyms

ABK	Average Base Knowledge
AGNES	Agglomerative Nesting
ARS	Average Retention Score
FS	Follow-Up Sequence
HSW	Hochschule Weserbergland
KDD	Knowledge Discovery in Databases
SM	SuperMemo
SM11	SuperMemo Version 11
SM2	SuperMemo Version 2
UUID	Universally Unique Identifier

Publications

During the presented research, different challenges and findings were identified and presented at several conferences and workshops. The following publications contain some of the scientific contributions, visualizations, and results as part of this doctoral thesis:

Conferences and Proceedings

- **F. Schimanke**, R. Mertens, O. Vornberger: *What to learn next? Content selection support in mobile game-based learning.* AACE E-Learn World Conference on E-Learning (E-Learn 2013), Las Vegas, NV, USA, 21.-24. October 2013, Proceedings pp. 2503-2512, (https://www.learntechlib.org/p/115267/)
  - My contributions to this paper:
    - Conception of the paper
    - Review of related work
    - Development of prototype language learning game
    - Integration of SM2 into the prototype game
    - Identification of early repetition problem and possible solutions
    - Conception of the FS auxiliary algorithm
  - This paper contributes to section 3.2.1 of this thesis and can be found in the Appendix.

- **F. Schimanke**, R. Mertens, S. Vollmer, O. Vornberger: *Multi Category Content Selection in Spaced Repetition Based Mobile Learning Games.* IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2013), Los Angeles, CA, USA, 09.-11. December 2013, Proceedings pp. 468-473, doi: 10.1109/ISM.2013.90
  - My contributions to this paper:
    - Conception of the paper
    - Review of related work
    - Development of prototype database learning game
    - Categorization of learning content
    - Adaption of the SM2 algorithm to use categorized content
  - This paper contributes to section 3.2.2 of this thesis and can be found in the Appendix.

- **F. Schimanke**, R. Mertens, O. Vornberger: *Designing for Motivation: Design-Considerations for Spaced-Repetition-Based Learning Games on Mobile Devices.* AACE E-Learn World Conference on E-Learning (E-Learn 2014), New Orleans, LA, USA, 27.-30. October 2014, Proceedings pp. 1770-1779, (http://www.learntechlib.org/p/148959/)
  - My contributions to this paper:
    * Conception of the paper
    * Review of related work
    * Identification of the area of conflict of too much motivation and spaced repetition
    * Development of strategies to mitigate the identified area of conflict
    * Conception of design strategies for spaced repetition enhanced mobile learning games
  - This paper contributes to section 3.2.5 of this thesis and can be found in the Appendix.

- **F. Schimanke**, R. Mertens, O. Vornberger: *Architecture Considerations for Spaced Repetition Based Mobile Learning Games on iOS.* IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2014), Taichung, Taiwan, 10.-12. December 2014, Proceedings pp. 363–368, doi: 10.1109/ISM.2014.17
  - My contributions to this paper:
    * Conception of the paper
    * Review of related work
    * Comparison of frameworks and plug-ins
    * Decision to use framework approach
    * Conception of algorithmic spaced repetition framework for integration into external learning games
  - This paper contributes to section 3.3.1 of this thesis and can be found in the Appendix.

- **F. Schimanke**, R. Mertens, O. Vornberger, F. Hallay, A. Enders: *Using a Spaced-Repetition-Based Mobile Learning Game in Database Lectures.* AACE E-Learn World Conference on E-Learning (E-Learn 2015), Kona, HI, USA, 19.-22. October 2015, Proceedings pp. 1610-1619, (https://www.learntechlib.org/p/152208/)
  - My contributions to this paper:
    * Conception of the paper
    * Review of related work
    * Development of prototype database learning game
    * Conception and evaluation of questionnaire
  - This paper contributes to section 3.2.3 of this thesis and can be found in the Appendix.

- **F. Schimanke**, S. Ribbers, R. Mertens, O. Vornberger, S. Vollmer: *Implications of Short Term Memory Research for the Design of Spaced Repetition Based Mobile Learning Games.* IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2015), Miami, FL, USA, 14.-16. December 2015, Proceedings pp. 571-576, doi: 10.1109/ISM.2015.13
  - My contributions to this paper:
    * Conception of the paper
    * Adaption of short-term memory implications on the spacing approach
    * Comparison of strategies to mitigate short-term implications on the spacing approach
  - This paper contributes to section 3.2.6 of this thesis and can be found in the Appendix.

- **F. Schimanke**, R. Mertens, O. Vornberger, J. Hillebrand: *Enhancing Mobile Learning Games with Spaced-Repetition and Content-Selection Algorithms.* AACE E-Learn World Conference on E-Learning (E-Learn 2016), Washington D.C., USA, 14.-16. November 2016, Proceedings pp. 1270-1279, (https://www.learntechlib.org/p/174071/)
  - My contributions to this paper:
    * Conception of the paper
    * Review of related work
    * Collaboration with external developer
    * Conception for replacing the Leitner System with SM2 and FS
    * Conception of interfaces for data exchange
    * Conception of the scheduling of on-screen notifications
    * Conception of the collection of playing data for later analysis
  - This paper contributes to section 3.3.5 of this thesis and can be found in the Appendix.

- **F. Schimanke**, L. Hill, R. Mertens, O. Vornberger: *Simulating Context in Mobile Learning Games for Testing and Debugging.* IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2016), San Jose, CA, USA, 11.-13. December 2016, Proceedings pp. 655-660, doi: 10.1109/ISM.2016.0140
  - My contributions to this paper:
    * Conception of the paper
    * Review of related work
    * Contribution to the conception of the Time Machine
    * Integration of Time Machine and prototype learning games
    * Testing the Time Machine
  - This paper contributes to section 3.3.2 of this thesis and can be found in the Appendix.

- **F. Schimanke**, R. Mertens, U. Schmid: *Spaced Repetition in Mobile Learning Games - A Cure to Bulimic Learning?* AACE E-Learn World Conference on E-Learning (E-Learn 2017), Vancouver, CAN, 17.-20. October 2017, Proceedings pp. 955-964, Outstanding Paper Award, (https://www.learntechlib.org/p/181278/)
  - My contributions to this paper:
    * Conception of the paper
    * Review of related work
    * Applying the implications of short-term memory on the spacing approach
    * Conception, evaluation, and interpretation of student questionnaire
  - This paper contributes to section 3.2.4 of this thesis and can be found in the Appendix.

- **F. Schimanke**, L. Hill, R. Mertens: *A Unit Testing Framework for Context Variant Code in a Mobile Learning App.* IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2017), Taichung, Taiwan, 11.-13. December 2017, Proceedings pp. 577-582, doi: 10.1109/ISM.2017.113
  - My contributions to this paper:
    * Conception of the paper
    * Contribution to the conception of the unit testing framework
    * Integration of the framework into prototype learning games
    * Testing the framework via Time Machine
    * Applying results on different context in learning games
  - This paper contributes to section 3.3.3 of this thesis and can be found in the Appendix.

- **F. Schimanke**, R. Mertens, U. Schmid: *Mobile Game-Based Learning in the App-Age – Where we are and where we want to be.* AACE E-Learn World Conference on E-Learning (E-Learn 2018), Las Vegas, NV; USA, 15.-18. October 2018, Proceedings pp. 1321-1330, (https://www.learntechlib.org/p/185098/)
  - My contributions to this paper:
    * Conception of the paper
    * Review of related work
    * Comparison of development of m-learning, e-learning, and game-based learning over the past decade
    * Conception, evaluation, and interpretation of user questionnaire
  - This paper contributes to section 3.3.4 of this thesis and can be found in the Appendix.

- **F. Schimanke**, B.S. Huck, R. Mertens: *Player Types in Mobile Learning Games – Playing Patterns and Motivation.* IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2018), Taichung, Taiwan, 10.-12. December 2018, Proceedings pp. 247-252, doi: 10.1109/ISM.2018.00035
  - My contributions to this paper:
    - Conception of the paper
    - Review of related work
    - Review of the collected playing data
    - Identification of different player types
    - Definition of different player types
  - This paper contributes to section 3.4.1 of this thesis and can be found in the Appendix.

- **F. Schimanke**, R. Mertens: *Measuring the Learning Success in Spaced-Repetition-Based Mobile Learning-Games: Suitable Techniques and Hurdles.* AACE E-Learn World Conference on E-Learning (E-Learn 2019), New Orleans, LA; USA, 04.-07. November 2019, Proceedings pp. 528-537, (https://www.learntechlib.org/p/211122/)
  - My contributions to this paper:
    - Conception of the paper
    - Review of related work
    - Development of strategy to measure learning success
    - Definition of learning success for the presented research
  - This paper contributes to section 3.4.2 of this thesis and can be found in the Appendix.

- **F. Schimanke**, B.S. Huck, R. Mertens: *Retrieval of Relevant Data for Measuring the Impact of Spaced-Repetition Algorithms on the Learning Success in Mobile Learning Games.* IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2019), San Diego, CA, USA, 09.-11. December 2019, Proceedings pp. 279-284, doi: 10.1109/ISM46123.2019.00063
  - My contributions to this paper:
    - Conception of the paper
    - Review of related work
    - Applying KDD process on the collected playing data
    - Definition of parameters of relevant data sets for later analysis
    - Identification of the relevant data sets
  - This paper contributes to section 3.4.3 of this thesis and can be found in the Appendix.

- **F. Schimanke**, R. Mertens: *Deriving Strategies for the Evaluation of Spaced Repetition Learning in Mobile Learning Applications from Learning Analytics*. IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2020), virtual, 02.-04. December 2020, Proceedings pp. 239-244, doi: 10.1109/ISM.2020.00049
  - My contributions to this paper:
    * Conception of the paper
    * Review of related work
    * Applying findings from learning analytics on the presented research
    * Applying classification scheme of learning outcomes
    * Adaption of findings on the data collection
  - This paper contributes to section 3.4.4 of this thesis and can be found in the Appendix.

- **F. Schimanke**: *The Impact of Spaced Repetition Learning on the Learning Success in Mobile Learning Games*. IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2021), virtual, 29. November - 01. December 2021, Proceedings pp. 275-280, doi: 10.1109/ISM52913.2021.00054
  - My contributions to this paper:
    * Entire idea, conception and writing of the paper
    * Review of related work
    * Data analysis and interpretation
    * Visualization of the findings
    * Conclusion of the impact of spaced repetition learning on the learning success in mobile learning games
  - This paper contributes to section 3.4.5 of this thesis and can be found in the Appendix.

Journals

- **F. Schimanke**, R. Mertens, O. Vornberger: *Spaced-Repetition Learning Games on Mobile Devices – Foundations and Perspectives*. International Journal of Interactive Technology and Smart Education (ITSE); 11(3), pp. 201 – 222, Emerald Group Publishing Limited, 2014, doi: 10.1108/ITSE-07-2014-0017
  - My contributions to this paper:
    * Conception of the paper
    * Review of related work
    * Applying findings from learning analytics on the presented research
    * Applying classification scheme of learning outcomes
    * Adaption of findings on the data collection
  - The full version of this paper can be found in the Appendix.

- **F. Schimanke**, R. Mertens, O. Vornberger: *Designing for Motivation: Design-Considerations for Spaced-Repetition-Based Learning Games on Mobile Devices.* International Journal on E-Learning (IJEL), July 2017, Volume 16, Number 3, pp. 287-311, https://www.learntechlib.org/primary/j/IJEL/v/16/n/3/
  - My contributions to this paper:
    - Entire idea, conception and writing of the paper
    - Review of related work
    - Data analysis and interpretation
    - Visualization of the findings
    - Conclusion of the impact of spaced repetition learning on the learning success in mobile learning games
  - The full version of this paper can be found in the Appendix.

- **F. Schimanke**, R. Mertens, O. Vornberger, L. Hill: *Simulating Context in Mobile Learning Games for Testing and Debugging.* International Journal of Semantic Computing (IJSC), Vol. 11 (2017), No. 01, pp. 21-41, doi: 10.1142/S1793351X17400025
  - My contributions to this paper:
    - Entire idea, conception and writing of the paper
    - Review of related work
    - Data analysis and interpretation
    - Visualization of the findings
    - Conclusion of the impact of spaced repetition learning on the learning success in mobile learning games
  - The full version of this paper can be found in the Appendix.

Part I

Synopsis of the Thesis

Chapter 1

Introduction

1.1 General Classification

Learning in general is a highly individual process [66], which means that everybody tries to find his or her best way of remembering a certain topic in order to have it at hand when it is needed. The latter might happen at a given point in time or occasionally during everyday life. The individual nature of the learning process leads to different strategies that are used by people to reach a certain goal through it. These strategies impact either the short-term or the long-term memory [46]. For the latter, there need to be individual intervals between the times the learning takes place for each learner in order to achieve the best possible results in terms of retention [9]. The individuality of learning also leads to the desire for a "learning on demand", which means that it should be possible that the learning can take place independently from time and place and anytime the learner wishes to learn [93]. Current technologies, such as mobile devices can foster this desire and can provide learning content in an ubiquitous way.

While learning occurs throughout the entire life, it is commonly seen as an activity that is connected to school, university, or the workplace. However, these learning environments are only one example for situations in which learning takes place. During everyday life, people also keep on learning, either consciously or unconsciously [55]. One main difference is that there are no fixed formal tests of knowledge retention for the latter. This usually means that people learn either because they are personally interested in a certain topic or that they do it unconsciously while performing everyday tasks, the things they do out of habits or things that on the surface are not connected to learning for them. One example for this is playing games.

In traditional learning environments like schools or universities, learning is to be seen as the main activity with the goal to acquire knowledge, which can be reproduced at a later point in time. The most common occasion when a certain knowledge is needed to be reproduced in these scenarios is a test or an exam. This usually leads to a result-driven learning approach, in which learners tend to rely on their short-term memory in order to be able to solve the exercises of the assessments [70]. The most common learning strategies for achieving this goal are called “massing” and “cramming”, which both refer to learning the topics in a very short amount of time to have it reproducible in one’s short-term memory. While massing is defined as a lot of learning events massed together in a short time, cramming is a special form of massing, in which something is learned intensely, often for the first time, in the days or hours before a test [8, 46]. Since both techniques are very result-driven and focused on the short-term memory, they usually have little sustainability [46]. Therefore, many educators and researchers are looking for ways to build a more long-term memory in the learners’ brains. One of the most promising approaches in this direction is called “spacing” or “spaced repetition” [47].

Strategy	Definition
Massing	Learning events are massed together in a short amount of time.
Cramming	Special form of massing; learning something intensely, often for the first time, in the days or hours before a test.
Spacing	Learning events are spaced apart over a longer period of time.

Table 1.1: Definition of learning techniques according to Kornell [46]

While “spaced repetition” is not a new approach in terms of learning techniques, its impact on content selection and to foster the learning process in a mobile learning game has not been researched, yet. Nowadays, almost everybody owns a smartphone and/or a tablet, which are devices that can be perfectly used for learning. Since these devices are usually connected to the internet and are portable and usable nearly everywhere and at any time, learning becomes ubiquitous with them, which supports the desire for a learning on demand. Furthermore, this makes them perfect learning companions in different places, such as at home, in a park, on the bus or while waiting for a train. Today’s smartphones and tablets are powerful devices that are able to perform different tasks, with powering motivating learning games being one of them. They also offer the possibility to show notifications and reminders on the display, triggered by apps, events, or schedules to remind people about different things. One of these things might also be to learn a certain topic at a specific time.

Learning games on the other hand are usually seen as a very motivational approach for learning tasks, which often rely on repetition [33]. However, one problem with repetitions that take place over a long time is that they may become boring quickly. Another problem is that learning in general is often seen as something negative, especially among students, because it is something that has to be done in order to pass a test or an exam and subsequently a course in school or university. Therefore, motivation is a key factor when trying to leverage a learning technique that is focused on continuous repetitions. One approach to tackle this problem can be game-based learning [23]. There are several advantages that come with game-based learning. While learning something statically, for example through reading a textbook, may be seen boring by a student, a learning game cannot only be fun to play but can also be capable of blurring the activity of learning completely. Both factors ideally make the learners return to the game (and therefore the learning process) continuously and take the boredom out of repetitions. Reward systems, such as the possibility to earn badges or rise in rankings which let players compare themselves with other players of the game, can further raise the motivational effect. Therefore, combining the long-term learning effect of spaced repetitions with the motivating effects of game-based learning and the ubiquity of mobile devices, such as smartphones and tablets, seems to be a promising approach for an ideal learning environment.

1.2 Research Questions and Goal of the Research

While the concept of spaced repetition has already been researched in different domains, the implications of using it in combination with mobile learning games as well as its effect on the learning success in this context was still unknown. Therefore, the goal of the research presented in this thesis tries to fill this gap and focuses exactly on this approach. In the past, spaced repetition was mainly used in combination with flash card learning. However, when using it with game-based learning, several influencing factors need to be considered that differ from those with flash cards. First, there is the question, to what extend it is technically possible to integrate a spaced repetition algorithm meaningfully into a learning game. Using flash cards, the algorithm only needs to schedule a repetition interval for a specific card. Within a game, it also needs to deal with things like motivation, different player types and a sophisticated content selection, especially when there is a huge amount of available content.

One problem that the spacing approach faces in all learning environments is that of early and late repetitions. While some spaced repetition algorithms try to deal with this problem by an even more sophisticated calculation of the repetition intervals, it is even more impacted by the motivation to play the game just for fun in game-based learning. This raises the question how to manage user motivation in order to mitigate early repetitions, since they are the bigger problem in terms of spaced repetition.

When trying to analyze the impact that integrating spaced repetition into a mobile learning game on the learning success has, the first question that arises is the question how exactly "learning success" is defined. There are different definitions from different stakeholders, which makes it necessary to establish a definition for the respective domain. After doing so, some strategies to measure the learning success can be derived from the field of learning analytics, which can then be used on collected real-world data from a learning game with a spaced repetition algorithm in place.

Besides the main question about the effect of using spaced repetition in mobile learning games on the learning success, this thesis also deals with the following sub-questions that are connected to the superordinate topic:

- How can spaced repetition be meaningfully integrated into a mobile learning game?
- How does motivation impact the combination of spaced repetition and game-based learning?
- What is the definition of "learning success" in the researched context?
- Which strategies can be derived from the field of learning analytics to measure the effect on the learning success?
- How can the collected real-world data be analyzed in order to find evidence on the impact of spaced repetition in mobile learning games on the learning success?

Summarized, the presented thesis deals with the mentioned questions and ultimately tries to answer the questions if and how spaced repetition can be meaningfully integrated into a mobile learning game and if this combination leads to a better learning success.

1.3 Structure of the Thesis

The research over the last couple of years was split into several parts and the results were presented at conferences and workshops and published in the associated proceedings and journals. These papers and the presented results build the core of this thesis. After some basic information and related work about the learning strategy spaced repetition and the used algorithms (chapter 2.2) as well as adjacent fields of research (chapter 2.1), the published papers are presented, and the core findings are summarized (chapter 3). This is split into three phases, which start with a view on the general possibility to integrate spaced repetition with game-based learning (chapter 3.2), followed by research on the technical implications of the integration (chapter 3.3) and finally a data collection and subsequent analysis (chapter 3.4). The latter concludes the research and answers the question if using a spaced repetition algorithm in a mobile learning game leads to a better learning success compared to playing the game without it (chapter 4). In the last part of this thesis the research results are summarized and discussed (chapter 5.2), and possible follow-up topics are presented (chapter 5.3).

Chapter 2

Research Context

2.1 Adjacent Fields of Research

The presented research was influenced by and is based on basic research from different fields. The four most important of these fields are game-based learning, software development & software design, learning analytics, and spaced repetition. In all these fields a lot of research has been done over the last couple of years in different directions.

In general, game-based learning is seen as a type of learning that can enhance cognitive processes on the one hand [4] and a motivating effect on the other [33]. This combination is therefore an ideal vessel to improve the learning experience of the players. Games can be used in various scenarios and have been proven to have a positive impact for example in mental rotation tasks through the game Tetris [87], as well as on taxonomic concepts [89] or on multitasking skills of pilots [36] to name just a few. In terms of motivation, Garris et al. [33] found that a learner is more motivated when he or she enjoys what he or she is doing and that this can be achieved through games. Malone and Lepper [53] propose a link between motivation and intrinsic learning that can be created by game-based learning. The motivating effects of games as well as the enhancements in terms of cognition make them an ideal environment in combination with repetitive learning, such as spaced repetition, which requires a continuous rehearsal of a specific topic and thus a persistent motivation. In the presented research, a specific type of games is used, in this case mobile learning games. This type of games enables a ubiquitous learning and is therefore especially helpful in a spaced repetition environment, in which the time and date of the repetitions is a crucial factor.

When trying to integrate a spaced repetition algorithm into an existing learning game, there are basically two ways to accomplish this: plug-ins (or synonymic add-ons) and frameworks. A plug-in is defined as a software component which can be added to existing software in order to add a specific feature to it [54] and is generally unknown to the main app at compile-time. This means that plug-ins are especially applicable for situations in which there is need for expansions during runtime [54] and that the main software needs to be designed in a way that the addition of plug-ins is allowed. A framework on the other hand consists of the abstract classes, the operations they implement, and the expectations placed upon the concrete subclasses [25]. Other applications can insert their own specialized code into those frameworks by constructing concrete subclasses that work together. For the presented research, so-called object-oriented frameworks are a good fit. According to Adair [3], the approach of instantiation and combination is primarily driven by the data, which is passed to the framework from the main app. This is in line with the approach taken in the presented research. Since the proof of concept and the subsequent data collection and analysis are executed on Apple's iOS platform, there are two programming language that can be used: Objective-C[1] or Swift[2]. Both languages are object-oriented and there is extensive documentation available.

In order to measure the impact of integrating spaced repetition into a learning game on the learning success, strategies from the field of learning analytics can be derived. It was defined as *"Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs"* by the 1st International Conference on Learning Analytics and Knowledge [86]. Therefore, a collection of data about learners' behaviors and learning patterns has to take place in order to analyze them and be able to draw conclusions from that [10]. The biggest problem to tackle is that there is no common definition for learning success [24], which makes it necessary to find a definition for each respective study. In addition to this, it also important to be aware of the type of knowledge that is learned. Kraiger et al. [49] have developed a classification scheme of learning outcomes to help with this task. According to this scheme, the focus of measurement for factual or declarative knowledge should be on cognitive outcomes and therefore on the amount of knowledge and the accuracy of recall. In order to measure the memorization of declarative knowledge, traditional test scenarios like multiple-choice or true-false tests are suggested [31]. Furthermore, it is proposed that knowledge tests should take place early in the knowledge acquisition process in order to identify knowledge gaps that may have an impact on the learners' performance in the future [1]. In addition to that, the general intelligence of the learner is a very critical factor for the acquisition of knowledge in the early stages of the learning process [2].

The learning strategy of spaced repetition is at the center of the presented research. As mentioned earlier, the effect of "spacing" has already been found at the end of the 19th century [27]. Since then, there has been ongoing research on this topic, which, among other things, led to spacing approaches, such as the Leitner System [52] and the different SuperMemo algorithms [96]. The latter one, more precisely the SM2 algorithm [97], was used for the integration into an established learning game in order to get an insight into the impact that spaced repetition has on the learning success when it is used in combination with mobile game-based learning. More information and related work about spaced repetition is presented in chapter 2.2.

[1] https://apple.co/3575TXo

[2] https://www.swift.org/documentation/

In addition to the four main fields that have influenced the presented work, there were also some minor fields, which were also considered. The following figure 2.1 visualizes the different influencing factors that the presented research is based on:

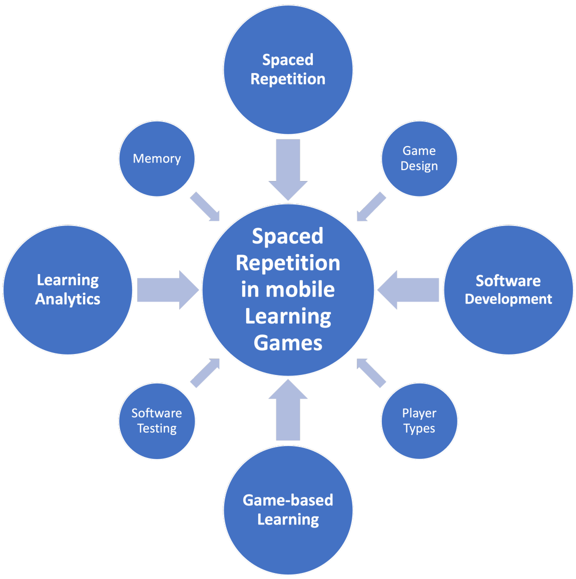


Figure 2.1: Basic research topics with influence on the presented research (own representation)

The research that is presented in this doctoral thesis started in late 2012, which means that a little more than nine years have passed since then. In the meantime, no relevant influencing factors changed in this field of research. The basics of spaced repetition still have their roots in research by Hermann Ebbinghaus [27]. There are many factors influencing the individual learning success and learning progress. These factors also have an impact on calculating the best possible intervals between presentations of a certain learning topic. There are many sophisticated algorithms, which try to achieve exactly that. For the presented research the SM2 algorithm was used, which was released in 1987 by Piotr Wozniak and which was the original computer-based version of the algorithm [96]. Although there have been subsequent versions, the SM2 algorithm has remained to be the most popular version and has been used in several spaced repetition applications, such as Anki[3] and Mnemosyne[4]. Newer versions of the SM2 algorithm tried to further improve the calculations through collected data from the respective previous versions. One major iteration of the SM algorithm was released in 2002 as SM11 and was the first version, which tried to deal with early and late repetitions, an issue that was also dealt with during the early stages of the presented research and which was tried to be solved by adding the self-developed FS algorithm, which takes over the content selection in case of

[3]https://apps.ankiweb.net/
[4]https://mnemosyne-proj.org/

an early repetition in order to avoid that the values used for the SM2 algorithm get compromised by this repetition [75]. While the FS algorithm protects the data, which is used for the interval calculation, the SM11 algorithm uses the data for an improved prediction of the intervals. However, ultimately the decision was made to stick with SM2 due to its popularity, its widespread adoption, a lack of training data, the simplicity of integrating it into an existing game, and the fact that the goal was to research the effect of sticking exactly to the calculated intervals on the learning success anyway. By using a game with a lot of content for the data collection it was furthermore possible to eliminate the problem that can arise with early repetitions due to a lack of a sufficient amount of content. Late repetitions are not as problematic as early repetitions because the algorithm would simply adjust the interval should the retention for the task at hand have been declined. The latest version of the SM algorithm is SM18, which was released in 2019 [99].

Since starting the research there has been an ongoing improvement of the technical capabilities of mobile devices, such as smartphones and tablets. However, these improvements do not have a direct impact on the researched effect of spaced repetition learning on them. One way to make use of the mentioned improvements in the future could be to use spaced repetition learning in combination with other mobile applications, such as augmented reality or virtual reality. Since the underlying spaced repetition calculations would still be the same in these applications, the technical improvements would have more influence on the application itself than on the spaced repetition approach.

2.2 Related Work: Spaced Repetition Basics

Using spaced repetition, learning takes place not only during the days and hours shortly before a test or an exam but "spaced" over time and at certain intervals infinitely in order to keep the information longer in one's memory and thus to build a long-term retention [46]. In general, during each and every learning process, a newly learned information is initially stored in the short-term memory [6]. If it is not rehearsed continuously afterwards, that memory vanishes quickly, gets lost and thus has to be learned completely anew when it is needed. One explanation for this effect is the so-called trace decay, which assumes that memories leave a physical or chemical change in the nervous system, called the memory trace. A continuous rehearsal strengthens this trace and finally leads to moving the information into the long-term memory where it then becomes knowledge [43]. However, even after this has happened, a continuous rehearsal has to take place in order to retain that knowledge. Spaced repetition (or short "spacing") is a learning approach that tries to fulfill exactly this. To do so, a learning topic is learned in a permanent way with continuous repetitions at certain intervals. The simplest form of this technique is commonly used in vocabulary learning with flash cards and a box-system, like the one proposed by Leitner [52]. The learners write the word in one language on one side of the card and in the other language on the other side. Then, there are several boxes in which these learning cards are sorted during each learning event, based on how well the learner was able to remember the word. For example, there may be four boxes as can be seen in the following figure 2.2:

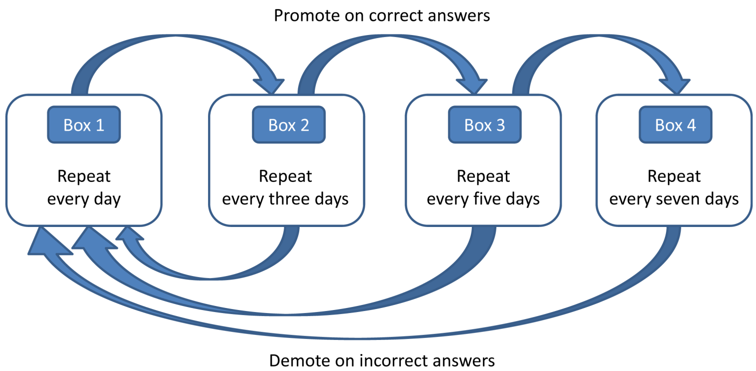


Figure 2.2: Spacing in the Leitner System (Own representation based on [52])

The words on the cards in box 1 are then rehearsed more frequently than those in box 2, the words on the cards in box 2 more frequently than those in box 3 and so on. This will lead to strengthening the memory trace of these words and move them into long-term memory over time. If this happens, which means that the learner is able to remember the word better in the following learning sessions, those cards can be moved to the next boxes and thus will be rehearsed less frequently but still sometimes. This way, the current learning status can be continuously adjusted by the learner. There are also different approaches to determine the intervals between the rehearsals in this approach. For example, words in box 1 may be learned again on the next day, words in box 2 after three days, words in box 3 after five days and words in box 4 after seven days as depicted in figure 2.2. A problem with this system is that every learner may adopt a different scheme of intervals and that those intervals are not supported by scientific findings, which means that they possibly do not make use of the ideal spacing for an efficient and sustainable learning.

One example for an implementation of the card box approach is the Leitner System, which is based on several boxes, that represent the learner's proficiency level [52]. For example, there may be three boxes. The first box holds the learning content that the learner is not very familiar with, while the third box contains the content that he is pretty familiar with. This system can also be extended to more boxes if needed and as described above. When a learned item can be remembered well, it is promoted to the next higher box. On the other hand, when an item cannot be remembered, it always is demoted back into the very first box. This way, the content, which the learners are having a hard time to remember, is repeated more often than the content that they remember well. The result is a more focused and efficient learning. On the downside, while this system makes use of the "spacing" approach, it only follows a static repetition scheme and is not connected to the learner's actual level of knowledge. Furthermore, it does not solve the problem with the determination of the most ideal repetition intervals.

However, there are already sophisticated algorithms that try to determine the best possible interval for each individual learner. These algorithms are based on psychological models and also take the learners' performance in the past into account to calculate the intervals for the presumably best possible learning success. One of the most common algorithms in this domain is called "SuperMemo" (SM) from Piotr Wozniak [96]. This algorithm has been in ongoing development since 1982 and takes several factors into account in order to determine the best intervals for each individual learner and adjusting these intervals continuously to adopt to the learners' performance on each repetition. However, the groundwork for this algorithm was still the aforementioned general learning approach of spacing or "spaced repetition". According to its lead developer, it is based on the following principle:

> *"Intervals should be as long as possible to obtain the minimum frequency of repetitions, and to make the best use of the so-called spacing effect, which says that longer inter-repetition intervals, up to a certain limit, produce stronger memories." [97]*

In other words, the goal is to find the *"Longest possible intervals that do not lead to forgetting"* [9], which means that remembering a learning topic should become as hard as possible to remember, while still being able to remember it in general. This approach is seen as one of the best ways to strengthen the aforementioned memory trace and thus build a long-term retention of knowledge.

One of the most used implementations of the SuperMemo algorithm is SM2. This algorithm calculates the intervals between presentations of a learning item based on the mentioned definitions. Since it recalculates the repetition interval after each learning session, it adapts to the actual performance of the individual learner and extends or shortens the interval until the next presentation of the learning item based on this. The SM2 algorithm was chosen as the algorithm for the presented research due to its widespread usage, its simplicity, and its easiness to implement. The only adjustment that was made was that the learning items were not graded by the players in the mentioned scale. Instead, the quality of the response was altered in a way that it was improved by one on correct answers until a maximum of 5 was reached and degraded by one on incorrect answers until a minimum of 0 was reached. All interval calculations were then made according to the SM2 principles. Based on the correctness of the given answer, this leads to the intervals becoming longer or shorter according to the level of knowledge the learner has reached for each learning item. Therefore, these intervals are unique for each learner and are fitted to their individual knowledge level. The concept of the SM2 algorithm [97] in pseudocode can be seen in the following algorithm 1.

Algorithm 1 SM2 Algorithm (pseudo code)

procedure SM2
    Split the knowledge into the smallest possible items.
    I(n) is the inter-repetition interval after the n-th repetition (in days).
    EF is the E-Factor of a given item.
    All items ← *E-Factor of 2.5.*

Repeat items using the following intervals:
    for $I(1) := 1$
    for $I(2) := 6$
    for $n > 2 : I(n) := I(n-1) * EF$
    → *If interval is a fraction, round it up to the nearest integer.*

After each repetition assess the quality of response in 0-5 grade scale:
    5 - perfect response
    4 - correct response after a hesitation
    3 - correct response recalled with serious difficulty
    2 - incorrect response; where the correct one seemed easy to recall
    1 - incorrect response; the correct one remembered
    0 - complete blackout.

After each repetition modify the E-Factor of the recently repeated item:
    EF':=EF+(0.1-(5-q)*(0.08+(5-q)*0.02))
    where:
    EF' is the new value of the E-Factor
    EF is the old value of the E-Factor
    q is the quality of the response in the 0-5 grade scale.
    → *If EF<1.3 then EF ← 1.3.*

If q<3, start repetitions for the item anew without changing the E-Factor (i.e. use intervals I(1), I(2) etc. as if the item was memorized anew).

However, even when the first implementations of SuperMemo were released, "spaced repetition" was not a completely new topic of research. In fact, the first steps in this domain were taken by German scientist Hermann Ebbinghaus at the end of the 19th century, when he found out about the so-called "forgetting curve" [27], which shows that learning and remembering is usually a matter of retention and time. In his original research, Ebbinghaus created a formula showing the degradation of memories:

$$\mathbf{R} = \mathbf{e}^{\mathbf{-t/S}}$$

In the formula, R is memory retention, S is the relative strength of memory, and t is time. The forgetting curve, which can be seen as the solid black line in figure 2.3, states that the retention of a newly learned information usually drops very quickly and vanishes completely over time if it is not continuously rehearsed. On the other hand, the forgetting curve gets flatter and flatter over time with each repetition that takes place, which leads to a better long-term retention, and which is depicted as the dotted red lines in figure 2.3. In can also be seen that the repetition intervals should ideally become longer over time up to a certain limit, since the learner should be able to remember the learning content better after each repetition.

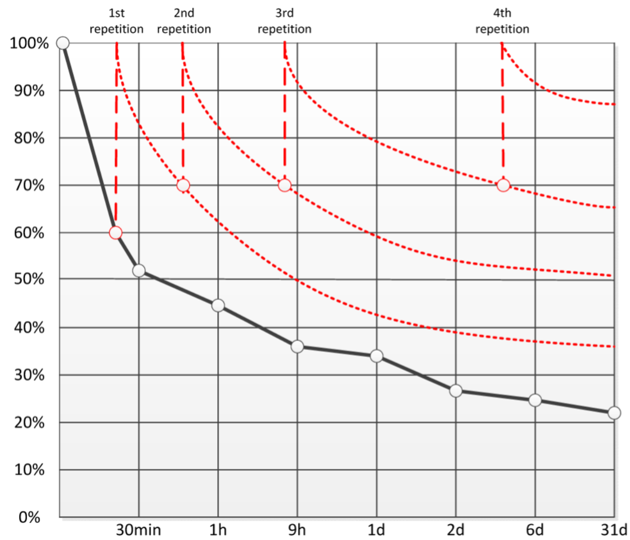


Figure 2.3: Alteration of the forgetting curve through repetition according to Ebbinghaus [27] and estimations from Paul [61].

Therefore, to make the best use of this so-called spacing effect, i.e. the improvement in long-term retention, the repetitions ideally need to take place shortly before the information is forgotten. As mentioned before, algorithms like SuperMemo in its different iterations try to ensure exactly that. To achieve an ideal quality for the calculated intervals, the algorithm needs to collect information about the actual knowledge level, which usually takes some time. Most important is that spacing leads to a long-term approach of learning, which means that learning takes place over a much longer time span than it does for example with massing and cramming.

2.3 Chronological Order of Publications

While dealing with the questions how a spaced repetition algorithm can be integrated into a mobile learning game and how its impact on the learning success will be, different challenges and findings were identified and presented and published at several conferences and workshops. The following figure 2.4 gives an overview over the different research steps and published paper titles that have ultimately led to this thesis.

Figure 2.4: Timeline of published paper topics (own representation)

Chapter 3

Using Spaced Repetition in a Mobile Learning Game

3.1 Different Phases of the Research

The composition of the research comprised three stages that are depicted in the following figure 3.1.

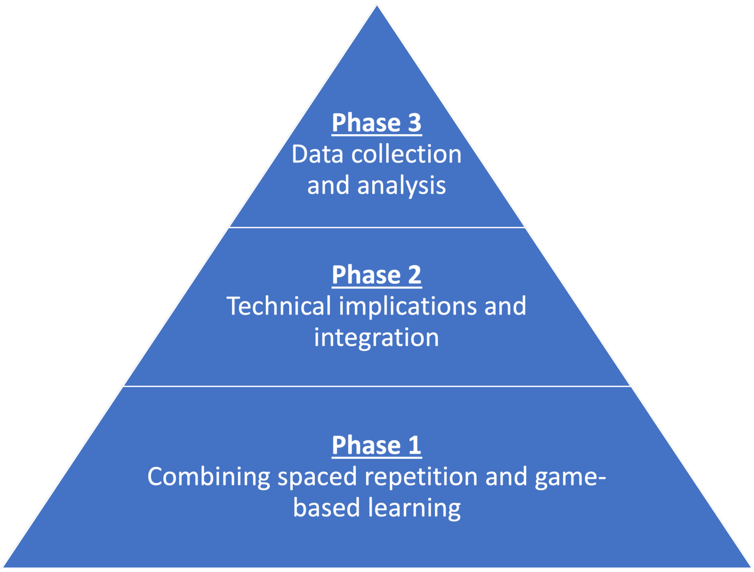


Figure 3.1: Phases of the research (own representation)

After building the groundwork by researching the general possibility and the implications when trying to integrate spaced repetition into a mobile learning game in phase 1, the technical aspects where determined, tested, and implemented in phase 2. In the final step (phase 3), a data collection and subsequent analysis took place in order to verify if the combination of spaced repetition and mobile game-based learning leads indeed to the expected improved learning success.

3.2 Phase 1: Combining spaced repetition and game-based learning

At the beginning of the research on the topic if the combination of spaced repetition and mobile game-based learning has the potential for a better learning success, there was the question if and how a spaced repetition algorithm can be meaningfully integrated into a mobile learning game. The first phase of the research dealt with this topic and included the development of prototype games, technical implications, and the handling of problems like too little content within the game and too much motivation, which can both result in early repetitions and therefore an adulteration of the values that are used for the interval calculation.

3.2.1 What to learn next? Content selection support in mobile game-based learning

In a first step [75, 74, 78] the focus was set on the content selection within a learning game based on the spaced repetition approach. For this, a simple prototype language learning game was developed for testing purposes in order to find out if and how the SM2 algorithm can be implemented into a learning game. The architecture of this game was designed as an all-in-one concept where the content, the data, the logic, the UI, and the algorithm reside all within one app. However, it was planned from the beginning that the goal of this research was not to develop a complete learning game including the spacing algorithms but to provide a spaced repetition framework, which can be integrated into existing learning games through simple and well-documented interfaces.

The prototype game, which included the content as well as the SM2 algorithm works in a way that when the user starts the game, it searches its database for the next content to be learned based on the algorithm. This selection relies on the SM2 algorithm which stores its data in a local database within the app on the user's device. An on-screen notification can optionally be used to remind the user that it is time to learn a specific content after a period of time, that was calculated by the algorithm.

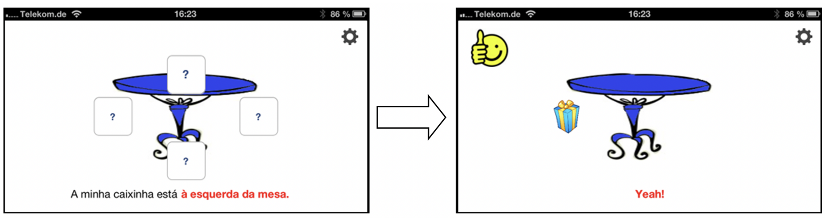


Figure 3.2: Example for task and solution in "Where is my Box?" [75]

In the developed simple language learning game, called “Where is my Box?” there are different categories of things to learn. An object, in this case a gift box, is placed in different positions on the screen depending on the current category and the player needs to find it. In order to do so he gets a task like “My box is left of the table” in Portuguese and then has to tap on the corresponding location on the screen to reveal the box. Depending on the correctness of the player’s answer, an interval for the next repetition of this task is calculated by the SM2 algorithm and the player gets immediate feedback whether the task was solved correctly or not. An example for this can be seen in figure 3.2.

Since the focus of the prototype game was on the technical implications when integrating the SM2 algorithm into it, there was only little content available. This quickly led to early repetitions when the player decided to continue playing although all tasks were already solved. It turned out that if the results of the subsequent rounds for the same learning item would be mirrored back to the SM2 database of calculation values, this would compromise the originally intended calculation of the repetition intervals as can be seen in figure 3.3.

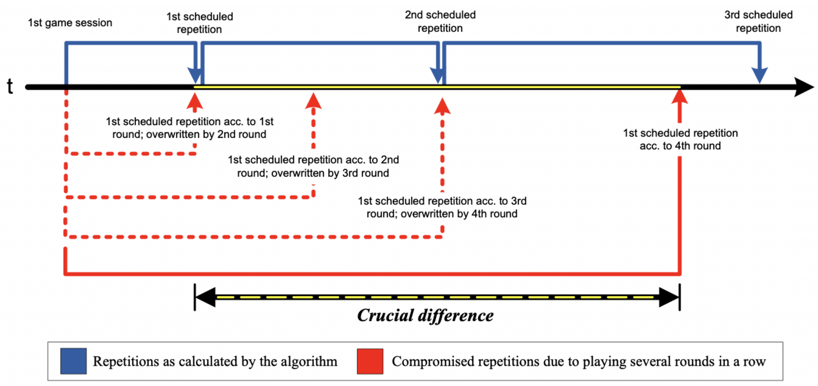


Figure 3.3: The effect of early repetitions [75]

To avoid the adulteration of the values that are stored and used for the interval calculation by the SM2 algorithm, an auxiliary algorithm was developed to avoid repetitions of the same item several times in a row and to keep the integrity of the SM2 algorithm. This algorithm is called the FS algorithm (FS = Follow-Up Sequence). In contrast to the time-based SM2 algorithm, the FS algorithm uses a round-based approach that is obviously more useful in a game with limited content that could be played several times in a row by the learner. The FS algorithm determines in a way similar to the SM2 algorithm which items should be repeated more frequently than others while avoiding back-to-back repetitions of the same item. The course of these events is shown in figure 3.4.

As can be seen, the SM2 algorithm is only in charge for the first round of each learning item when a user starts the app. After this initial round, the FS algorithm takes care of the possible following rounds. The results generated in these additional rounds are not mirrored back to the SM2 algorithm since this would lead to a corrupted scheduling of the next repetitions according to the

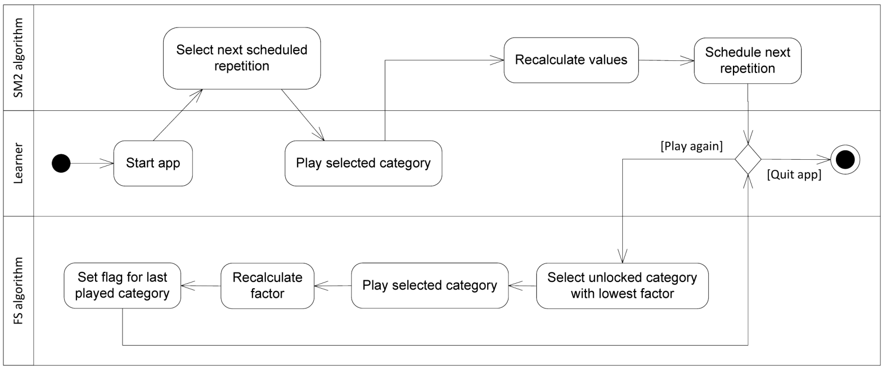


Figure 3.4: Activity diagram of the prototype learning app [75]

spaced repetition approach. To avoid presenting the same content back-to-back while still focusing more on the less well-known content, the FS algorithm saves its own data about the learner's performance in a separate database. This data is then used to keep the game interesting by presenting content in a similar way as the SM2 algorithm would do. However, the FS algorithm is only capable of mitigating the problem and to avoid the incorrect alteration of the SM2 values. The early repetitions remain to be not in line with the spaced repetition approach and undermine the idea of building a long-term retention through the spacing effect.

As mentioned, in contrast to the SM2 algorithm the FS algorithm should not work in a time-based but a round-based manner because a learner might play several rounds in a row during the same session. Therefore, there is no time-based scheduling but a ranking which sorts the learning topics by how well the learner has answered them after each round of play. Just like with the SM2 algorithm the FS algorithm increments the score (i.e. the quality of answer) on right answers and decrements the score on wrong answers. Additionally, there is another value called "relevance" which is increased for the current item by 1.5 on right answers and by 1 on wrong answers. All other categories get decreased by 0.2. The sum of score and relevance is called the rank. The rank then determines the content to play in the next round if the player should choose to carry on playing. Additionally, a flag is set for the last played content. This flag is not present for the other content. Therefore, the unflagged item with the lowest rank is to be played next.[75]

Another problem that arises regarding the content selection is the question how or where to start with from the very beginning, which is also known as the "cold-start problem". The content selection improves as the learner continuously plays the game since the algorithms delivers better and better results with each repetition based on the learned data. However, the starting point remains a key question when using the app for the first time(s). The same problem can be found in other fields of research, for example in techniques like “social bookmarking” or “social tagging” [60]. One approach to tackle this problem would be to use techniques like automatic text analysis and opinion classification [63] or community-based recommendations [85]. The spaced repetition algorithm collects data from users of the game to gain information for the repetition interval. This data might also be used to detect, which learning items are more difficult to remember for learners from a certain group than other materials, though everybody learns and remembers differently. Furthermore, the SM2 algorithm sets the Easiness Factor to an initial value of 2.5 and has two fixed initial values for scheduling repetitions when the user plays a specific topic for the first time. The first repetition will always take place on the following day. The second repetition will take place six days after the first repetition. After that the algorithm determines the next date for a repetition depending on how well the learner remembers the item based on the collected data.[98]

3.2.2 Multi category content selection in spaced repetition based mobile learning games

After the first prototype language learning game [75], another prototype game was developed, which focused on database learning, and which was provided to students of a database lecture at the Hochschule Weserbergland (HSW) [74]. A screenshot from the game can be seen in figure 3.5. Besides examining how this game can be used as an additional supporting tool in a traditional learning environment, it also focused on a categorization of the learning content, which can help to focus more on concepts rather than on bare facts [64]. While the idea behind the SM2 algorithm is to split the knowledge into the smallest possible items [98], a categorization of the learning content seems like a better fitting option when integrating the algorithm into a mobile learning game. For example, with vocabulary learning it is obviously useful to split the knowledge into single words. However, in other scenarios it is more useful to categorize learning content, for example with regard to a specified topic, in which different tasks aim at imparting the concept behind this topic.

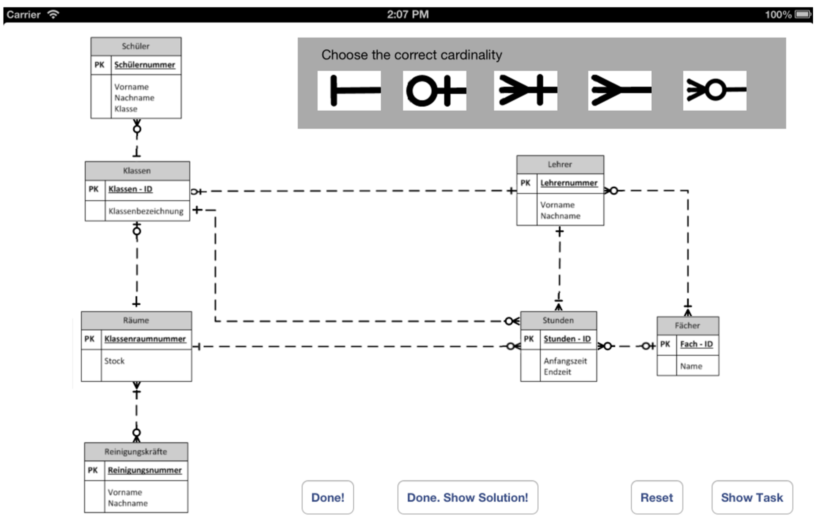


Figure 3.5: Cardinality selection in prototype learning app [74]

It turned out that categories can well be used for content selection in the same way in which atomical learning items can be used. However, depending on the learning topic it might also be useful to have a mechanism that not only selects the correct category for repetition at the correct time but also the correct task within the selected category at that respective time. Nevertheless, it was shown that also categories may be used in combination with a spaced repetition algorithm if the topic at hand should make this necessary or reasonable.[74]

3.2.3 Using a spaced-repetition-based mobile learning game in database lectures

After having used the database learning game during a lecture at the HSW, a small-scale survey among those students who used the game was conducted [78]. The intention was to get an insight into how the students liked the concept of spaced repetitions in a learning game and how they made use of it during a traditional lecture at the university. On the one hand, algorithmic- and performance-data within the game was collected from which the way how the students used the game during the observation period was derived. On the other hand, a survey among the participants was concluded in which they were asked questions about whether they already knew the concept, how they liked it or if they think that they learned better with it [78]. While most participants agreed that spaced repetition is a promising approach for better learning, there was no consent if this technique would be a good way to prepare for a test or an exam. From the collected algorithmic- and performance-data it was obvious that almost none of the students played the game strictly according to the calculated intervals. From the experiences with the prototype language learning game and from the survey among the students it became clear that this was almost completely due to the small number of different tasks within the game and therefore a lack of content. Almost all the participants confirmed that the game would be more helpful and the motivation to use it would be higher is there was more content available [78]. This so-called "content dilemma" can also be found in other areas of e-learning, as several studies from all over the world have already found out. In Malaysia for example, Norazah et al. found out that one of the main challenges for students to deal with e-learning is a lack of proper content [59]. The same applies to e-learning in Bulgaria [94] and is also seen as a general challenge [65].

For the evaluation, the database learning game was distributed to all interested students in a class about database concepts. Ultimately, ten students from the second semester participated in the study. Five of them used their own iPads, five other iPads were given to students who also wanted to participate but did not have an own device. The game consisted of 25 stories and related interactive tasks from the database domain. After the given task is answered, the students get immediate feedback whether their solution was correct or incorrect. If the answer was right, the learners will automatically be taken to the next task as calculated by the respective algorithm. In case of an incorrect answer, the learners are given the opportunity to review the task, the correct solution, and their own provided answer and can then manually switch to the next task. In the background, the SM2 algorithm scheduled the next repetition based on the individual player's performance. The participating students were also provided with a short explanation about how to use the game and a little background information about the idea behind spaced repetitions. However, there were no rules established within the game about using it according to this approach and the students were free to use the game whenever they liked. To get an insight into how and when they used the game, the following data was logged each time a task was played. Most of these values were also used for the final data collection and subsequent analyses in phase 3 of the research.

Value	Meaning
Time of play	The date and time when the current task was played
Next scheduled repetition	The date and time for which the next repetition was scheduled for the current task
Times played	The number of times the current task was already played
E-Factor	The value of the Easiness Factor for the SM2 algorithm
Score	The value of the quality of response for the SM2 algorithm
Relevance	The value of the relevance, used by the FS algorithm for content selection
Algorithm in charge	Indicator which algorithm was in charge for content selection in this round of play

Table 3.1: Values logged while playing the learning game

As expected, all participating students did not strictly stick to the scheduled repetitions and mostly played the game outside of the calculated intervals. This resulted in early as well as in late repetitions. Some reasons for that were later revealed in a questionnaire after the evaluation period (full results in appendix A.1), which was completed by all participants from which six were male and four were female students. All of them were between 18 and 27 years of age, studied business informatics in the second semester and did not take a class in database concepts before. From the questionnaire about the spaced repetition concept in general it was learned that 30% of the participating students had heard about the spaced repetitions approach before and 80% of them believe that this concept leads to a more efficient learning while 60% believe that it leads to a generally better learning. All participants think that doing spaced repetitions over a longer period of time leads to a better long-term memory. However, there are quite mixed opinions about whether spaced repetitions alone are helpful for preparing for a test and whether the intervals between the learning sessions are too long or too short. In the latter case, there is a tendency that the students found the intervals to be slightly too long.

Regarding the provided game, all students found it to be fun and intuitively usable. There were no technical issues in using the game and all of the participants believed that using a learning game similar to the one provided is generally motivating.

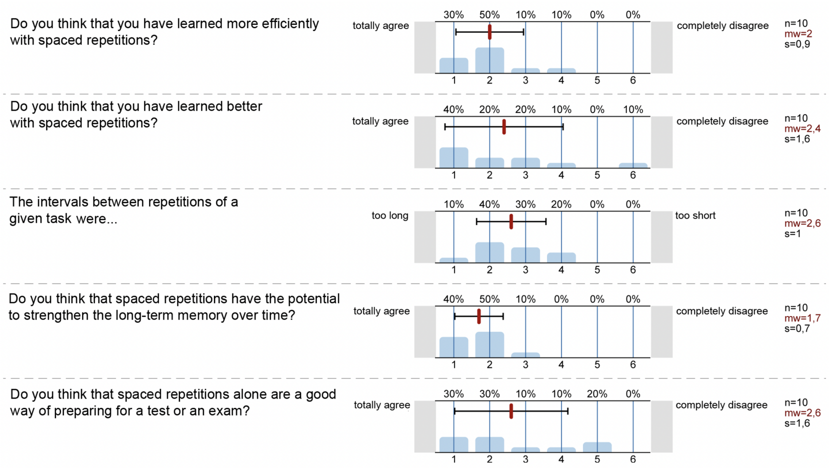


Figure 3.6: Results from questionnaire about spaced repetitions [78]

According to the results of the questionnaire, the key aspects for a good learning game are good user interface design, good visualization, and a motivating effect. Other important aspects are easy navigation and general user friendliness. With that said, there are some improvements to be made in the database learning game. 60% of the participants felt that the graphical design was "satisfying" but not good and half of the participants think that there should be some improvements to be made about the on-screen notifications when a learning day has been reached. Furthermore, all of the participating students also believe that a bigger number of tasks and therefore more variation would be helpful both for the learning success as well as for the motivation to use the game. The provided amount of content in the game (25 tasks) was only seen as "satisfying". Furthermore, the level of difficulty of the tasks should be reviewed. While 60% of the participants rated the level to be just right, 40% found the tasks too easy while none of the students found them too hard [78].

From a technical point of view, both used algorithms (SM2 and FS) worked as intended with the SM2 algorithm only being in charge when there was a scheduled repetition, according to the collected data. Otherwise, the FS algorithm took over the content selection and did not alter the values used by SM2. Therefore, the repetition intervals stayed correct according to the spaced repetition approach used by SM2 and were not rescheduled when the FS algorithm was in charge.

The evaluation results revealed several interesting facts about how students used the learning game and which impact this usage had on their learning progress. It can be said that students often follow their very own learning rhythm, which is only now and then in line with the calculated learning times according to the spaced repetitions approach and also that every student follows a different learning strategy. The appeal of the short-term success, which can be achieved using massing or cramming often leads to preferring them over

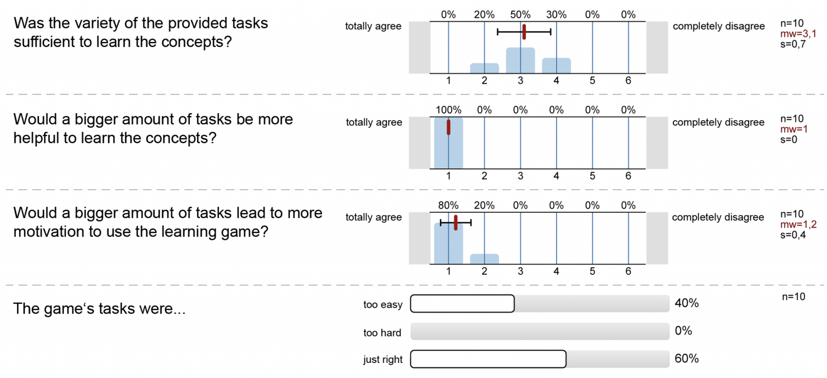


Figure 3.7: Results from questionnaire about the content in the prototype learning game [78]

more long-term oriented strategies like spaced repetitions. While 40% of the participants believe that spaced repetitions alone are not a good way to prepare for a test, 90% them believe that it has a better impact on long-term memory than the other two strategies. However, in a result-oriented society, learning obviously takes often place very result-driven but regarding the obvious effects on long-term memory, spaced repetitions in mobile learning games are still a promising approach to facilitate those effects [78].

3.2.4 Spaced repetition in mobile learning games – A cure to bulimic learning?

Learning strategies like massing or cramming are often connected to the colloquial term of "bulimic learning" [70]. They are often found among learners in traditional learning environments like schools or universities where learning is usually result-driven because students are aiming at passing a test or an exam (ideally with good grades) and try to be prepared just for this special occasion [56]. Using massing or cramming, students start learning intensely during the last days and hours before an exam and usually forget most of the learned content shortly after it. While this strategy is in most cases sufficient for achieving good grades or at least for passing a course, it usually does not lead to long-term retention. The actual effect is however, that most information that was obtained through these techniques is forgotten rather quickly after it was needed [46]. The reason for the illusion that massing and cramming are more effective can be related to tricking one's mind into thinking that one should be familiar with a topic after a massing session of studying information a couple of hours. Looking at that same information only a few hours later then creates a feeling of familiarity because one can still remember a good part of it from short-term memory [8]. Building on that, for most students it seems easier to learn something during the last hours before a test than to start preparing for it weeks or months ahead. However, there are already studies that found out that spacing is more effective than cramming for 90% of the participants of the study [46]. Still, 72% of them thought that cramming is the more effective way of learning. There are, however, domains in which massing may indeed be a good solution to learn. Kornell and Bjork [48] quote a colleague with the words *"spacing is the friend of recall, but the enemy of induction"*. What he means is that while spacing fosters long-term retention of information, massing can be more helpful when trying to induct concepts and patterns. In an example, Kornell states, *"spacing presentations of individual paintings by a given artist will make it more difficult to notice any characteristics that define the artist's style because spacing increases the chances that those characteristics will be forgotten between successive presentations."* On the other hand, the positive effects of spacing have been proven in several domains, such as learning vocabulary [27] or educational materials [12, 22]. It can be derived from this, that massing can be a good solution when trying to optimize the induction of patterns and concepts, while spacing is the better way to learn factual knowledge and transfer it to long-term memory. However, cramming, i.e. learning something intensely in the days or hours before a test, still is not helpful in building a sustainable long-term knowledge.

In order to get a better insight into learning strategies that are currently used by learners in the mentioned traditional learning environments a survey about this topic was conducted among students at the Hochschule Weserbergland [70]. In this survey the participating students were asked about things like their usual learning behavior, whether they are aware that result-driven learning is usually short-lived and if they would consider changing their behavior to a more long-term related one. Overall, there were 135 students participating in the survey from all six semesters and between 18 and 31 years of age. 68.1% of the participants were male, 31.9% were female. Only very few of them deliberately make use of learning strategies, such as the Leitner system or spaced repetition in general. 94.8% of the participants just repeat the learning topics randomly to have them available during a test.

The results showed that the vast majority of the participants in the survey (94.8%) do not deliberately use a specific learning strategy, which includes spacing approaches, but only repeat the topics randomly.

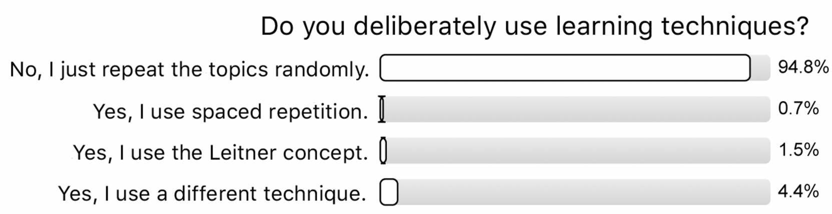


Figure 3.8: Survey among HSW students – Learning techniques [70]

While almost all participants claim that the learning intensity becomes higher towards the end of the semester, there is a huge variety about when the students start to learn during it. However, about 70% of the participants declared that they start learning for a test during the last week before a scheduled test.

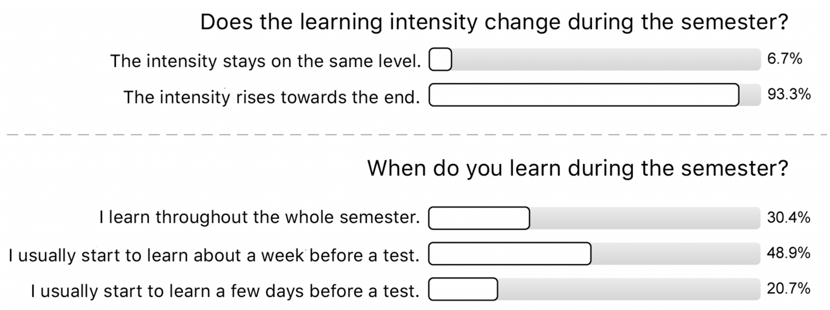


Figure 3.9: Survey among HSW students – Learning times [70]

It was also shown that most of the participating students are indeed learning goal-driven by a test or an exam. In the survey, 85.8% of the participants indicate that the effect of their learning strategy is aimed directly at passing a test, while only 14.2% believe that their strategy has a lasting effect after finishing the course. This is also supported by another question in which the participants were asked about their estimation whether their learning strategy is sustainable or not. 19.5% of the participants believe that they learn sustainably, while 80.4% do not think so or answered neutrally. Furthermore, 83% of the participants are learning during the semester with the goal of getting good grades, 44.4% want to just pass the course and only 38.5% have the goal of acquiring a sustainable knowledge.

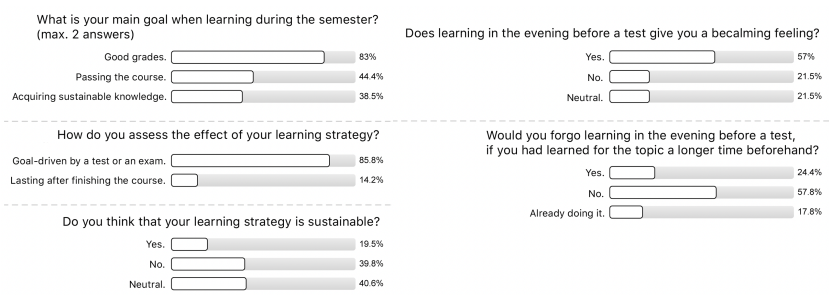


Figure 3.10: Survey among HSW students – Sustainability of learning [70]

The full results of the survey can be found in appendix A.2 and clearly show that learning in traditional learning environments is still result-driven and based on the goal of passing a test or achieving good grades. Since learning during the evening before a test does give the learners a becalming feeling and they would not forgo it even if they had learned the topic for a longer time, it seems likely that they will keep on following this strategy even if they had used spaced repetition beforehand. However, while giving oneself a becalming feeling will most likely be something that cannot be taken away from the learners, using spaced repetition in addition should still have a positive effect on the long-term retention of knowledge when used continuously. Therefore, using spacing as the main learning strategy, which is than supplemented by massing or cramming in the days before a test will still lead to long-term retention, while still fulfilling the illusion of learning success created by the "bulimic learning" strategies [70].

3.2.5 Designing for motivation: Design-considerations for spaced-repetition-based learning games on mobile devices

During the experiences that were made with the two self-developed prototype learning games [75, 74, 78] it was shown that the amount and the variation of the content in the learning game is very crucial – especially for the motivation to use it continuously. This motivation is one of the key aspects why game-based learning is used [19]. However, while motivation is one of the main reasons to use game-based learning, too much motivation stemming from the learning game may be at odds with the sophistically determined intervals between the learning sessions for a certain topic of the spaced repetition algorithm [76]. This means that if the used learning game is too fun to play, the learners may play it just out of fun outside the calculated intervals, which undercuts the idea behind doing rehearsals shortly before a memory vanishes and thus undermines the whole spacing effect. While traditional design for fun games and for learning games likewise usually aims at keeping the player engaged in the game for a long period of time by creating an intrinsic motivation [20], this scenario can be counter-productive for the spaced repetition approach. Especially when there is only limited content available in the game, chances that the learner plays several rounds of the game in a row are pretty high. Since it is not desirable to completely prohibit continuing to play the game, which may annoy or demotivate the players and make them quit the game for good, different solutions need to be found. Other proposed ways to make sure that the players stick to the calculated intervals could be to implement a reward system that motivates the learners to play only at certain times or an explained countdown that tells the learner that he or she needs to wait until the time for learning the topic at hand has come. The latter idea could be based on the same concept that is used by the popular "Farmville" game, in which the player needs to wait a certain amount of time for example between seeding a plant and harvesting it. There are also newer versions of SuperMemo, which take early and belated repetitions into account when calculating the repetition intervals. However, since the goal was to make it as easy as possible for game developers to implement the algorithmic framework containing the SM2 and the FS algorithm in order to get an insight into how and to what extend spaced repetition can be integrated into a learning game, the decision was made to stick with SM2. Furthermore, as mentioned before, the problem with early repetitions is often related to a limited availability of different learning tasks in the game. Having a wider choice of tasks, like it is the case with established learning games, offers the possibility to present other tasks than those that have already been scheduled for repetition in the meantime, which reduces the chance that those are presented to the learner again too early.

When designing a mobile learning game, there are several decisions to be made before thinking about the design itself. First, the type of the game has to be defined. This is usually strongly influenced by the respective learning topic. While one type of a game might be appropriate for one learning topic, it might be inappropriate for others. Another consideration has to be made about the skills the game should teach. Anderson's revision [5] of Bloom's taxonomy [13] represents a good model to classify those skills. However, every kind of a learning game should always be designed to be motivating and engaging to the learners. Davidsson et al. have published an extensive list of design patterns for mobile games [18] which is based on a different hardware, but which still applies to most of the mobile games today. Designing mobile learning games should take those design patterns in consideration as groundwork depending on the type of the game. As an addition, the current hardware used by today's mobile learners should be taken in consideration. Most of those devices rely on touch displays to interact with the user and offer a wide range of technical options like for example gyroscopes, one or more cameras and GPS tracking which may also come into use in learning games.

Talking about motivation, it has to be distinguished between intrinsic and extrinsic motivation. Intrinsic motivation to play the learning game can be created through good game design or it can stem from a learner's interest in the learning topic. While elements like rewards or a great gameplay are extrinsic motivators in the first place, they also have the power to convert this extrinsic motivation to intrinsic if they become meaningful to the learner [21]. One interesting function of mobile devices regarding a time-based learning concept like spaced repetition is the possibility to display on-screen notifications when the time to learn has arrived. These notifications represent an extrinsic trigger to motivate the students to play the game.

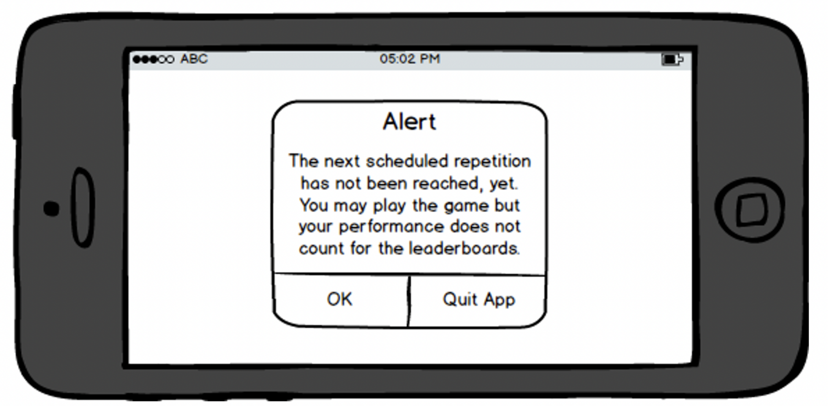


Figure 3.11: User alert about playing at a not scheduled repetition-time [76]

Other examples for extrinsic motivational triggers are rewards, goals or praise. However, there might be a downside in the latter approaches. According to Deci et al. rewards have the potential to lower intrinsic motivation [21]. Furthermore, extrinsic motivation is highly focused on a certain outcome, like a reward, which is in contrast with intrinsic motivation that is focused on the activity itself [68]. Therefore, the reward-system needs to be sophistically designed in order to keep the focus on immersing the learning content. Ideally, the design of the game should contain a balanced combination of extrinsic and intrinsic motivation in order to keep the players engaged for a long time, while also ensuring that building a long-term retention of knowledge stays the main focus of using it. Figure 3.12 shows the different influencing factors for this.

Figure 3.12: Intrinsic vs. extrinsic motivation [76]

As mentioned above, there is a field of conflict between too much intrinsic motivation to play the game and the spaced repetition approach. When the learners get lost in the game because it is too addictive and creates a high level of intrinsic motivation, they might play it at times that are not in line with the calculated spaced repetition intervals. One way to solve this conflict could be to present alerts to the learners when they have accomplished the scheduled tasks for this session and should stop playing the game. However, according to good practices in game-design they should still be able to decide for themselves whether they want to play another round of the game or if they want to quit [57]. Those alerts could then be designed in a way to motivate the users to finish the current round of play while also motivating them to return to the game not earlier than on the next scheduled point of time. This could be achieved by a reward-system which creates an extrinsic motivation to climb for example in a ranking or leaderboard, but which is also directly connected to the scheduling of the learning intervals. This way the learners can only climb in the ranking when playing the game at a scheduled repetition-time and only once (i.e. in the first round of play) at each usage of the game at that time. Overall, spaced-repetition-based learning games should be designed in a way that they motivate learners to play mostly at the calculated times according to the spacing effect. This distinguishes classic learning games from spaced-repetition-based learning games as can be seen in table 3.2.

On a sidenote a proper design for spaced-repetition-based mobile learning games also affects the way the content is categorized. When learning things like vocabulary using learning cards, the content is very atomic, which means that there are single words as learning content. However, in learning games there is often content in different tasks which is very similar to each other, and which

Classic Learning Game	Spaced-Repetition-based Learning Game
Intrinsic triggers make the learner play the game.	Extrinsic reminders make learners play the game at a calculated time.
Played at unspecified times.	Alerts remind learners to play the game when it is best according to the spacing effect.
Engage learners over a long period of time.	Keep learners in the game only for one round of the scheduled content.
Bloom's taxonomy (Anderson revision [5]): Apply, Analyze, Evaluate, Create	Bloom's taxonomy (Anderson revision [5]): Remember, Understand
Raise intrinsic motivation through extrinsic triggers like rewards, every time the learners reach a certain goal	Create motivation to play the game at given times by only rewarding learners when playing according to calculated intervals

Table 3.2: Comparison of design considerations for different kinds of learning games

can therefore be combined in categories. It was already shown how this can be done and what considerations have to be made [74].

3.2.6 Implications of short-term memory research for the design of spaced-repetition-based mobile learning games

As mentioned before, one consequence from early repetitions (especially in games with only little content) that may occur when there is too much motivation to play the game is that learning tasks might be presented ahead of the scheduled time and can therefore be answered from short-term memory. Often times, relying on this memory is even used as a learning strategy in traditional learning environments such as school or university in preparation for a test or an exam. For example, massing and cramming are based on short-term memory. In general, one of the main differences between different learning strategies is the question if they affect more the short-term or the long-term memory. The implications of these two characteristics also have an impact on the design of spaced repetition mobile learning games. Therefore, research literature on the interaction of short term and long-term memory was reviewed in order to determine how short-term memory effects can be coped with in the context of spaced repetition-based learning games [79]. One important finding was the Atkinson-Shiffrin-Model [6], which describes the process of learning from gaining access to some information and processing it over different stages until it is stored in long-term memory and which can be seen in figure 3.13.

First, the information is processed by the sensory memory, which means, that something is seen, heard, or felt. Depending on how much attention this information is given, it gets transferred into short term memory or is imme-

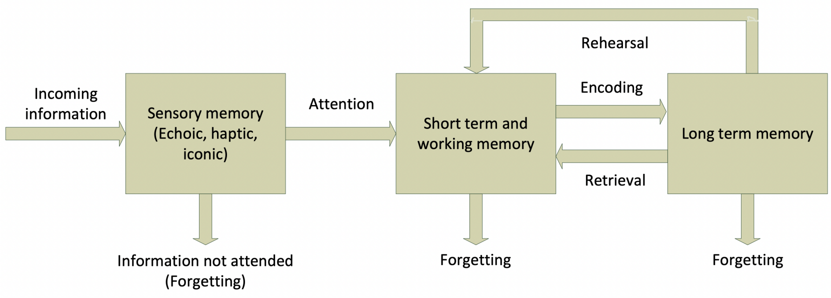


Figure 3.13: Atkinson-Shiffrin-Model [6]

diately forgotten. From short term memory the information is either encoded and transferred to long term memory through the process of consolidation or it is forgotten if consolidation does not take place. Once in long term memory, a continued rehearsal assists in strengthening the memory trace for that information and thus ensures that it is not forgotten over time. The main goal of using a spaced repetition-based mobile learning game is to strengthen the long-term retention of knowledge. In order to achieve this, there has to be a continued rehearsal of knowledge in the long-term memory, which moves it back into the short-term memory from where the process between short-term and long-term memory starts anew. However, since information generally gets transferred from short term memory into long term memory by rehearsal, no matter at which interval this takes place, as long as they are short enough, using intervals other than those proposed by the spaced repetition approach will also strengthen the memory, although at another efficiency. The most important mechanism by which information is moved from short term memory into long term memory is consolidation [43]. While consolidation, among other things, is facilitated by sleep, it must be assured that the rehearsal is not interrupted before it took place. Therefore, the rehearsal must be repeated several times, optimally interrupted by certain intervals [6]. Those are the intervals that are aimed for by the calculations of the SM2 algorithm.

Without an ongoing rehearsal and therefore consolidation, the information vanishes and ultimately becomes forgotten. In general, forgetting information from short-term as well as from long-term memory can be explained by several effects. One explanation for forgetting information from short-term memory is trace decay. This assumes that memories leave a physical or chemical change in the nervous system which is called the memory trace [14]. According to this theory, forgetting occurs as a result of the automatic decay of the memory trace, which may happen in short term memory already between 15 and 30 seconds after learning the information unless it is rehearsed. Forgetting from long term memory on the other hand can be explained by a lack of consolidation. This process needs a certain amount of time to take place. According to Hebb [39] consolidation takes approximately 30 Minutes, but Pinel [62] found out that the process of consolidation can also continue for a number of years.

In general, rehearsals strengthen the long-term memory and therefore enhance the retrievability of information over time. In other words, learners need to practice retrieval of information because it is basically the same process as required on a test [43]. According to Roedinger and Karpicke [67] an excellent way to enhance memory is repeated testing, which is basically what happens through the tasks in a learning game. Cepeda, et al. [16] add that one way of enhancing long term memory is to distribute its study over time rather than trying to study all information in a short amount of time. Furthermore, according to Craik & Lockhard [17], one is more likely to remember the studied information the deeper the level of processing is.

On-Screen Notifications

Strengths: - easy to implement - familiar interface	Weaknesses: - might get ignored - improper times
Opportunities: - available on all devices - anywhere, anytime	Threats: - might get switched off - annoyed learners

Lots of content

Strengths: - rich in variety - repetitions in same session unlikely	Weaknesses: - highly depending on the respective learning topic - difficult content production
Opportunities: - content appears only at calculated times	Threats: - content gets too atomic just to have a lot of it

Leaderboards

Strengths: - only calculated times count	Weaknesses: - prone to cheating
Opportunities: - motivating	Threats: - students cheat to improve rise on the leaderboard

Rise of difficulty

Strengths: - demotivate learners to play out of sequence	Weaknesses: - learners may find the game too difficult
Opportunities: - learners stick to calculated times	Threats: - learners may lose interest in the game completely

Table 3.3: SWOT-analysis of mitigations strategies

In order to achieve this deep level of processing, spaced repetition aims at taking care of a continuous rehearsal and the above-mentioned design strategies can help to tackle the problem of early repetition, which usually make use of the short-term memory rather than the long-term memory. Table 3.3 shows the results of a SWOT-analysis [41] for informing the learners with notifications when it's time to learn or that all learning tasks are done for that session, for the possibility to provide enough content inside the game which makes it unlikely that the same content is presented multiple times, for a class- or university-wide leaderboard in which improvements can only be gained, when using the game at the calculated times, and for using the level of difficulty to discourage the learners to keep on playing [79].

However, the easiest way to avoid early repetitions is still to simply have a lot of content in the learning game, which enables the spaced repetition and content selection algorithms to present different tasks to the player until the time for the next scheduled repetition of a specific content has arrived.

3.3 Phase 2: Technical implications and integration

The second phase of the research focused mainly on technical implications and tests for a later correctly working integration of the spaced repetition and content selection algorithms into an existing learning game.

3.3.1 Architecture considerations for spaced-repetition-based mobile learning games on iOS

The easiest way to have access to a huge amount of content is to make use of already established learning games. Therefore, it was planned to do so from the beginning of the presented research. Furthermore, this opens the possibility to collect a sufficient amount of data, which could subsequently be analyzed to gain insight into the general impact of integrating a spaced repetition algorithm into a mobile learning game. In order to provide external developers an easy way to implement the necessary algorithms and everything else into their games, some architecture considerations were made beforehand [77]. While the two self-developed prototype games [75, 74] were all-in-one concepts in which all components, including the used algorithms reside within one application, the latter need to be provided in a way to external developers, which makes it as easy as possible for them to integrate them into their existing learning games, while still ensuring that all components work together flawlessly. In order to make code easily reusable in several similar software projects, there are two different approaches. Those are usually called frameworks and plug-ins or add-ons. Therefore, a decision had to be made, which of those two approaches should be used.

According to Deutsch [25] a framework consists of the abstract classes, the operations they implement, and the expectations placed upon the concrete subclasses and is constituted by the collection of abstract classes, and their associated algorithms. Other applications can insert their own specialized code into those frameworks by constructing concrete subclasses that work together. Frameworks can be further divided into architectural frameworks and object-oriented frameworks. According to Buschmann and Meunier [15] architectural frameworks provide a set of predefined subsystems as well as rules and guidelines for organizing the relationships between them and thus express a fundamental paradigm for structuring software. Object-oriented frameworks on the other hand can be used in two different ways according to Adair [3]. The first approach is called architecture-driven or inheritance-focused. In this case the main applications can use the framework by deriving classes and overriding operations of it. The second approach is called instantiation and combination and is primarily driven by the data, which is passed to the framework from the main app. The latter kind of frameworks is usually easy to use but limited in its adaptability. However, since it fits perfectly for what was planned in the presented research, this is the kind of framework, which is the most suitable in this context.

A plug-in (or synonymic an add-on) on the other hand is defined as a software component which can be added to existing software and which adds a specific feature to it or enhances it by doing so [54]. This also seems to be a fitting technology for what is planned to be accomplished. Plug-ins are commonly known for example to add a certain feature, like a search engine to a web browser. Since this feature is added afterwards, a plug-in is generally unknown to the main app at compile-time. Therefore, no references to the plug-in are hard-coded into the main app's source code. According to Mayer et al. [54] plug-ins are especially applicable for situations in which there is need for expansions during runtime. They also have the power to modularize huge systems in order to reduce complexity. However, to be able to use plug-ins, the main software needs to be designed in a way that allows the addition of plug-ins or add-ons and the latter need to have a certain knowledge about how to integrate with the main software.

While frameworks and plug-ins offer ways to add reusable code in a convenient way to software, both seem to be suitable for integrating spaced repetition and content selection algorithms into an existing learning game. This approach is based on a separation between the UI and the logic of the existing app on the one hand and the provided content selection and spaced repetition algorithms on the other. Since third-party games already have a dedicated part that is responsible for the game's UI and its logic, the algorithms are provided as a framework that contains everything that is needed in order to implement spaced repetition into the game. This includes the SM2 algorithm, the FS algorithm, an SQLite database for each of the algorithms and the defined and documented interfaces for data exchange between the game and the framework. These interfaces require only minimal adjustments and additions to the code of the apps which want to make use of the framework. The developed overall structure of the provided framework can be seen in figure 3.14.

This way, the changes that need to be made to the existing game are very few. While the game continues to work as before, the implemented framework is responsible for the content selection and repetition scheduling. It also creates an on-screen notification that appears on the player's device when the time has arrived to play a certain topic based on the calculations of the SM2 algorithm. This is an optional feature that makes it easier for the players to stick to the

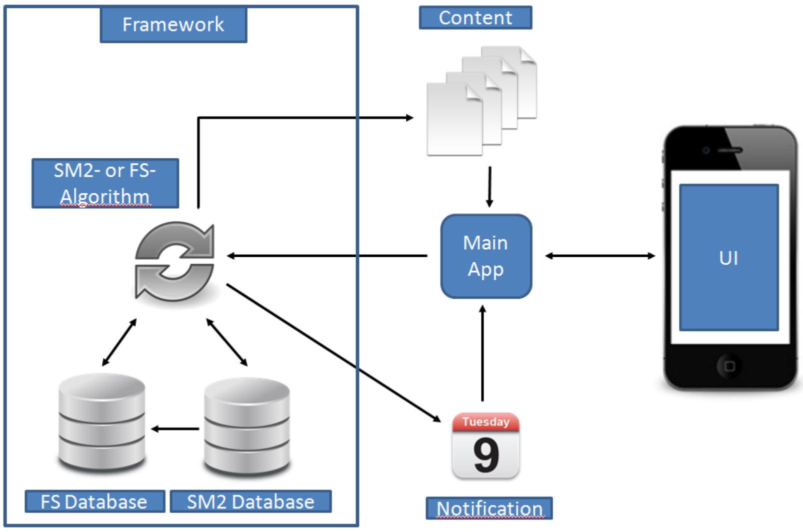


Figure 3.14: Using a framework design [77]

calculated intervals but has to be explicitly enabled by the player. When the game is launched with spaced repetition enabled, it asks the framework, which content should be played. The framework answers with the ID of the selected content, which is then presented to the player by the main part of the game. After that, there is a constant communication between the main part of the game and the framework in order to select the next content to be played according to the spaced repetition approach. After each played task, the framework schedules a local notification on the device, which is presented to the user when the time for the next repetition of that task has arrived.

3.3.2 Simulating context in mobile learning games for testing and debugging

Before using the developed algorithmic framework in real-world scenarios, it was necessary to validate its correct work. While the general integrations of the SM2 algorithm and the FS algorithm into simple mobile learning games were able to show that it is possible to combine spaced repetition and game-based learning, it was hard to test the long-term effects of the algorithms and thus their correct work. Especially in settings that are influenced by a certain context, in this case the time, this context has to be somehow replicated for testing and debugging. Therefore, an external web-based context manipulation interface was developed, which determines when to present a learning item in a learning game based on the SM2 and FS algorithms [81]. Through this interface it is possible to simulate the work of the algorithms at any specific point in the future by creating virtual time hops. This was especially useful since inter-presentation-intervals can be in the range of days to months, which is why system testing in this environment cannot be conducted in a conventional manner. Through the web-based interface, called "Time Machine", it is possible to select a random date in the future that is sent to the connected learning game, which then causes the algorithmic framework to recalculate the next scheduled repetition. The idea behind this is to simulate how the framework will react if the player plays the game at a specific date in the future. Through this tool it became possible to verify that the calculation of the intervals and the content selection work correctly and if the correct algorithm takes over at a certain point in time. For instance, if the next scheduled repetition is still not reached at the configured date, the FS algorithm will be in charge to select the appropriate content. If not, the SM2 algorithm is responsible for this.

The SM2 and the FS algorithm are provided to a learning game as a joint framework. This involves the need for certain interfaces, which are used for data exchange between the game and the framework [77]. Depending on how much is known about what happens behind an interface, this implementation is either seen as a black-box or as a white-box. In a black-box implementation, completely nothing is known about what happens inside the framework. On the other hand, in a white-box implementation, the operations inside the framework can be observed [90].

The work of the two algorithms in the framework happens completely in the background and thus transparent to the using learning game and to the user. While this is intended, it also makes it difficult for researchers and educators to get an insight on whether and how the implementation performs in a real-world scenario. Having this insight is very important to ensure that the algorithms are implemented correctly and that the correct algorithm is in charge at the right time.

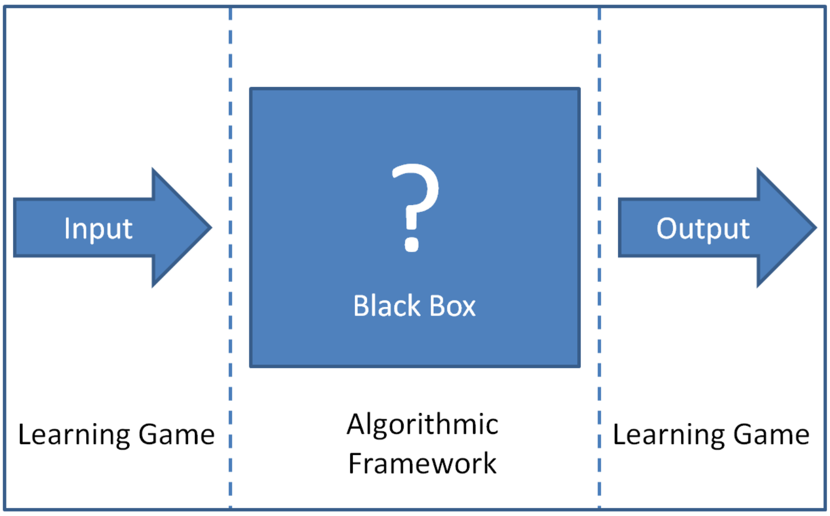


Figure 3.15: Black-Box concept [81]

An abstract example about how a black-box works can be seen in figure 3.15. While the framework consists of the SM2 algorithm as well as the FS algorithm and the corresponding data stores, the learning game on the other hand is responsible for providing the content and the user interface. In this context the algorithmic framework is represented by the black-box. The communication between the framework and the learning game is realized through defined interfaces which correlate to the input and the output of the black-box concept shown in figure 3.15. In this case, the black-box accepts data from the learning game in terms of the learner's performance on an individual task (Input). This data is then processed inside the box which either leads to a next repetition date scheduled by the SM2 algorithm according to the spaced repetition approach or to a content selection based on the FS algorithm if it was an early repetition. This information is then given back to the learning game which presents the respective content accordingly. However, while everything that happens inside the black-box should be implemented as intended, there is no way of simulating and visualizing this, despite waiting for the interval to elapse.

To mitigate the mentioned problem, the aforementioned "Time Machine" web app was developed, which can be connected to the learning game and serve as a trigger for virtual time hops as well as an interface, which can visualize the activities that take place inside the black-box, i.e. the algorithmic framework. The "Time Machine" consists of two components: the TimeMachineServer and the TimeMachineClient which visualizes the algorithm's data. The server application receives and emits the socket.IO-messages and distributes messages to all connected clients. No further logic is implemented in this component. The TimeMachineClient picks up the server's messages and analyzes and converts them into graphical representations. Additionally, it offers the opportunity for interaction with the iOS app. For example, it is possible to simulate the algorithm's work by changing the time and calculate a new next repetition date [81].

Web sockets are used for the connection between the "Time Machine" and the learning game due to their characteristic that they interact in almost real-time. The client is developed with the following frameworks:

Used framework	Purpose
socket.IO	Communicate with the server and to implement web sockets
Angular.js	Used to realize data-binding and user-interaction
Bootstrap	Used to build a fast and responsive user interface

Table 3.4: Used frameworks for the Time Machine

Since spaced repetition is a time-based approach, a way to simulate elapsed repetition intervals was needed in order to demonstrate that the SM2 algorithm works according to the spaced repetition approach and that the FS algorithm takes over in case of an early repetition. Therefore, the "Time Machine" needs to work in a bi-directional way as can be seen in figure 3.16. On one hand there has to be a way to trigger the time hops through the Time Machine and push this data to the learning game, on the other hand there has to be a way to push some values back to the Time Machine from the learning game to illustrate the correct functionality.

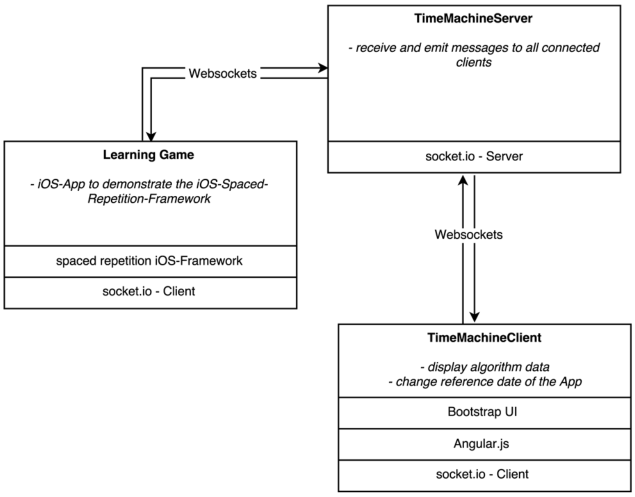


Figure 3.16: Time Machine architecture [81]

As mentioned above, the connection between the web app and the learning game is realized through a web socket, which also serves as the interface for the data exchange. The web app has some advantages over a native application because of the platform independency and the lightweight implementation. The communication between the web app and the learning game is also realized in a lightweight way through JSON. The following figure 3.17 shows a screenshot of the web-interface of the "Time Machine" for choosing a certain date in order to simulate the algorithmic framework's work.

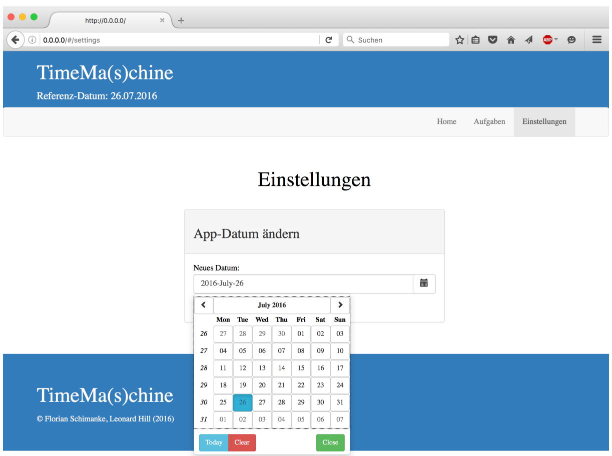


Figure 3.17: Time Machine date setting [81]

Since it is not possible to externally alter the system time on an iPad, on which the prototype learning game runs, the virtual time hop is realized based on the time difference between the current date and the date set by the "Time Machine". This will affect the display notifications as well as the selection of the correct algorithm and the content selection. The approach has shown to be a valuable debugging and testing aid and can be extended for other contextual information sources [81].

3.3.3 A unit testing framework for context variant code in a mobile learning app

In a next step, the approach of context-based simulations was extended to a unit testing framework in order to make sure that the units under test behave in a reproducible and deterministic way. This is again especially important because of external context factors, such as the time, which plays an important role in the spaced repetition approach. In general, unit testing is a method, which aims at testing the different modules of a software's functions [11]. During this process, the software is split into so-called units, which represent a subset of the software's functions. The units are then tested independently from each other. The unit that needs to be tested in the presented case is a black box, which in this case represents the algorithmic framework. Building on the aforementioned "Time Machine" approach, its functionality was extended to also support unit testing [82]. The "Time Machine" is therefore the main interface for the inputs and outputs of the tests, which are defined by the values that are needed by the framework to select the correct algorithm and subsequently the correct learning content. During the unit-tests, the black box is fed with a definable number of dummy tasks through the "Time Machine" web-app, which are then played automatically by the test-framework and are answered randomly correct or incorrect. The results of the test show if all components work as intended, i.e., according to the spaced repetition approach.

As mentioned earlier, the interfaces for input- and output-values between the black-box (i.e., the framework) and the main app were kept to a minimum. All that the framework needs are the number of tasks to be tested, which also works as a unique identifier for the task of each round of play, the current time and date of play and the correctness of the answer given by the learner. Table 3.5 gives an overview of the input- and output-values, which will be used for the unit testing.

During the unit-tests, the black box is fed with a definable number of dummy tasks, which are then played automatically by the test-framework and are answered randomly correct or incorrect. Before a task is played, the test-framework also sends a date value to the framework containing the SM2 and the FS algorithm. This date is seen as the new current date by the algorithmic framework, which makes it possible to check the output of the framework on that specific date. The latter is important because the algorithms in the framework use this information to choose the correct algorithm and, in case of SM2, to calculate the next repetition date. Therefore, using unit testing in combination with the Time Machine offers the possibility to test the games'

Input	Description
Number of tasks	Number of tasks that will be played automatically by the Test-Framework during the unit-test
Correctness of answer	Boolean value of either right or wrong, that can be extended to more sophisticated answers in the future
Date	Time and date value to create virtual time hops during the unit-test
Output	Description
Used algorithm	Either SM2 or FS, based on whether there was an early repetition or not
Next repetition date	Value calculated by the SM2 algorithm
Next task	Either determined by next repetition date (SM2) or by selection of FS algorithm

Table 3.5: Input- and output-values of a spaced repetition black box

behavior even in the future, which is very crucial for a time-based approach like spaced repetition. Entering the three input values from table 3.5 manually in the test-framework through the web interface furthermore makes it possible to test different scenarios in different intervals and with changing values for the correctness.

The sequence diagram in (figure 3.18) shows all steps that take place during the unit-tests in a chronological order. The first step when a unit-test is started is to send a configurable number of tasks to the framework that contains the spaced repetition and content selection algorithms, as well as a date value, which defines the date on which the task is played simulated. Playing the tasks is represented by just sending a Boolean value of right or wrong to the algorithmic framework. Based on this information, it calculates the next repetition date and selects the next task to be played. These values are sent back to the test-framework, which compares them with the expected results. After that, the test-framework sends a new date value to the algorithmic framework, which sets this value as its new system date. In a next step, the test-framework "plays" the task that was calculated by the algorithmic framework earlier on, i.e. it sends a new Boolean value of right or wrong. This sequence of events is repeated as many times as defined by the number of simulated questions, which was chosen during the first step. The most important part in this sequence is to always compare the actual result that is sent back to the test-framework from the algorithmic framework with the expected result. Only if these results are equal, the unit-test was successful.

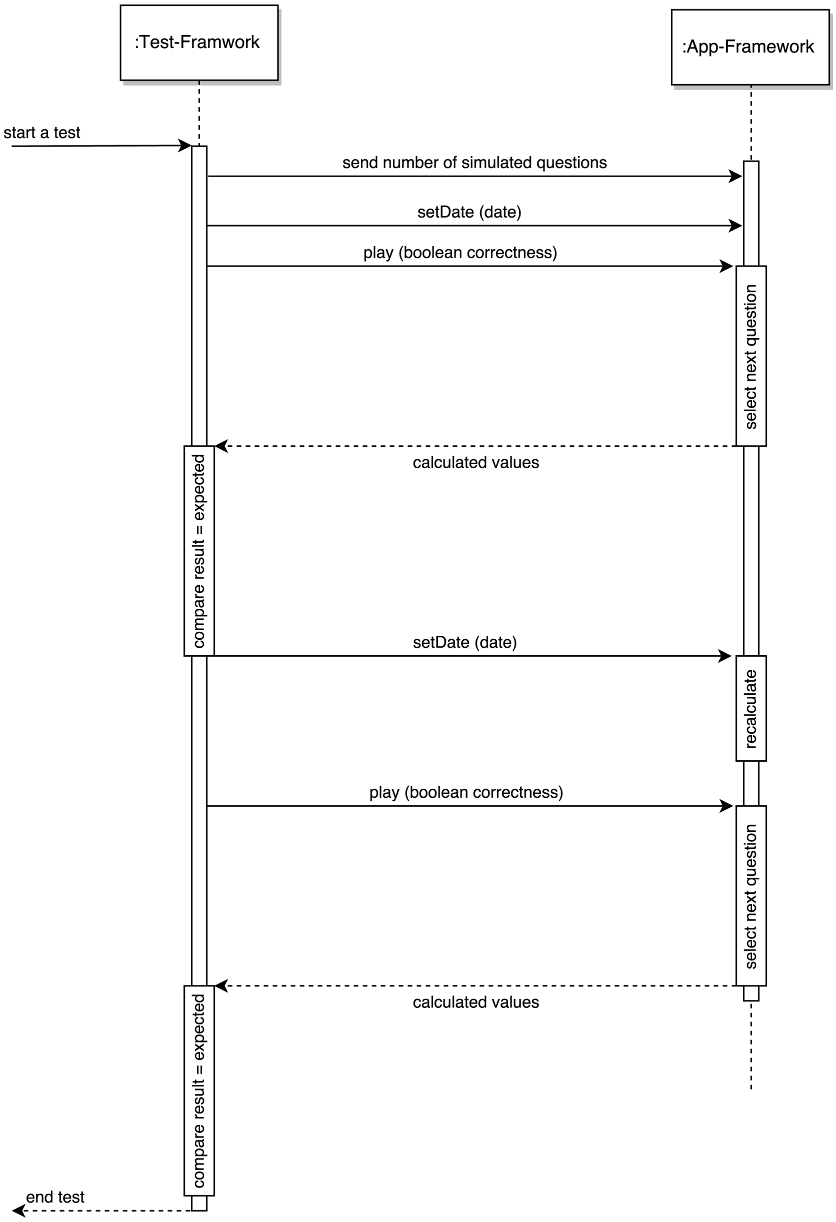


Figure 3.18: Sequence diagram of the unit-tests [82]

When comparing the results, there are some distinct values, which indicate if the algorithmic framework works as intended or not. In any case, the test results need to contain the used algorithm (i.e., SM2 or FS). If an early repetition occurs, the result should be the FS algorithm. If the learner plays the game according to a scheduled repetition, SM2 should be in charge in order to calculate the next repetition date for the played learning item. This next repetition date as well as the content that will be presented to the learner next would be the output in this case. If FS was the algorithm in charge, the output should be the next content to be played as well but this time based on the FS algorithm's own calculations. The next repetition date for the played content should not have been altered in this case and has thus to be the same as before.

The results of the unit-tests are visually presented to the testers through the web-based interface of the extended Time Machine. Since the result can only be either successful or unsuccessful, a green (correct) or red (not correct) indicator distinctly depicts it on first sight.

```
{
 "id": 1,
 "name": "first test",
 "numberOfTasks": 2,
 "Steps": {
  "1": {
   "Input": {
     "time": 1501545600000,
     "boolean": true
   }, …  // Further steps following
```

Figure 3.19: JSON test case example [82]

While the presented approach for unit testing is especially useful for a time-based scenario, in which there is a need for virtual time hops, there are also several other use cases for applications with context variant code. One example would be a location-based application. In this case, the unit-tests would require a certain location to be provided by the test-framework in order to find out if the results from the algorithmic framework are in line with the expectations.

3.3.4 Mobile game-based learning in the app-age – Where we are and where we want to be

After verifying that the algorithmic framework works as intended, it was ready to be implemented into an existing mobile learning game. Before that, the current state of mobile learning was researched and compared it to the state of mobile learning around the early 2000s [71]. While mobile learning (or short m-learning), which usually describes learning that is not bound to a certain location and happens on mobile devices, has been around for many years, it is still not widely used on purpose and supported by sophisticated strategies. Since the beginning of the century, the developments in technology, user interface design and software distribution have changed dramatically, which also impacts the design and the sophistication of learning applications on mobile devices. Especially the ubiquity of personal mobile devices, their computing power and their permanent connection to the internet allow and call for new ways of mobile learning that can easily be integrated in the users' daily routines and stand up to the competition of numerous distractions that come with modern mobile phones. In the Apple App Store as well as in the Google Play Store, games are the most popular kind of apps by far [88]. Combining m-learning with game-based learning, led to mobile game-based learning, which brings the motivational aspects of learning games together with the ubiquitous learning possibilities with mobile devices [34].

e-learning	m-learning (2005)	m-learning (2018)
Computer	Mobile	Mobile
Bandwidth	GPRS, 3G, Bluetooth	WiFi, 4G, Bluetooth
Multimedia	Objects	Multimedia
Interactive	Spontaneous	Interactive
Hyperlinked	Connected	Hyperlinked
Collaborative	Networked	Collaborative
Media-rich	Lightweight	Media-rich
Distance learning	Simulated training	Learning anywhere, anytime
More formal	Informal	Informal
Simulated situation	Realistic situation	Simulated and realistic situations
Hyperlearning	Constructivism, situationism, collaborative	Connectivism, constructivism, situationism

Table 3.6: Complemented from Laouris and Eteokleous [50]

As one can see from table 3.6, while m-learning in 2005 needed to do away with the elements that were connected to traditional e-learning due to technological restrictions of the used mobile devices, today's m-learning can be seen more in a way in that we also see traditional e-learning – with the exception that it is now done on mobile devices. This change offers new opportunities in terms of provided content, communication, and collaboration among the learners. However, simply transferring the traditional e-learning over to mobile devices because of today's devices capabilities should not be the goal. Instead, it is important to adapt the way m-learning is used today to the way mobile devices are used and to the way people learn nowadays. That learning often takes place in cooperation with other learners, which is fostered by the several communication features of mobile devices. On those devices, the local learning often takes place within apps, which is why the current time is also recognized as the "app-age".

Using apps (or more specific games) on the mobile devices, there is often an implicit learning taking place. Of course, the existence of implicit learning is not something that is new or that is unique to m-learning. Implicit learning usually takes place throughout the whole day while we make experiences and interact with our environment. The new capabilities and widespread usage of today's mobile devices make it possible to foster implicit learning by injecting content and learning materials in elements that are used by the mobile device's users without the intention of explicitly wanting to learn something. However, there is still a lot of discussion among researchers about how to evaluate if something is done implicitly or explicitly [50]. In general, this depends on a certain intention or awareness of what one is actually doing. One example would be if somebody sits down to read a textbook about biology, this behavior would usually be seen as an explicit way of learning. On the other hand, if somebody sits down to read a novel in which the human body is described very detailed for contextual purposes, the intention is usually not explicitly to learn. However, just by reading the novel and thus the detailed description of the human body, the reader learns implicitly about this biologic topic.

One field in which implicit learning frequently takes place is game-based learning. While there are already some games available, which are intended to be used explicitly for learning (so-called learning games), the majority of available games does not have this background. However, a lot of these games offer an implicit learning in some kind of way. One example might be a regular role-playing game, which is only available in English. If a non-native English speaker plays this game, he or she needs to deal with the foreign language in order to be able to proceed with the game. The same applies to a geography game in which the player needs to spot a certain location on a map according to various tasks. While the player may most likely have the intention of receiving a high score or gaining virtual rewards, a lot of implicit geography learning happens in the background.

When talking about mobile learning games, also the "mobile" aspect needs to be considered. Mobile devices can be used ubiquitously, which means that the conveyance of the implicit learning content needs to be adapted to that. People use their mobile phones or even tablets frequently in short timeslots [38] like for example when they are commuting on a bus or train. Therefore, the amount of time in which they can play the learning game is pretty short, which means that the tasks within the game need to be small enough to be solved within that time. This needs to be taken in consideration when designing the game in order to benefit from nowadays' ubiquitous usage of mobile devices [76].

Summing up these considerations, there are several aspects around mobile devices, game-based learning, motivation, and intrinsic learning that can be combined in order to create a learning environment that adapts to today's learners and that can also be used for a life-long learning approach. These aspects are [71]:

- Ubiquity of learning/playing opportunities due to available technologies
- Wide-spread deployment of mobile devices, such as smartphones and tablets
- Motivation through well designed game-based learning
- Repeated learning due to motivational aspects of learning games
- Implicit learning by playing (learning) games

In addition to literature research on the topic of mobile game-based learning a survey among students at the Hochschule Weserbergland and random participants was conducted to get an insight into how widespread mobile devices and learning games are used among these two groups. Overall, there were 126 students and 314 random people participating in this survey. The students were from all six semesters and from all offered study programs at the HSW. 61.11% of those participants were male, 34.13% were female, 4.76% did not answer to that question. From the random people group, 84.39% were male, 8.6% were female and 7.01% did not answer. Among all participants, there is a wide-spread distribution of smartphones as well as tablets. Furthermore, if a participant owns a tablet, he or she usually also owns a smartphone. Overall, 72.93% of the general people answered that they own and use a smartphone as well as a tablet on a regular basis, while 16.88% use only a smartphone. Among the HSW students, this is a little different. While the number of students who own and use a smartphone is at 66.67%, only 26.98% of them use both mobile devices.

Regarding the number of installed apps on the devices, there is a huge difference between the students at the HSW and the random participants. While the vast majority of the students (91.27%) have a maximum number of 150 apps installed on their devices, the answers from the general participants spread almost equally across all ranges that were offered as possible answers with even 15.29% having installed more than 200 apps. However, the number of games among those apps is surprisingly small in both groups. Among the students, the number is at a maximum of 20 games for 84.92% of the participants, among the random people, the number is at a maximum of 20 for 73.56% of the participants.

As expected, the number of learning games among the installed games is only a fraction of the absolute number. On the students' devices, the number of learning games among the games is below 10 for 89.68% of the participants, while the number is below 10 for 78.03% of the random participants. Furthermore, these learning games are used less than an hour per day by almost all participants from both groups. It is interesting though, that the number of participants who use learning games less than an hour per day is higher among the students (88.89%) than it is among the random people (77.71%), which means that the latter spent more time of the day using a learning game. Furthermore, while 46.18% of the random participants use learning games at least partially deliberately for acquiring knowledge, this applies to only 36.51% of the students. On the other hand, 74.6% of the latter would like to see a wider integration of mobile learning apps in their studies.

When asked about their estimation about the learning success with learning games, the answers were quite mixed among all participants of the survey with a tendency towards little success. Among the students, 30.83% think that the success is little or very little, 37.5% judge the success as medium and 23.33% as high or very high. A similar picture can be observed among the general participants. In this group, 25.57% believe that the success with learning games is little or very little, 35.6% think that there is a medium success and 21.04% see a high or very high success.

Asked if they think that a scientific approach behind a learning game would raise the likelihood that they would use those games more and whether they would expect a better learning success from this, 49.04% of the random participants expect a better learning success, while 25.16% rate this effect medium an 8.28% expect little or very little change regarding the learning success. A similar picture can be observed among the students, of whom 58.4% expect a positive impact on the learning success, 24% expect a medium effect and 9.6% see a little or very little impact. Integrating a scientific approach in a learning game, however, would motivate more people to use learning games. Among the random people, 67.73% rate the likelihood medium, high or very high. Among the students, this number is even higher with 77.6%. In order to be able to see one's success when using learning games, immediate feedback is seen as an important factor. 73.81% of the participants in the student group rate immediate feedback as important or very important, while this is also the case for 53.18% of the random participants.

The full results from the questionnaires can be found in appendix A.3 and A.4. Summing up the findings, it is quite interesting that both groups would like to see a scientific approach to be integrated into learning games in order to raise the motivation to use them. This is also in line with the expectation that such an approach would be able to improve the learning success of games. Therefore, evaluating this actual success would be the next logical step in the research [71].

3.3.5 Enhancing mobile learning games with spaced repetition and content selection algorithms

After making the decision to provide the algorithms as a framework [77] and validating its correct work through the mentioned simulation [81] and unit testing [82], an appropriate third-party game needed to be found, which fulfilled the mentioned requirements, whose developer was willing to implement the algorithmic framework and from which data could be collected in order to analyze the effect of integrating spaced repetition into a mobile learning game on the learning success [80]. This ultimately led to a cooperation with Jonathan Hillebrand, a game developer who already had a geography game called "Where is that"[1] in the Apple iOS App Store with versions for iPhone and iPad. The game already had more than six million downloads and thus a huge userbase that would presumably generate enough data for the planned analysis at the time the cooperation started. It contained over 2,000 locations in more than 33 categories all over the world and thus is equipped with a huge amount of content, which helps keeping the players' motivation high over a long time, while minimizing the chance of early repetitions and therefore problems with short-term memory effects on the spacing approach.

At the beginning of the cooperation, the game already had a learning mode implemented, which was based on a simplified version of the "Leitner System". This system also belongs to the spaced repetition learning approach and is based on several boxes, which represent the learner's proficiency level [52]. For example, there may be three boxes. The first box holds the learning content that the learner is not very familiar with, while the third box contains the content that he is already pretty familiar with. This system can also be extended to more boxes as needed. In "Where is that", it consisted of four boxes, which led to the

[1]https://apps.apple.com/app/where-is-that/id492240967

effect that locations the players are less familiar with get presented more often than those they already seem to know. On the downside, this system did not track any historical data and does not calculate the optimal intervals between repetitions. Furthermore, if a task was answered incorrect, this would lead to it being demoted into the first box, which is not ideal with regard to the spacing effects discovered by Ebbinghaus.

Therefore, the Leitner System was replaced with the developed framework of spaced repetition and content selection algorithms to make even better use of the spacing effect. The developer agreed to provide anonymized data about how the players played the game, which ultimately enabled the possibility to analyze the impact of spaced repetition in a mobile learning game on the learning success [80].

Within the game the players can chose between different game modes. The used mode for the data collection was the "Quiz Mode" in which the players see a map on which a location is marked and need to choose the correct name of the location from four given options. Having four options is especially helpful in this case, since this reduces the possibility of correct guesses to 25%, which is better than having a 50/50 chance of guessing the correct answer when there are only two possibilities. Going further, since the research is based on a repetition-based concept, the chance to make three correct guesses in a row is at only about 1.6% when having four possible answers compared to 12.5% when having only two options. Furthermore, the playing data was only sent to the data collection server if the player selected the "Learning" mode in combination with the "Quiz Mode" within the game which was based on the "Leitner Spaced Repetition System" in the past and which was changed to the provided algorithmic framework containing the SM2 and the FS algorithms before the data collection.

In general, geography learning, which is at the core of the used game, deals with factual, declarative, or verbal knowledge. The focus of measurement for this kind of knowledge should be on cognitive outcomes and therefore on the amount of knowledge and the accuracy of recall according to a common classification scheme for learning outcomes from Kraiger et al. [49]. The proposed evaluation methods for this measurement are recognition and recall tests, which is what the game does with its multiple choice / single answer approach. Especially the categorization of the content was very helpful for the presented research, since this categorization could be used to characterize the player types based on their playing behavior on the one hand and to see a difference in knowledge development on the other hand. For example, a player from middle Europe will supposedly be familiar with the capitals of Europe and therefore will be more successful in this category, while he might be struggling to find the capitals of countries in Africa or even the correct countries at all and therefore be less successful in this category.

The game's developer was provided with the algorithmic framework and the documentation of the interfaces for data exchange between the main part of the game and the framework, which was then integrated by him. For the game to be able to select a task from a proper category according to the calculated intervals, there has to be a data exchange between the framework and the main game. This data exchange will be done through defined interfaces which can be kept to a minimum in this case due to the overall structure of "Where is that?". Therefore, the basic set of interfaces that are needed to ensure the correct scheduling of repetitions and content selection according to the spaced repetitions approach can be seen in the following iOS source code snippet:

```
//Retrieve Data
//All Categories
(NSString*)getUUID;

//By Category
(NSString*)getUUIDByCategoryID:(NSString*)categoryId
(CGFloat)getEasynessByUUID:(NSString*)uuid
(NSArray*)getEasynessArrayForCategoryId:(NSString*)categoryId
(NSInteger)getBoxForUUID:(NSString*)uuid

//Set Data
setUUID:(NSArray*)uuids forCategoryId:(NSString*)categoryId
setResult:(BOOL)correct forUUID:(NSString*)uuid
setResultQuality:(int)quality forUUID:(NSString*)uuid

//Notitications
-(NSDate)getNotificationDateForCategoryId:(NSString*)categoryId
```

As mentioned above, in a first step the framework must be made aware of all present categories by passing the corresponding Universally Unique Identifier (UUID) to it. After this initial data exchange, all that the game must provide to the framework is the correctness of the played task. The framework will then schedule the next repetition of this task based on this correctness, the number of times the task at hand had already been played and the stored historical playing data that is used for the interval calculation, such as the Easiness Factor. This information as well as the information needed to create the display notification is then passed back to the main app. Since the FS algorithm for content selection basically uses the same data as the SM2 algorithm, there do not need to be special interfaces for it in place. However, in order to keep the integrity of the data used for calculating the correct intervals, the FS algorithm has its own data storage. Passing the needed data for content selection back to the main game is done through the same interfaces that are also used by the SM2 algorithm.

While "Where is that?" already had a learning mode implemented, which was based on the Leitner System, the experience can presumably be even more enhanced by using the provided algorithmic framework, which takes the individual player's historic performance into account for the interval calculation and thus provides a more accurate learning regarding the spacing approach. The tangible effect will later be analyzed from the provided playing data[2].

[2] An example for the data sent from the learning game to the data collection server for analysis can be seen in figure B.1 in the appendix.

3.4 Phase 3: Data collection and analysis

In the third and last phase of the research, the focus was on the collection and the subsequent analysis of real-world playing data in order to find proof if using spaced repetition in mobile learning games leads to an improved learning success.

During the observation period more than 26 million data sets were collected from which about 12 million unique usable data sets were filtered out [3]. Each data set represents a unique game that was played by a specific player. All in all, there were 23,122 unique players during the observation period, from which 68 were identified who played the locations from at least one category strictly according to the intervals that were calculated by the SM2 spaced repetition algorithm. These players were selected to comprise the test group for the data analysis. The reason why the number of players in the test group is only 0.3% from the overall number of players in the data collection can be explained by the used criteria for the selection. Since the main goal was to prove that playing strictly according to the calculated spaced repetition intervals leads to an improved learning success, players needed to be identified for the test group who did exactly that. Therefore, only players who played at least in one category strictly according to the scheduled repetitions were selected from the collected data. It turned out that there were only 68 of all the players during the observation period who fulfilled this criterion.

In an early insight into the collected data [83] different player types were identified, such as “Learners”, “Confirmers”, “Leisure Players” and “Sporadic Players”, who are characterized by different playing patterns and behaviors. While the players in the test group are naturally completely from the player type “Learners”, the players that comprise the control group are also completely from this player type to make both groups as comparable as possible to each other. For the control group another 68 players who were playing the game with the obvious intention of knowledge acquisition in mind (“Learners”) but who did not play according to the suggested intervals were selected. After finding both groups, their learning progress over a time span of six months was compared to find out if playing according to the spaced repetition intervals leads to a better learning success than not playing according to those intervals. While the data collection took place over a stretch of 20 months, the decision to compare only the progress over six months was made because the intervals calculated by the SM2 algorithm become quite long after several correct answers for a location and would then exceed several months.

As mentioned before, the mobile learning game that was used for the presented research was thankfully enhanced by its developer with a framework containing the SM2 algorithm and the FS algorithm. The developer provided access to the needed data collection via a JSON string that was sent to a data collection server for each game played during the observation period. The collected data consisted solely of playing data and did not contain any demographical and other information about the players due to privacy restrictions. However,

[3] Usable means that those were the data sets that remained after filtering out collected data that was irrelevant for the analysis because it was created by players who only played the game once or twice during the observation period.

it was sufficient to get an insight into the playing behavior of individual players and the difference between those players who played strictly according to the calculated spaced repetition intervals and those who did not. An example of a JSON string sent by the "Where is that?" app to the data collection server where it created an entry in a mySQL database for each played location can be seen in figure B.1 in the appendix. The content of the JSON string represents the relevant information that was needed for the subsequent analysis regarding the verification that using spaced repetition in a mobile learning game leads to a better learning success.

The data collection was divided into two periods. The first period lasted three months, during which the algorithmic framework containing the SM2 algorithm, and the FS algorithm was not present in the game. This period was used to filter out players who have played the game during the first observation period from the data that was later used for the analysis. Hereby it was possible to ensure that the observed players in the test group and the control group did not play the game at least during the three months before the data collection during the second observation period and thus did not have previous experience with the game and did not gain geographical knowledge through playing it. This made the two groups even more comparable to each other.

3.4.1 Player types in mobile learning games – playing patterns and motivation

As mentioned above, after receiving the first sets of playing data from "Where is that?", an early analysis took place with the goal of getting an insight into the different player types, their playing patterns, and their assumed motivation to use the game [83]. In a first step the playing-data of the users was collected over a period of three months and subsequently analyzed in order to get a better insight in how the game was played. Since the motivation of the players to use the game is very crucial for the final analysis, a characterization of player types was established to make the collected data more comparable. First, an ABC analysis was conducted on the 12 million data sets in the database to get a better insight on how they were spread among the players [95]. Those 12 million data sets each represent a unique played game. An ABC analysis is based on the "Pareto Principle", which states that 80% of an effect is caused by only 20% of the influencing factors (80/20 rule). Based on the respective field of research, the original 80/20 rule can also be adapted to other splits, such as 70/30 or 90/10 [26]. This method is used to find out how the collected data is spread among three groups, where group A should contain the "important few" while group C contains the "trivial many" [42] players that are responsible for the logged games.

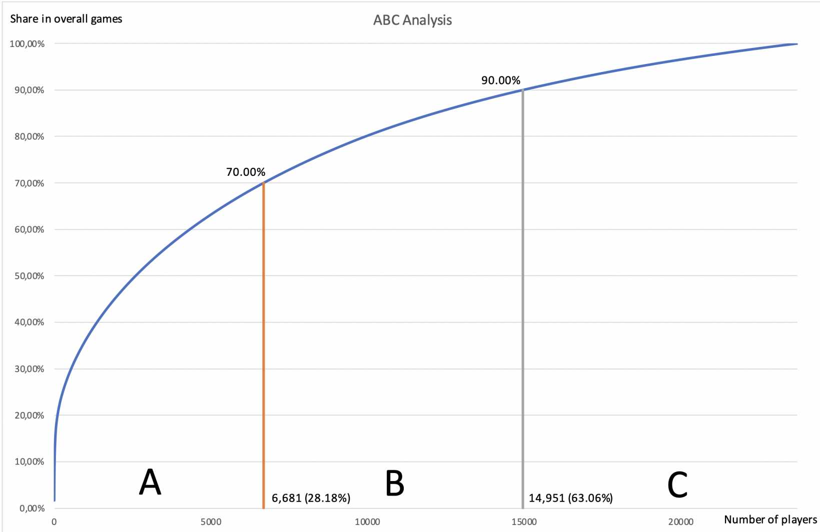


Figure 3.20: ABC analysis of players in relation to played tasks [83]

The analysis showed that 6,681 players (28.18%) were responsible for 70% of the over 12 million played tasks. Building on that, 14,951 of the players (63.06%) were responsible for 90% of the played tasks. This means that 36.94% of the overall players were only responsible for the remaining 10% of the played tasks. The dispersion of the played tasks among the players can be seen in figure 3.20. In group A one can find players that played the game a lot during the time of the data collection, in group B are the players that played the game on a regular basis, while in group C there are a lot of players, who even started the game only once or twice and therefore do not provide any relevant data for the later overall analysis.

After concluding the ABC analysis, the playing habits of players from group A and group B were analyzed in order to build clusters of different strategies and patterns. Players from group C were left out because their contribution to the overall data is not relevant for the analysis. 50 random players from each of the were selected and analyzed regarding their playing habits. More specifically, it was analyzed how many tasks each player played, at which intervals they played, how they performed in each played category and how many different categories they played overall.

Based on the number of different categories and on the number of games that a given player has played, on the average success and on the total time spend in a certain category, four different types of players were identified [83]:

Player type	Description
Learner	A player who plays a huge number of different categories and returns to them even if he is not successful in them.
Confirmer	A player who tends to play mainly the categories in which he has been successful in the past and quits categories quickly when he is not successful in them.
Leisure Player	A player who plays the game frequently but obviously simply out of enjoyment and does not have a certain goal like learning or gaining confirmation in mind.
Sporadic Player	A player who plays the game rather sporadically and does not follow an obvious strategy or pattern.

Table 3.7: Player types in spaced-repetition-based mobile learning games

As long as a game is not specifically designed to work as a dedicated learning game, there will always be different player types with different intentions in playing it. However, even players that do not have learning as their primary intention can profit from the integration of a spaced repetition algorithm into the game. Having a content selection and presentation scheduling for the game's tasks based on the players' performance in the past has the potential to present them the content they should play next, regardless of their intention. Hereby, the learning effect and also the whole game experience can be improved even for "Leisure Players" and "Sporadic Players" who do not follow an obvious learning goal or strategy when playing the game. Regarding the Confirmers, getting them to play the tasks they should play in terms of the spaced repetition calculations, will be a difficult undertaking. Those players primarily aim at gaining confirmation for the knowledge they already have and not so much at acquiring new knowledge. Therefore, these players will very likely select categories they are familiar with and not those that are selected by the spaced repetition algorithm. However, by repeating the already common tasks, players from this category also experience a learning effect, which helps them to strengthen the already available knowledge in their long-term memory.

The most interesting player type for the final analysis if using spaced repetition in a mobile learning game leads to a better learning success is the "Learner", who uses the game with the general goal of knowledge acquisition in mind. Therefore, the playing data of players from this group is going to be the data which will later be analyzed [83].

3.4.2 Measuring the learning success in spaced-repetition-based mobile learning games: Suitable techniques and hurdles

In order to measure learning success based on the collected data, there are several considerations to be made beforehand. It turned out that measuring "success" is one of the most challenging tasks in this context [72]. The main reason for this is that learning is a highly individual process, which is subject to several influencing factors. In addition to this, there is no common definition for learning success in general [24], which makes it necessary to find a definition for each respective study. Sophisticated spaced repetition algorithms already take into account that learning is an individual process, which means that there cannot be fixed repetition intervals for every learner. Instead, those intervals need to adapt to the learning progress of each learner individually. While this can be done by using collected data on the learning performance over time, the measurement of success needs to be directly connected to the definition of "success". In traditional learning environments, measuring and evaluating learning outcomes is usually done through tests or exams. However, this approach can only represent a snapshot of the learners' knowledge at a specific point in time. Moreover, the results of these methods are often diluted through learning strategies by the learners like massing or cramming that are usually result-driven and therefore focus on the reproduction of knowledge from the short-term memory. On the other hand, learning success is not only hard to measure but also a matter of expectations. While some learners may define their learning success as passing a test or exam, others may define it as doing so with good grades, and others again may define it as gaining a long-term memorization. This needs to be considered when establishing a definition for learning success.

In general, evaluating training or a learning process is described as a collection of data points that are related to success [35]. This success is strongly connected to learning outcomes, which is the value that needs to be assessed. In order to do so, a measured outcome needs to be put in relation with the intended learning objectives, such as cognitive, affective, or skill capacities as Kraiger et al. [49] define them. One of the most popular learning evaluation models was proposed by Kirkpatrick [44] and is based on four levels of evaluation: learner reactions, learning itself, behavior, and organizational results. However, even this model measures learning success through assessing the extent to which learners have acquired the given facts or skills through a traditional multiple-choice test, which is more or less the same as writing a test or an exam. Building on the work of Bloom [13] and Gagne [32], who both propose a multidimensional approach for measuring learning outcomes, Kraiger et al. [49] propose three categories of those outcomes and thus possible measures for learning success:

Category	Characteristics
Cognitive	• Verbal knowledge • Knowledge organization • Cognitive strategies
Skill-Based	• Compilation • Automaticity

Affective	• Attitudinal
	• Motivational
	   – Self-efficacy
	   – Goal-setting

Table 3.8: Characteristics of learning outcome categories

Each of the categories touches also on the research presented in this thesis, since the used mobile learning games are focusing primarily on providing factual knowledge, which would be the cognitive part of the classification. From the gained knowledge, the learners should be able to transfer the learning outcomes to real-world or other scenarios, which fits the skill-based category. Since learning games are used to create motivation to learn, this would be the affective part of the classification. While all three of the categories are somehow linked to the approach of spaced repetition enhanced mobile game-based learning, the cognitive category is the one that is most important for the closer measurement of learning success. Cognition is defined by the American Heritage Dictionary of the English Language (4th edition) [40] as follows:

> *"Cognition. The mental process of knowing, including aspects such as awareness, perception, reasoning, and judgment. 2. That which comes to be known, as through perception, reasoning, or intuition; knowledge."*

This definition shows that the key aspect of cognition is to gain knowledge, which is basically also the fundamental aspect of a learning process. After having learned a certain information, this information becomes knowledge, which means that it should reside permanently in the learner's memory. There are, however, different types of knowledge [2], such as declarative knowledge (information about what), procedural knowledge (information about how), or strategic knowledge (information about which, when, and why). In this research the focus is primarily on declarative knowledge, which is commonly seen as a prerequisite for a higher knowledge development [2]. Therefore, evaluating the retention of declarative knowledge should usually take part in the early stages of learning.

According to Gagne [31], traditional test scenarios like multiple-choice or true-false tests are the best fit for measuring the memorization of factual or declarative knowledge. However, in order to get viable results from this kind of evaluation when mapping multiple-choice assessments on the time-based approach of a spaced-repetition-based learning strategy, some adjustments need to be made. Besides the time factor, which is especially interesting in terms of effectiveness, situational and individual factors have also an influence on the efficiency of learning as Noe [58] and Tannenbaum et al. [91] found out. Both factors are explicitly important for the presented research since they touch the concept of spaced repetition mobile game-based learning. While situational factors can stem from the mobile part of the research or the environment in which the learning game is used and cannot be influenced, individual factors are already tackled by calculating the repetition intervals based on the individual learner's performance in the past. Taking this individual factor into account becomes even more important as the general intelligence of the learner seems to be the most critical factor for the acquisition of knowledge in the early stages of the learning process [2]. This means that learners with a higher level of intelligence will most likely perform better in traditional multiple-choice or true-false tests than those with a lower level.

The mentioned traditional test scenarios can also be projected on a spaced repetition based mobile learning game to measure the learners' performance at each time when they play the game. However, this evaluation cannot be made at a specific point in time in this scenario but needs to continue over a certain time span. Ideally, the learners' forgetting curve will improve over time which leads to better results and thus, an improved memorization, which would imply a learning success. Since spaced repetition is a time-based approach, indicators for learning success can be a better performance with each repetition as well as a better efficacy, which means getting better results with fewer repetitions. Therefore, the evaluation of the learning success through the mentioned multiple-choice tests needs to take place over time in order to see the development of the learners' performance. Summarized, the established definition of learning success regarding the presented research reads [72]:

"A quickly improved and lasting retention of knowledge over time."

In other words, a faster and therefore better improvement of retention of players who played according to the calculated spaced repetition intervals compared to a group of players that did not play in accordance with the calculated intervals would show that using spaced repetition leads to a better learning success in comparison [72]. While evaluating and measuring the learning success as well as the efficiency of a learning strategy can very well be done based on a big set of collected data, there are, however, still some soft factors that cannot be measured and that will always have an impact on the individual learner. These factors include several aspects such as individual learning strategies, learning environments, personal preferences, or discipline.

3.4.3 Retrieval of relevant data for measuring the impact of spaced repetition algorithms on the learning success in mobile learning games

Having established a definition for learning success and reviewing different influencing factors for measuring learning success, the next question was how to retrieve the relevant data from the huge amount of overall collected data [84]. While having access to such a huge set of data available for analysis is quite fortunate, it also contains several entries that are either irrelevant or useless for it. Some of the entries may dilute the results, even if the data from group C of the ABC analysis were already filtered out. In order to filter out only the relevant and useful data sets to be able to make discoveries from the data, several iterations of the preprocessing of the data need to be done. One approach for his activity can be the KDD process (knowledge discovery in databases) [30], which in general is a variation of data mining and tries to make sense of the information that is contained in huge databases. According to Fayyad et al. [30], the KDD process is made up of several defined steps, interactive and iterative. It is divided into the following steps:

1. Understanding the purpose of the data analysis
2. Select target data set for the respective purpose
3. Preprocessing the data
4. Data reduction and projection
5. Matching the KDD goals to a data-mining method
6. Exploratory analysis and hypothesis selection
7. Data mining; search for patterns of interest
8. Interpret the mined patterns
9. Using the discovered knowledge

The sequence of these steps can also be seen in figure 3.21. In the first step, an understanding of the purpose of the data analysis has to be built. After that, a target data set has to be selected on which the analysis is going to be made. Third, the data has to be cleaned from any irrelevant or useless data. This step is also called preprocessing. In the fourth step, the preprocessed data has to be reduced for the later analysis, for example based on a prototype data set. After that, the goal behind the KDD process is matched to a data-mining method before an exploratory analysis and hypothesis selection is made in the sixth step. In the following step, the actual data-mining process takes place, in which there is a search for patterns of interest. Based on these findings, the patterns get interpreted in step eight and documented or used further in step nine. Depending on the progress and on the findings in each of the aforementioned steps, it may become necessary to return to any of the previous steps to further narrow down the data. This can be done from each step in the process. In the presented research the KDD process was used to retrieve the relevant data for measuring the impact of spaced repetition algorithms on the learning success in mobile learning games.

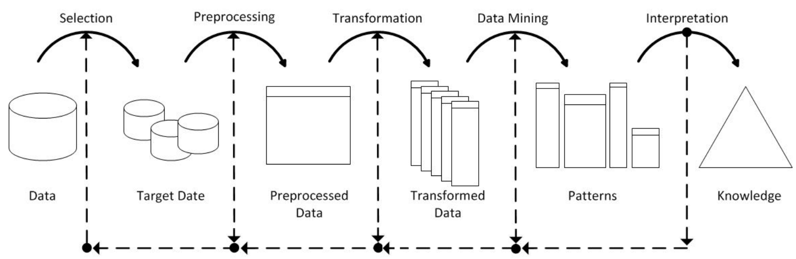


Figure 3.21: Steps of the KDD process according to Fayyad [30]

Since not all collected data was relevant for the planned analysis, a strategy was needed to be found to filter out only the relevant data. Two relevant decisions were already made at this point of the research. First, all players from the mentioned group C were eliminated from the database [83]. Second, only players from the group of the “Learners” [83] will be used for the analysis. The database contains many information and attributes that can be used to filter out even more irrelevant datasets. Before this can be done, a definition for the datasets to be used needs to be made derive some rules for the preprocessing from that. The key-factors for this purpose are:

- players are from the "Learners" group
- playing the game in several sessions
- players have played strictly according to the calculated SM2 intervals for the test group
- a sufficient number of sessions was played
- a sufficient number of games was played per session
- a sufficient number of repetitions was played

Summarized, the early rule-proposals for preprocessing the data with regard to determine the learning success in a spaced repetition mobile learning is intended to find players from the "Learners" group...

- ... with at least 250 unique datasets.
- ... with at least 10 game sessions.
- ... with at least 25 played tasks per session.
- ... who stuck strictly to the calculated repetition intervals.

Figure 3.22: Minimum values for prototype players (own representation)

For a comparison of the established definition of learning success between players who stuck to the calculated repetition intervals and players who played at more or less random times a test group and a control group needed to be compiled. As mentioned earlier, both groups should contain players from the "Learners" type to make them more comparable. The first obvious characteristic for players in both groups is the question whether they played according to the spaced repetition approach or not. Therefore, the test group should contain only players who played at least one category in the game strictly according to the calculated SM2 intervals, while the control group should contain "Learners" who played not strictly according to those intervals. A comparison of the data of the two groups will then show if learning with a spaced repetition algorithm is more successful than learning in a more or less unstructured way.

3.4.4 Deriving strategies for the evaluation of spaced repetition learning in mobile learning applications from learning analytics

Before diving into the actual analysis one last step was to review some related work from the field of learning analytics in order to comply with the findings from that field of research [73]. During this review, the problem with a missing common definition for learning success became obvious again. Therefore, as Laura Patsko [24] points out, “As learners define their learning success quite differently, research has to respect this variation.” When researching learning success, it is not only important to establish a respective definition but it also important to be aware of the type of knowledge that is learned. For example, there have to be other approaches when trying to measure the development of things like gained motoric skills compared to measuring success when learning factual or verbal knowledge. Kraiger et al. [49] have developed a classification scheme of learning outcomes. The used learning game deals with factual knowledge in the field of geography. In general, the quality of an answer is judged by right or wrong. This makes it easy to see a progress over time even if this is not a finely detailed determination of the learners’ knowledge. According to the mentioned classification scheme, the focus of measurement for this factual or declarative knowledge should be on cognitive outcomes and therefore on the amount of knowledge and the accuracy of recall.

Two of the most important influencing factors when analyzing the learning success for spaced repetition in mobile learning games are the correctness of the given answer and the time factor. Combined, these two factors give an indication about how much the amount of correct answers improved over time and thus the learner built a long-term memorization of the knowledge. Since the SM2 spaced repetition algorithm also considers the number of times a task was presented, this value can also be seen as an indicator for the efficiency of learning with spaced repetition. Put simply, the less often a task was presented during the observation period, the better the learner was able to remember the correct answer. This would not only be a proof that the spaced repetition algorithm worked as intended (i.e. calculated the best intervals for the respective learner) but also an indicator for an improved efficiency compared to having a lot of repetitions in the same period. The initial plan for the length of the observation period was three months. However, while examining the collected data it became quickly obvious that a lot of the calculated spaced repetition intervals in the test group already exceeded this period very quickly when the learners performed well from the beginning, which is why the observation period was extended to six months for the test group as well as for the control group. Figure 3.23 depicts the development of the interval for a task in days when the player answers correctly every time.

In order to be able to make a valid statement from the compared data, a common baseline from where to start needed to be found for both groups. To find this baseline, all the answers from the first game of the analyzed data from the test group as well as from the control group were taken and an average score for each group was calculated based on whether the answer was given correct or incorrect. This value will be seen as the “Average Base Knowledge” (ABK)

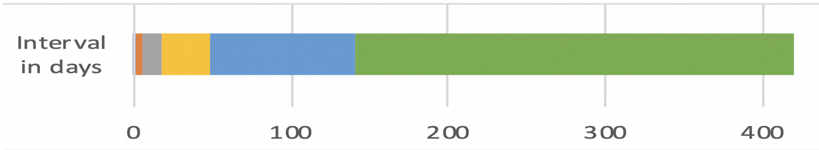


Figure 3.23: Interval development on solely correct answers [73]

for each of the two groups from which the development in knowledge retention and thus in long term memorization over the observation period was analyzed.

3.4.5 The Impact of spaced repetition learning on the learning success in mobile learning games

After building a detailed groundwork for measuring the impact of spaced repetition learning on the learning success in mobile learning games over several years, in a final step an analysis on the collected real-world data was conducted in order to find out, if the proven positive effect of spaced repetition learning in general can also be transferred to learning with a mobile learning game [69]. As planned, the data of players who played the used geography learning game strictly according to the calculated intervals and of players who also played with the goal of learning in mind but at randomly chosen times were compared under the stipulation of the previously made definition.

It turned out that from the overall number of unique players only 68 played the locations from at least one game category strictly according to the calculated SM2 intervals. These players comprise the test group for the data analysis. The same number of players was selected from the remaining members of the player type "learners" to build the control group. In order to make the results of both groups as comparable as possible, a base-knowledge for the test group was determined by calculating the average correctness for the first observed game of all members of the group. The 68 players from the test group achieved a combined ABK of 0.51. This base-knowledge, which also defines the starting point for the observation of the development of the learning success for both groups, was also used when selecting players for the control group. The 68 players in this group achieved a combined average score of 0.50, which means that the players from both groups have a highly comparable base knowledge. The selection of the players for the control group happened randomly and comprises the chronologically first 68 players fulfilling the selection criteria from the overall group of learners. While there are more players that comply with the criteria, it could be ensured that there is a sufficient amount of created playing data by selecting those first 68 players during the observation period. Furthermore, a data collection that took place during the three months before the actual collection was performed to make sure that all analyzed players did not use the game before the observation period.

Starting from the calculated base knowledge the development in retention was observed and the "average retention score" (ARS) for the following games was calculated. For both groups, all games from the 3^{rd} to the n^{th} game were used for this calculation. The second game was intentionally left out since this repetition will always take place on the day after the first game according to the used SM2 spaced repetition algorithm and might therefore water-down the calculation through a short-term retention that may occur on the day after the first presentation of a learning item. Furthermore, only the first fourteen repetitions in each group were observed. While there were partially more repetitions in the control group, the maximum number of repetitions in the test group was fourteen due to the restriction in the observation period, which did not allow more repetitions according to the spaced repetition intervals.

Building on the average base knowledge of about 0.5 for each of the two groups, the subsequent games are the games that build on that and were expected to show an improvement in retention over time for both groups. The better this improvement is, the better is the learning success by the established definition. In other words, a higher ARS at the end of the observation period represents a better learning success. Overall, the players in the test group achieved an ARS of 0.94 during this period, while the players in the control group achieved an ARS of 0.75.

A significant difference between both groups cannot only be seen in the ARS at the end of the observation but especially in the development of the retention over time. As mentioned before, the first repetition always takes place on the day after the initial presentation of a learning item according to the SM2 algorithm. While this repetition was left out for the calculation of the ARS, it can be seen in the visualization of the analyzed data in figure 3.24 that the retention during this repetition was very high within the test group, presumably due to the short-term memory effect resulting from the relatively short interval between the first two games. It reached an average score of 0.88 among the members of the test group, while the average correctness for the second game in the control group has only improved marginally from 0.51 to 0.56. The reason for this is most likely that the first repetition did not necessarily take place on the day after playing the game for the first time but on a day chosen by the player. In fact, the average interval between the first two games in the control group was 3.8 days and therefore almost four times longer than in the test group.

For this visualization, the filtered-out data was exported into two excel files and used to create a boxplot output in R[4], which can be seen in figure 3.24. Boxplots are one of the most frequently used graphical tools for the visualization of analyzed data. They are constructed from the Minimum (or 0th percentile), the Maximum (or 100th percentile), the Median (or 50th percentile), the First Quartile (or 25th percentile) and the Third Quartile (75th percentile). In addition to the box, there can be so-called whiskers, which are basically lines extending from the box's top and bottom and which indicate a variability outside the upper and lower quartiles [45]. For the presented research, the code in B.2.1 in the appendix was used to create the boxplots in R.

[4] "R is a powerful language and environment for statistical computing and graphics." [92]

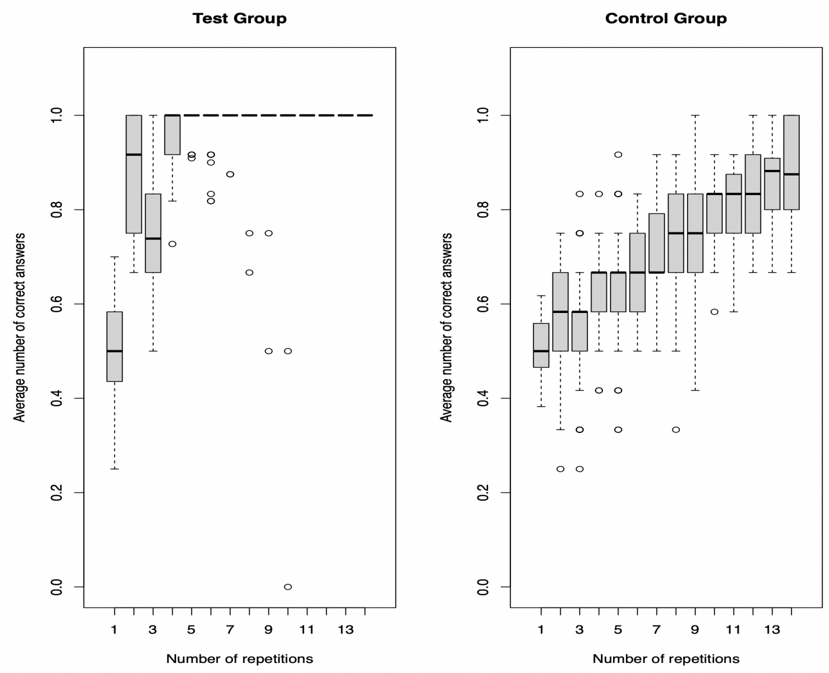


Figure 3.24: Development of the average retention score (ARS) [69]

While the difference in the interval between the first two games likely explains the difference in the performance between the two groups, it also emphasizes the relatively steep decline in retention after learning an item for the first time without short-term repetitions. According to the SM2 algorithm, the second repetition takes place six days after the first repetition of a learning item per definition. Especially for items that were already not well remembered in the first repetition, this is a quite long interval, which makes it likely that the retention will drop even further until the second repetition. This comparatively long period compared to the first interval led to an expected drop in retention compared to the previous repetition from 0.88 to an average retention score of 0.75 for the test group. However, the retention for the second repetition (i.e. the third overall game) in the test group is still better compared to the control group, where it remains at 0.56 correct answers on average, which is the same value as for the first repetition and still only a modest improvement compared to the base knowledge score of 0.50. The average time difference between the second and the third repetition for the control group was 7.63 days. Therefore, the time between the first and the third presentation for the test group was seven days, while the time between the mentioned presentations for the control group was 11,43 days and therefore considerably longer.

For the test group, the SM2 algorithm's calculation based on the learners' performance kicks in after the second repetition, which means that the intervals between the repetitions are no longer fixed but actually calculated by the algorithm. Therefore, the following intervals represent the presumably best possible intervals for the respective learning item and the individual learner. This has already an effect on the overall performance in the test group, which achieves an ARS for the fourth game (i.e. the third repetition) of 0.95. This is a considerably better retention than in the control group, which reaches an ARS of 0.63 for the same repetition without the interval calculation by the SM2 algorithm. The effect can also be seen in the following repetitions, during which the test group achieves an average correctness between 0.98 and 1.0. This shows that the test group does not only reach this quality of retention very fast, it also stays at this level due to the sophisticated calculations of the spaced repetition algorithm, based on the individual performance of the respective learner. This is completely in line with the established definition of learning success for the presented research.

The presented analysis of the collected playing data shows that using a spaced repetition algorithm to calculate the presumably optimal learning intervals in a mobile learning game has a reasonable impact compared to playing the game at more or less randomly chosen times. While it was shown that the members of the control group were able to improve their retention on average with each repetition, they did not reach the values achieved by the players of the test group, who played the observed games strictly according to the intervals that were calculated by the SM2 algorithm. Furthermore, while the test group reaches an almost perfect result already with the third repetition, the control group reaches the highest value of 0.87 correct answers on average with their 14th game, which was the last observed game in the analysis.

Chapter 4

Structure Exploratory Analyses

4.1 Starting Situation

The analysis of the collected data is the main part of the conducted research to finally see if the integration of spaced repetition into mobile learning games has a positive impact on the learning success when using them. While one analysis was already presented in the previous chapter 3.4.5, this chapter presents more structure exploratory analyses to verify those findings. Details about the data collection process can be found in section 3.4 ("Phase 3: Data collection and analysis") of this thesis.

According to the established definition, learning success when using spaced repetition in a mobile learning game is a faster or steeper improvement of the average correctness of given answers over time compared to the average correctness when playing a location of a certain category in the game for the first time. Furthermore, this retention also stays at a high level over time. The expected effect can clearly be seen in the analyzed data, which was collected from the geography learning game. While the retention also improved for players in the control group, the improvements for players in the test group happen faster, steeper and with fewer repetitions and also stay at a high level afterwards.

For the data analysis different methods were used. Since there was no test or training data, no prediction model was built for both groups. Instead, the results of the test group and the control group were compared regarding a distinct difference in their retention progress. The whole analysis is therefore descriptive and not predictive. Besides the analysis in chapter 3.4.5, the generally better learning success according to the established definition when using a spaced repetition algorithm compared to a more or less unplanned learning can also be observed in another visualization of the data in the form of a scatterplot, which can be seen in figure 4.1. A scatterplot uses Cartesian coordinates to display the values of two variables in a diagram. These values are depicted as a collection of points, which represent the value of one variable on the horizontal axis and of the other variable on the vertical axis [28]. In this thesis, the playing data of the test group as well as the data of the control group are displayed in the same diagram and are distinguished from each other by color. The code in B.2.2 was used to create the scatterplot in figure 4.1 in R.

While both groups had an Average Base Knowledge (ABK) of about 0.5, the Average Retention Score (ARS) during the learning games (games 3 to n) is much higher in the test group with a value of 0.94 compared to the control group with a value of 0.75. As can be seen in figure 4.1, there is a difference in the variance of the two groups, which stems from the data selection. While the test group has defined itself by the learners playing strictly according to the calculated intervals, the selection of the members for the control group was based on different criteria. One of these criteria was that the control group needed to have a similar ABK like the test group. Furthermore, the players from the control group played more games during the observation period because they played at random patterns, while the players from the test group played strictly according to the intervals that were calculated by the SM2 algorithm and therefore had a limited number of games during the observation period. This leads to the members of the control group (the blue dots in figure 4.1) gathering more centric around the value of 0.5 for the first played game in the following figure.

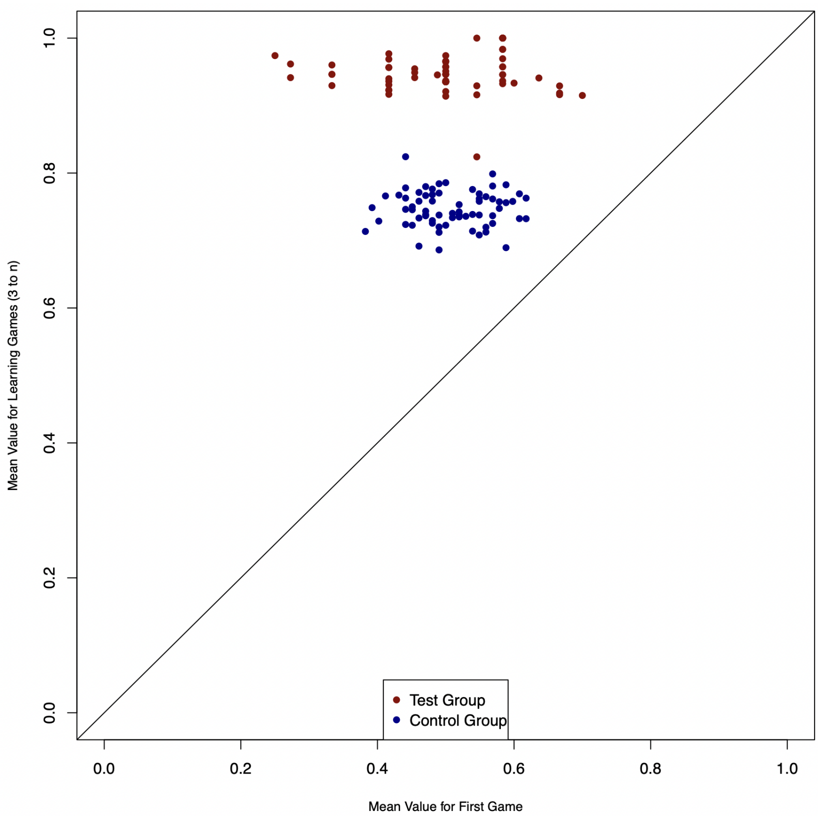


Figure 4.1: Visualization of the base knowledge and the ARS from the learning games (own representation)

4.2 Cluster Analysis

Further evidence for an obvious difference between the test group and the control group was gained by using a structure discovery process on the collected data, in this case an average cluster analysis, which confirmed the first impressions. The goal of a cluster analysis is to find objects that belong together based on a set of descriptive variables [45]. On the other hand, objects from different groups should be distinct from each other. For the research presented in this thesis, the agglomerative approach was used (AGNES). This means that in the beginning, each object represents its own cluster. After that, pairs of clusters are made in each of the following steps, until all objects are merged to one cluster. This approach is also called the "bottom-up" approach. As the result, it should be clearly visible that players from the test group belong together and that players from the control group belong together and that both groups are distinct to each other. Using the average linkage, the average distance between all pairs of objects defines the distance between the two groups [37]. The code in B.2.3 was used to create the dendrogram in figure 4.2 from the collected data in R.

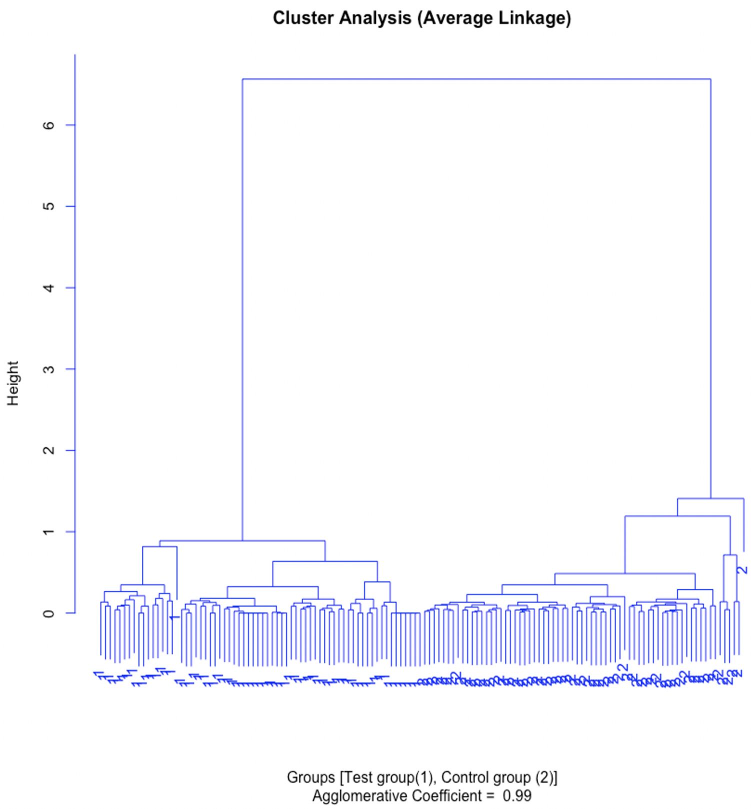


Figure 4.2: Dendrogram of AGNES cluster analysis (own representation)

As can be seen from the dendrogram in figure 4.2, a distinct classification into the two groups can be observed, where the data points labeled with a 1 (all on the left side) are the players from the test group and the data points labeled with a 2 (all on the right side) are the players from the control group. The length of the vertical lines describes the distance between the clustered pairs and follows the rule: the longer the line, the bigger the distance. Therefore, with every step the groups should become more heterogenous, which is the case with the used data. Assigned to this data it can be observed that the biggest distance is between the cluster describing the test group and the cluster describing the control group, which further proves that there is a clear distinction between both groups. As variables for the presented cluster analysis, the average result for the first game, the average result for the 3^{rd} to the n^{th} game, and the average number of games per location were used. While also any other combination of the variables leads to a distinct clustering, the result when using all three of them is the clearest.

4.3 Logistic Regression

As a third method, a logistic regression was used on the data in order to find a functional correlation between different influencing factors. Contrary to a linear regression, the response variables in a logistic regression only have two possible values [7]. In the presented research these values are defined by either having played according to the calculated repetition intervals according to the SM2 algorithm or not. What both approaches have in common is that a response variable is explained through a set of independent variables. For the latter the following variables were used for both groups.

- The average result for the first game.
- The average result for the 3^{rd} to the n^{th} game ARS.
- The average number of games per location.

These three variables were chosen because they are the best fit for describing the learning progress starting from the average base knowledge. The average number of games per location is also an indicator for the efficacy of the chosen learning strategy. This variable may be influenced by the decision to use a spacing approach in general, as well as by the learning effect. While the ideal effect should be that the number of repetitions is minimized by the spacing algorithm, it could also happen that this effect does not occur for a learner, which would lead to a high number of repetitions. On the other hand, the number of repetitions is expected to be higher when not using spaced repetition. Therefore, these variables are the best indicators to research whether playing a mobile learning game in line with the spaced repetition approach leads to a better learning success according to the presented definition. While comparing different numbers and combinations of these variables for the logistic regression, it turned out that the only descriptive variable that can be used for the logistic regression is the ARS. In any other combination the algorithm does not converge.

It turned out that the selected model allocates the analyzed playing data correctly to the test group and the control group at a rate of nearly 100% in the logistic regression. While the players from both groups started with a comparable average base knowledge of about 0.5, this shows a clear distinction between them. The code in B.2.4 was used to perform the logistic regression from the collected data in R. The functional relationship that was found is depicted in the following equation 4.1:

$$P(Y = 1) = \frac{1}{1 + e^{-(-185.6 + 225.2 \cdot ARS)}} \tag{4.1}$$

The logistic regression produces a maximum residuals value of 1.19094, which shows that the model can make an accurate prediction, which player belongs to which of the two groups. The lower this value is, the more accurate is the assumption of the model [29]. Adding weight to this, the following classification table, which resulted from the executed logistic regression shows again the nearly 100% correct allocation of the players to the respective group based on the collected data of 68 players in each group.

Actual Group/ Assumed Group	Test Group	Control Group
Test Group	67	1
Control Group	1	67

Table 4.1: Confusion matrix for group assumption (fitted 0.5)

Summarizing the results of the different analyses that were conducted, there is a clearly visible difference between the test group and the control group. On the one hand, this difference shows that playing the mobile learning game according to the intervals calculated by the SM2 algorithm leads to a different learning progress compared to playing the game at more or less randomly selected points in time. On the other hand, the analysis has shown that playing according to the spaced repetition intervals leads to a better learning success according to the established definition.

Chapter 5

Discussion, Conclusion and Outlook

5.1 Results and Research Contribution

As already outlined in the previous chapter, the superordinate research question about the effect of using spaced repetition in mobile learning games on the learning success was answered positively. While this effect was already known from other learning environments, such as flash cards, the evaluation of the collected real-world data proved that this also applies to game-based learning as long as some influencing factors are considered.

One of these factors is motivation, which is implicitly generated by using a learning game. While the motivational effect is a desired aspect in game-based learning, it turned out to be potentially counterproductive when using it in combination with spaced repetition. If the learners develop too much motivation to play the game just for fun, this can lead to early repetitions of learning content, which is not in line with the calculated repetition intervals. While some spaced repetition algorithms try to mitigate this problem by even more sophisticated calculations, it is not that big as long as there is enough content available in the learning game. If this is the case, the chance of early repetitions is not as big as the chance of late repetitions. However, the effect of the latter will be mitigated by the spaced repetition algorithm, which will simply adjust the scheduling of the intervals if the learning content should have been forgotten in the meantime. Learning games with only a little amount of content can either use newer versions of the SM algorithm or an auxiliary algorithm, such as the self-developed FS algorithm (see chapter 3.2.1) to prevent the values used for the interval calculation by the spaced repetition algorithm from being corrupted by the early repetitions. Games with a sufficient amount of content on the other hand will not be negatively impacted by those early repetitions, since they can simply present other content to the player until a repetition date for a specific content has arrived. This also applies to the used SM2 algorithm, which was not developed to be used in game-based learning in the first place. Since the self-developed FS algorithm is only useful when there is only little amount of content, its development was not continued after the initial phase of the research.

On the technical side, it became clear early that the integration of a spaced repetition algorithm (in this case the SM2 algorithm) in a mobile learning game can be done without any bigger problems. There have to be, however, some considerations to be made in order to ensure the intended functions. While the used algorithm can be hard-coded into a learning game from the beginning, the presented research had the goal to provide existing games with the opportunity to integrate spaced repetition retroactively. To do so, the options of plug-ins and frameworks were researched. It turned out, that frameworks are the better option for this (chapter 3.3.1). An algorithmic framework, containing the SM2 and the FS algorithm was created and provided to a cooperating developer to integrate it into his geography learning game. To further support an easy integration of the framework into an existing game, there have to be well documented and easy to implement interfaces. These interfaces should ensure a minimal data exchange between the main application and the framework, which in the presented work is simply based on already existing identifiers in the learning game. The framework then not only schedules the next repetition date but also provides the main application with the necessary data to create a local notification, which is displayed on the players device in order to remind him or her that the time for a repetition of a specific content has arrived (chapter 3.3.5).

Before providing the framework to the developer, several tests were carried out to ensure the correct work of the algorithms. Especially when using context variant code, like it is the case with time in a spaced repetition environment, testing tools that can simulate this context are important, since it is not useful to always wait until the calculated repetition interval has passed in order to see if the calculation was correct. Therefore, the "Time Machine" was developed (chapter 3.3.2), which makes it possible to make virtual time hops and see how the algorithms behave at different points in time. This tool was later extended to also support unit testing (chapter 3.3.3). Another example for context variant code would be location-based approaches.

After receiving the first real-world playing data from the geography learning game, an early insight already provided some interesting information about different player types using the game. While there are players that clearly use the game with the intention of gaining knowledge ("learners"), there are also a lot of players who play the game simply out of joy (chapter 3.4.1). Furthermore, it became clear that a definition needs to be established for the presented research in order to be able to measure the "learning success" based on the collected real-world data (chapter 3.4.2). This definition was ultimately established as:

"A quickly improved and lasting retention of knowledge over time."

For the analysis of the data, the research field of learning analytics was examined and some strategies were derived from it. Especially, the classification scheme of learning outcomes by Kraiger et al. [49] was useful in order to find the best way of evaluating the learners' progress while playing the game. While judging the quality of the answer given by the players simply by right or wrong may not appear to be very detailed, it turned out that multiple choice tests, which are also based on this type of judgement, are the best fit for the goal of the presented research (chapter 3.4.4). Therefore, it still fits well in this context. They may, however, be refined in future work, for example by taking the time between the presentation of a task and the answer given by the player into consideration.

In the final step of the research, different structure exploratory analyses were conducted. Theses analyses were descriptive and not predictive and were based on the collected real-world playing data from a geography learning game. Besides a cluster analysis (chapter 4.2) and a logistic regression (chapter 4.3), also different visualizations of the data were created. All used approaches showed that playing the content of the used game according to the calculated spaced repetition intervals leads to a better learning success compared to playing the game in an unstructured way according to the established definition.

5.2 Problems and Discussion

During the final steps of the research, some minor hurdles occured, which did not undermine the analysis and the findings, but which made it necessary to adjust some of the previously planned approaches and strategies.

After receiving the data from the second observation period (i.e., with the framework containing the spaced repetition and content selection algorithms added to the game) it turned out that the implementation of the FS algorithm was not working correctly. The problem was that the initial values for the FS algorithm were not set correctly and only sent to the framework after the first presentation of a learning item according to the FS algorithm. Since those initial values were missing after implementing the algorithms, a content selection according to the FS algorithm never took place. However, due to the huge amount of content within the game and since the work and effects of the FS algorithm were not part of the research goal, this problem was not relevant for the analysis. While the content selection of the FS algorithm did not happen, its presence still solved the problem that early repetition would have compromised the data used by the SM2 algorithm. The framework correctly selected the algorithm to be used and altered only the data for this algorithm in the database.

Another problem, which occurred was that there is a confounding variable, which may have an influence on the data analysis. When starting the game, the players can choose whether they want to play the game just for fun or in an offered learning mode, which enables the use of the spacing algorithm. This decision may be relevant in terms of the general intelligence and the character of the respective player. This means that if a player decides to play the game in the learning mode, this decision may already mean that the player understands he may achieve better learning results when selecting this mode, which might then result in a better learning success compared to players who do not make this conscious decision. However, the impact of this factor should not be significant due to two allaying circumstances. First, as mentioned multiple times before, learning is always a highly individual process, which includes decisions if, when, and how to learn. These decisions naturally have an influence on how people learn and what learning success they might achieve. The individual definition of learning success by every learner as mentioned before might also play a role for the decision. Of course, the decisions about if, when, and how to learn are also reflected when selecting the game mode in the used geography learning game. Second, the test group and the control group are still comparable since only players for both groups were chosen which had selected to play the game in the learning mode and have therefore made the same decision. The main difference between both groups stays that the players from the test group played at least one category from the game strictly according to the calculated spaced repetition intervals, while the players from the control group did not stick to these intervals.

After reviewing the data, that was received from the learning game, also some adjustments to the planned analysis were necessary. While it was previously planned to analyze only data over a time span of three months, it turned out that this time span was too short to make a thorough analysis. The main problem was that the calculated intervals between two presentations of a learning item may become quite long very quickly due to the calculations of the SM2 algorithm if the player answers the tasks correctly multiple times right from the start. Therefore, the observation period from the originally planned three months was extended to six months in order to get a more meaningful insight into the long-term development of retention.

Some other adjustments needed to be made regarding the filter criteria that was previously defined for selecting the "prototype players" after getting a better insight into the data [84]. To make the results more comparable, especially regarding the mentioned confounding variable, only players were selected for the control group, which also had the learning mode of the game enabled. This, combined with the criteria characterizing the players as "learners" and the same base knowledge created a control group, which is highly comparable to the test group and therefore represented the best data for the analysis.

While it was previously planned to compare players from the first observation period, in which the algorithms were not present in the geography learning game, with players form the second observation period, this approach was later changed. Instead, solely data from the second observation period was used and the members of the test group and the control group were both completely from this data collection. Through this adjustment it was possible to filter out all players that had already played the game in the previous three months, which comprised the first observation period. By filtering out those players, it was possible to select only players, which had not played the game and thus built any knowledge through it before the start of the analysis, at least not in the previous three months. It was therefore possible to calculate a base knowledge for both, the test group and the control group without this base knowledge being influenced by any previously played rounds in the game. Furthermore, as mentioned before, since the observation period was extended from three months to six months and the first period lasted only three months, analyzing just the data from the second period delivered more comparable results for both groups.

5.3 Conclusion and future work

During the work on the presented research topic over the past years, a deep insight into several aspects and influencing factors around it was gained. The goal was to find out if a combination of the motivating effects of mobile game-based learning with the sophisticated calculation of the best possible learning intervals as stated by the spaced repetition learning strategy leads to an improved and long-term learning success for the players. The research was started by examining if and how the concept of spaced repetition can be used in combination with mobile learning games and it turned out that the technical integration of the used SM2 algorithm itself can be done without any bigger problems, especially when using a game that is based on categories and tasks that can be valued by right or wrong. However, besides the technical aspects, there are several things that need to be considered. One critical factor is the amount of available content in the learning game. If there is only a small number of different tasks in combination with a high motivation stemming from the used game, this very likely leads to early repetitions, which are no longer in line with the spacing approach. If those early repetitions are not addressed, the values used by the spaced repetition algorithm for calculating the intervals may get corrupted and the intervals do not get calculated correctly afterwards. Late repetitions on the other hand are not as critical as early repetitions since they would presumably simply lead to a further decline in retention, which would be addressed by a shortened interval from the used SM2 algorithm the next time the task is played if it was not answered correctly.

Furthermore, little amount of content can also lead to boredom with the learners, which could cause them to stop using the game quickly. To address the problem with early repetitions, an auxiliary algorithm was introduced for the first prototype games, which took over the content selection based on its own values in case of early repetitions to make sure that the values used by the SM2 algorithm do not get corrupted. However, the problem with a lack of sufficient content within the game was mainly a problem with the first two self-developed prototype games, which had the main goal to examine the technical implementation of the algorithm and should not be given when using established learning games, which usually carry a huge amount of already available content.

While the problem with early repetitions can stem from a lack of content as mentioned before, another possible cause for this problem was also discovered. If the learning game is designed in a way that it is just fun to play, even without the goal of learning in mind, this could also lead to players playing the game more often than planned and therefore not in line with the spaced repetition idea. There are several ways to address this problem and it should not be too significant when there is a huge amount of content available, which can be presented to the players instead of the algorithmically scheduled tasks when they decide to play outside of the calculated intervals. If this should not be the case, the players can be motivated to play the game according to the calculated intervals for example by implementing a reward system or through waiting strategies like used in the popular "Farmville" game, in which the player needs to wait a certain amount of time between seeding a plant and harvesting it.

To investigate the effect of using spaced repetitions in a mobile learning game on the learning success, the SM2 algorithm and the FS algorithm were provided as a framework for third-party developers, which can be easily integrated into already established games. The data exchange between the main game and the framework takes place through just a few interfaces and with minimal effort by the developer. Through this integration a huge amount of data was collected with an established learning game, which was subsequently analyzed regarding a possible difference in the learning success for players sticking to the calculated intervals and those who do not. Before providing this framework to a developer it was tested with different self-developed tools to make sure that the algorithms, interval calculations, data exchange interfaces and content selection worked as intended.

For the presented research, different learning strategies that are used mainly in traditional learning scenarios, such a schools or universities, were considered. In those environments learning is usually heavily result-driven. This means that the learners usually have a self-defined goal like passing a test or an exam in mind when they learn. This leads often to less sustainable learning strategies like massing or cramming, which can be summarized as "bulimic learning". These strategies lead to a comparatively quick decline in memorization after the event in which the knowledge was needed and therefore stand in contrast to the spaced repetition approach which aims at building a long-term memorization of knowledge. While the motivating effect of game-based learning may help to mitigate the danger of "bulimic learning" and to keep the learners engaged in the game for a long time, learners in the mentioned traditional learning environments are still very likely to learn the topics on the night before a test or an exam, at least just to give themselves a becalming feeling. However, as long as they keep playing the game (and thus learning through it) afterwards according to the calculated intervals, this fact can be neglected.

The ultimate goal of the presented research was to find out if the use of spaced repetition in a mobile learning game leads to a better learning success compared to refraining from using the calculated intervals according to the spaced repetition approach. To achieve this goal, the algorithmic framework containing the SM2 and the FS algorithm was provided to an external developer of a geography learning game who agreed to provide some anonymized data about how and when his users played the game and if they answered the tasks correctly or wrongly. Before analyzing the data, established strategies about how to measure the learning success from learning analytics were considered. The main problem with measuring learning success is that there is no common definition of this kind of success. In fact, everybody may define learning success in a different way. It was therefore necessary to define learning success for the presented research in order to be able to measure it for the test group who stuck to the calculated learning intervals compared to the success in the control group who did not do this.

Based on this definition and the collected and filtered data it was possible to prove that there is indeed a measurable positive effect on the learning success in the mobile geography learning game when the user decides to stick to the calculated intervals for playing the game. Not only did those players reach a high level of retention for the learning objects, but the retention also stayed on a high level afterwards during the observation period, which denotes a long-term retention of the learned content and therefore the goal behind spaced repetition. The test group as well as the control group started with a base knowledge of about 0.50 based on the answers given for the first task of the observed category. After that, the answers from the 3^{rd} to the n^{th} played game in this category during the observation period were used to calculate the "Average Retention Score" (ARS) for both groups. While the control group reached an ARS of 0.75, the test group achieved a score of 0.94. From this it can be concluded that integrating spaced repetition into a mobile learning game leads to a better learning success.

While the goal for the presented research was achieved, there is still room for some future work. One problem that always exists in scenarios like the one that was examined is the so-called cold-start problem, which was already briefly discussed in the first published paper [75]. This problem describes the challenge where to start with, occurs for example when tasks need to be categorized by their difficulty and is especially problematic when there is no historical data from any users available. The SM2 algorithm tries to solve this problem through fixed intervals for the first two repetitions and by setting the Easiness Factor to an initial value of 2.5. For the content selection by the FS algorithm this problem is even bigger since this selection cannot be based on any initially fixed values. An anonymized data collection within the learning game could be used in the future in order to calculate an average difficulty for categories or tasks and use the result for the content selection of the FS algorithm. However, this data collection would always be limited to a specific learning game.

Furthermore, currently the valuation of the players' performance is simply based on correct and incorrect answers. While this is always the base information, further indicators may be used in order to judge the learners' performance even more detailed. One approach to this could be to analyze the time between the presentation of the content and the learner's answer and then draw conclusions from that about how hard the learner needed to think about his or her answer. However, this time may also be diluted by any short distraction that can occur at any time.

While the presented research used a mobile learning game from the field of factual knowledge to find out if using a spaced repetition algorithm can improve the long-term learning success, this approach may fit for other domains of knowledge and learning as well. One aspect that needs to be addressed even more in this case would be how to judge the learners' performance when it cannot only be valued by right or wrong. This judgement needs to be developed and evaluated for every type of application separately in order to consider the respective requirements.

Another possible future field of research besides classic learning in schools or universities and learning in spare time could also be the usage of the spaced repetition approach with mobile learning games in medical environments. One interesting field in this domain might be Alzheimer's disease, where spaced repetition applications might be a promising solution to mitigate the effects of this disease by providing ongoing repetitions of different topics based on the learner's performance to keep the memory trace alive and to stimulate the learner's memory against forgetting [51].

While the presented research in this thesis could not examine demographic influences and differences due to privacy restrictions, considering this data would also make an interesting field for further research. While the current data analysis is solely descriptive, it could be extended into a predictive approach in future work, which could then be used to verify or refute an established hypothesis.

Bibliography

[1] P. Ackermann and L. Humphreys. Individual differences in theory in industrial and organizational psychology. *In: Handbook of industrial and organizational psychology*, 2nd Edition, Vol. 1:223–282, 1990.

[2] P. L. Ackermann. Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin*, 102:3–27, 1987.

[3] D. Adair. *Building Object-Oriented Frameworks.* AIXpert (February and May), 1995.

[4] A. F. Anderson and D. Bavelier. Action game play as a tool to enhance perception, attention and cognition. In S. Tobias and J. D. Fletcher, editors, *Computer Games and Instruction*, pages 307–330. Information Age, Charllotte, NC, 2011.

[5] L. W. Anderson, D. R. Krathwohl, and P. W. Airasian. *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Allyn and Bacon, Boston, 2000.

[6] R. C. Atkinson and R. M. Shiffrin. Human memory: A proposal system and its control processes. In *The psychology of learning and motivation (Vol. 8).* Academic Press, London, 1968.

[7] K. Backhaus, B. Erichson, W. Plinke, and R. Weiber. *Multivariate Analysemethoden - Eine anwendungsorientierte Einfuehrung.* Springer Gabler, Berlin, 15 edition, 2018.

[8] A. D. Baddeley and D. J. A. Longman. The influence of length and frequency of training session on the rate of learning to type. *In Ergonomics (Vol. 21)*, pages 627–635, 1978.

[9] H. P. Bahrick and E. Phelps. Retention of spanish vocabulary over 8 years. *In: Journal of Experimental Psychology*, 13:344–349, 1987.

[10] F. M. Bienkowski, M. and B. Means. Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. *U.S. Department of Education Office of Educational Technology*, October 2012.

[11] D. Bird and C. Munoz. Automatic generation of random self-checking test cases. *IBM Systems Journal*, 22(3):229–245, 1983.

[12] R. A. Bjork. An information-processing analysis of college teaching. *Educational Psychologist*, 14:15–23, 1979.

[13] B. Bloom. *Taxonomy of educational objectives: The cognitive domain*. Donald McKay, New York, 1956.

[14] J. Brown. Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, 10:12–21, 1958.

[15] F. Buschmann and R. Meunier. A system of patterns. In *Proceedings of the First Conference on Pattern Languages and Programming*. Addison-Wesley, 1994.

[16] N. J. Cepeda et al. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3):354–380, 2006.

[17] F. Craik and R. S. Lockhart. Levels of processing: A framework for memory research. *Journal of Memory and Language*, 11(6):671–684, 1972.

[18] O. Davidsson, J. Peitz, and S. Bjoerk. *Game design patterns for mobile games*. Project report to Nokia ResearchCenter, Finland, 2004. URL `https://www.researchgate.net/publication/221217599_Game_Design_Patterns`. (last accessed on 2022-02-28).

[19] S. de Freitas. *Learning in immersive worlds: A review of game-based learning*. Joint Information Systems Committee (JISC) E-Learning Programme, Bristol, UK, 2007.

[20] E. L. Deci. *Intrinsic motivation*. Plenum Press, New York, 1975.

[21] E. L. Deci, R. Koestner, and R. M. Ryan. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6):627, 1999.

[22] F. N. Dempster. The spacing effect: A case study in the failure to apply the results of psychological research. *In: American Psychologist*, 43:627–634, 1988.

[23] G. Denis and P. Jouvelot. Motivation-driven educational game design: Applying best practices to music education. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pages 462–465, Valencia, Spain, 2005.

[24] L. Dentelo. What does success look like? evaluating success for language learners in the 21st century. November 2017. URL `https://www.cambridge.org/elt/blog/2017/11/13/what-does-success-look-like`. (last accessed on 2022-02-28).

[25] L. Deutsch. Reusability in the smalltalk-80 programming system. In *ITT Proceedings of the Workshop on Reusability in Programming (reprinted in Tutorial on Software Reusability)*, pages 72–76. IEEE Computer Society Press, 1987.

[26] Q. Dunford R., Su and E. Tamang. The pareto principle. *The Plymouth Student Scientist*, 7(1):140–148, 2014.

[27] H. Ebbinghaus. *Memory: A Contribution to Experimental Psychology.* Dover, New York, 1885.

[28] F. et al. *Statistik – Der Weg zur Datenanalyse.* Springer Spektrum, Berlin, 8 edition, 2016.

[29] L. Fahrmeir, T. Kneib, and S. Lang. *Regression – Modelle, Methoden und Anwendungen.* Springer, Berlin, 2 edition, 2009.

[30] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 1996.

[31] R. M. Gagne. *The conditions of learning.* Holt, Rinehart & Winston, New York, 1977.

[32] R. M. Gagne. Learning outcomes and their effects: Useful categories of human performance. *American Psychologist*, 39:377–385, 1984.

[33] R. Garris, R. Ahlers, and J. E. Driskell. Games, motivation and learning, simulation & gaming. *An Interdisciplinary Journal of Theory, Practice and Research*, 33(4):441–467, December 2002.

[34] J. P. Gee. *What Video Games Have to Teach us About Learning and Literacy.* Palgrave/Macmillan, New York, 2003.

[35] I. L. Goldstein. *Training in organizations: Needs assessment, design, and evaluation.* Brooks/Cole, Monterey, CA, 1986.

[36] D. Gopher, M. Weil, and T. Bareket. Transfer of skill from a computer game trainer to flight. *Human Factors*, 36:387–405, 1994.

[37] R. Hatzinger, K. Hornik, H. Nagel, and M. J. Maier. *R – Einführung durch angewandte Statistik.* Pearson, München, 2014.

[38] S. Haug, R. P. Castro, M. Kwon, A. Filler, T. Kowatsch, and M. P. Schaub. Smartphone use and smartphone addiction among young people in switzerland. *Journal of Behavioral Addictions*, 4(4):299–307, 2015. URL `https://doi.org/10.1556/2006.4.2015.037`. (last accessed on 2022-02-28).

[39] D. O. Hebb. *Organizations of Behavior.* Wiley, New York, 1949.

[40] A. Heritage. *Dictionary of the English Language*, volume 4. Houghton Mifflin Harcourt, Boston, 2000.

[41] S. G. V. (Hrsg.). Gabler wirtschaftslexikon. *Stichwort: Akkreditiv*, 2018. URL `https://wirtschaftslexikon.gabler.de/definition/swot-analyse-52664/version-275782`. (last accessed on 2022-02-28).

[42] J. M. Juran. The non-pareto principle; mea culpa. *Quality Progress*, 8:8–9, 1975.

[43] J. W. Kalat. *Biological Psychology.* Cengage Learning, 2013.

[44] D. L. Kirkpatrick. Evaluation of training. In R. L. Craig, editor, *Training and development handbook: A guide to human resource development*, pages 301–319. McGraw-Hill, New York, 1987.

[45] U. Kockelkorn. *Statistik für Anwender*. Springer Spektrum, Berlin, 2012.

[46] N. Kornell. Optimizing learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, 23(9):1297–1317, 2009.

[47] N. Kornell and R. A. Bjork. Optimizing self-regulated study: The benefits and costs of dropping flashcards. *Memory*, 16(2):125–136, 2008.

[48] N. Kornell and R. A. Bjork. Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science*, 19:585–592, 2008.

[49] K. Kraiger, K. Ford, and E. Salas. Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology*, 78(2):311–328, 1993.

[50] Y. Laouris and N. Eteokleous. We need an educationally relevant definition of mobile learning. In *Proceedings of 4th World Conference on Mobile Learning mLearn 2005*, pages 1–13, Cape Town, South Africa, 2005.

[51] S. B. Lee, C. S. Park, J. W. Jeong, J. Y. Choe, Y. J. Hwang, C.-A. Park, J. H. Park, D. Y. Lee, J. H. Jhoo, and K. W. Kim. Effects of spaced retrieval training (srt) on cognitive function in alzheimer's disease (ad) patients. *Archives of gerontology and geriatrics*, 49(2):289–293, 2009.

[52] S. Leitner. *So lernt man lernen. in Angewandte Lernpsychologie – ein Weg zum Erfolg*. Verlag Herder, Freiburg im Breisgau, Germany, 1972.

[53] T. W. Malone and M. R. Lepper. Making learning fun: a taxonomic model of intrinsic motivations for learning. *In Aptitude, Learning, and Instruction: III. Cognitive and Affective Process Analysis (eds R.E. Snow and M.J. Farr)*, pages 223–253, 1987.

[54] J. Mayer, I. Melzer, and F. Schweiggert. Lightweight plug-in-based application development. In *Net. ObjectDays: International Conference on Object-Oriented and Internet-Based Technologies, Concepts, and Applications for a Networked World*, pages 98–102. Springer, 2002.

[55] P. Murphy and R. McCormick. *Knowledge and Practice: Representations and Identities*. SAGE Publications Ltd, Thousand Oaks, 2008.

[56] T. O. Nelson, J. Dunlosky, A. Graf, and L. Narens. Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science*, 5:207–213, 1994.

[57] J. Nielsen. *Usability Engineering*. Academic Press, Boston, 1993.

[58] R. A. Noe. Trainees' attributes and attitudes: Neglected influences on training effectiveness. *Academy of Management Review*, 11:736–749, 1986.

[59] M. N. Norazah, A. E. Mohamed, and A. W. Zaidan. Integration of e-learning in teaching & learning in malaysian higher education institutions. *e-Learning in Malaysian Higher Education Institutions: Status, Trends & Challenges*, pages 81–98, 2011.

[60] D. Parra-Santander and P. Brusilovsky. Improving collaborative filtering in social tagging systems for the recommendation of scientific articles. In *Proceedings of 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology Toronto*, pages 136–142. IEEE, 2010.

[61] K. Paul. *Study Smarter, Not Harder.* Self-Counsel Press, 2007.

[62] J. P. J. Pinel. *Biopsychology.* Allyn and Bacon, Boston, 1993.

[63] D. Poirier, F. Fessant, and I. Tellier. Reducing the cold-start problem in content recommendation through opinion classification. In *Proceedings of 2010 IEEE/WIC/ACM International Conference on Web Intelligence*, pages 204–207, Toronto, Canada, 2010.

[64] G. J. Posner and K. A. Strike. A categorization scheme for principles of sequencing content. *Review of Educational Research*, 46(4):665–690, 1976. doi: 10.3102/00346543046004665.

[65] R. Remes. Learning management system. In *WDS'05 Proceedings of Contributed Papers*, pages 207–212, 2005.

[66] L. J. Rennie and D. J. Johnston. The nature of learning and its implications for research on learning from museums. *Science Education*, 88:4–16, 2004.

[67] H. L. Roediger and J. D. Karpicke. The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3):181–210, 2006.

[68] R. M. Ryan and E. L. Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55:68–78, 2000.

[69] F. Schimanke. The impact of spaced repetition learning on the learning success in mobile learning games. *IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2021)*, pages 275–280, November 2021. doi: 10.1109/ISM52913.2021.00054.

[70] F. Schimanke and R. Mertens. Spaced repetition in mobile learning games - a cure to bulimic learning? In *AACE E-Learn World Conference on E-Learning (E-Learn 2017)*, pages 955–964, Vancouver, CAN, October 2017. AACE.

[71] F. Schimanke and R. Mertens. Mobile game-based learning in the app-age – where we are and where want to be. In *AACE E-Learn World Conference on E-Learning (E-Learn 2018)*, pages 1321–1330, Las Vegas, NV, October 2018. AACE.

[72] F. Schimanke and R. Mertens. Measuring the learning success in spaced-repetition-based mobile learning-games: Suitable techniques and hurdles. In *AACE E-Learn World Conference on E-Learning (E-Learn 2019)*, pages 528–537, New Orleans, LA, November 2019. AACE.

[73] F. Schimanke and R. Mertens. Deriving strategies for the evaluation of spaced repetition learning in mobile learning applications from learning analytics. *IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2020)*, pages 239–244, December 2020. doi: 10.1109/ISM.2020.00049.

[74] F. Schimanke, R. Mertens, S. Vollmer, and O. Vornberger. Multi category content selection in spaced repetition based mobile learning games. *IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2013)*, pages 468–473, December 2013. doi: 10.1109/ISM.2013.90.

[75] F. Schimanke, R. Mertens, and O. Vornberger. What to learn next? content selection support in mobile game-based learning. In *AACE E-Learn World Conference on E-Learning (E-Learn 2013)*, pages 2503–2512, Las Vegas, NV, October 2013.

[76] F. Schimanke, R. Mertens, and O. Vornberger. Designing for motivation: Design-considerations for spaced-repetition-based learning games on mobile devices. In *AACE E-Learn World Conference on E-Learning (E-Learn 2014)*, pages 1770–1779, New Orleans, LA, October 2014.

[77] F. Schimanke, R. Mertens, and O. Vornberger. Architecture considerations for spaced repetition based mobile learning games on ios. *IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2014)*, pages 363–368, December 2014. doi: 10.1109/ISM.2014.17.

[78] F. Schimanke, R. Mertens, F. Hallay, A. Enders, and O. Vornberger. Using a spaced-repetition-based mobile learning game in database lectures. In *AACE E-Learn World Conference on E-Learning (E-Learn 2015)*, pages 1610–1619, Kona, HI, October 2015. AACE.

[79] F. Schimanke, S. Ribbers, R. Mertens, and O. Vornberger. Implications of short term memory research for the design of spaced repetition based mobile learning games. *IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2015)*, pages 571–576, December 2015. doi: 10.1109/ISM.2015.13.

[80] F. Schimanke, R. Mertens, and J. Hillebrand. Enhancing mobile learning games with spaced-repetition and content-selection algorithms. In *AACE E-Learn World Conference on E-Learning (2016)*, pages 1270–1279, Washington, DC, November 2016. AACE.

[81] F. Schimanke, R. Mertens, O. Vornberger, and L. Hill. Simulating context in mobile learning games for testing and debugging. *IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2016)*, pages 655–660, December 2016. doi: 10.1109/ISM.2016.0140.

[82] F. Schimanke, L. Hill, and R. Mertens. A unit testing framework for context variant code in a mobile learning app. *IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2017)*, pages 577–582, December 2017. doi: 10.1109/ISM.2017.113.

[83] F. Schimanke, R. Mertens, and B. S. Huck. Player types in mobile learning games – playing patterns and motivation. *IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2018)*, pages 247–252, December 2018. doi: 10.1109/ISM.2018.00035.

[84] F. Schimanke, R. Mertens, and S. Huck. Retrieval of relevant data for measuring the impact of spaced repetition algorithms on the learning success in mobile learning games. *IEEE Multimedia Technologies for E-Learning (MTEL-Workshop held in conjunction with IEEE ISM 2019)*, pages 279–284, December 2019. doi: 10.1109/ISM46123.2019.00063.

[85] S. Shaghayegh and W. Cohen. Community-based recommendation: A solution to the cold start problem. In *Workshop on Recommender Systems and the Social Web (RSWEB)*, pages 40–44, Chicago, October 2011.

[86] G. Siemens and P. Long. Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5), September/October 2011.

[87] V. K. Sims and R. E. Mayer. Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology*, 16:97–115, 2002.

[88] Statista. (last accessed on 2022-02-28). URL https://www.statista.com/statistics/270291/popular-categories-in-the-app-store.

[89] Y.-T. Sung, K.-E. Chang, and M.-D. Lee. Designing multimedia games for young children's taxonomic concept development. *Computers and Education*, 50(3):1037–1051, 2008.

[90] C. Szyperski, D. Gruntz, and S. Murer. *Component Software. Beyond object-oriented programming.* Addison-Wesley, London, 2002.

[91] S. I. Tannenbaum, J. E. Mathieu, E. Salas, and J. B. Cannon-Bowers. Meeting trainees' expectations: The influence of training fulfillment on the development of commitment, self-efficacy, and motivation. *Journal of Applied Psychology*, 76:759–769, 1991.

[92] P. Torfs and C. Brauer. A (very) short introduction to r. 2014. URL https://cran.r-project.org/doc/contrib/Torfs+Brauer-Short-R-Intro.pdf. (last accessed on 2022-02-28).

[93] E. Trondsen and K. Vickery. Learning on demand. *Journal of Knowledge Management*, pages 169–180, 1997.

[94] G. Tuparov and D. Tuparova. Didactical issues of e-learning- problems and future trends. In *Proceedings of International Conference on Computer Systems and Technologies – CompSysTech*, 2005.

[95] A. Ultsch and J. Lötsch. Computed abc analysis for rational selection of most informative variables in multivariate data. *PLoS ONE*, 10(6), 2015. URL `https://doi.org/10.1371/journal.pone.0129767`. (last accessed on 2022-02-28).

[96] P. A. Wozniak. 1990. URL `https://www.supermemo.com/de/archives1990-2015/english/ol/sm2`. (last accessed on 2022-02-28).

[97] P. A. Wozniak. 1994. URL `http://www.supermemo.com/english/princip.htm`. (last accessed on 2022-02-28).

[98] P. A. Wozniak. 2018. URL `https://supermemo.guru/wiki/SuperMemo_1.0_for_DOS_(1987)#Algorithm_SM-2`. (last accessed on 2022-02-28).

[99] P. A. Wozniak. 2019. URL `https://supermemo.guru/wiki/Algorithm_SM-18`. (last accessed on 2022-02-28).

Part II

Appendix

Appendix A

Full questionnaire results

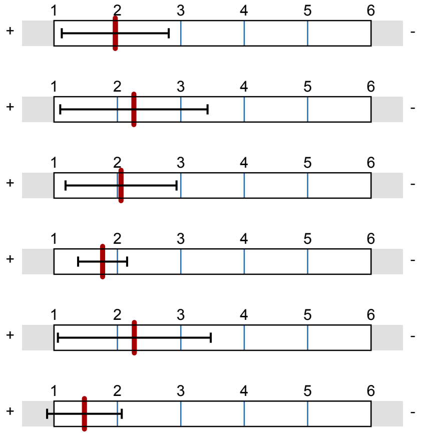
A.1 ER Maker (Database learning game)

The following pages contain the full results from the questionnaire among the participating students using the ER Maker database learning game (Chapter 3.2.3 and paper "Using a spaced-repetition-based mobile learning game in database lectures").

Evaluation ER-Maker

Schimanke ()
Erfasste Fragebögen = 10

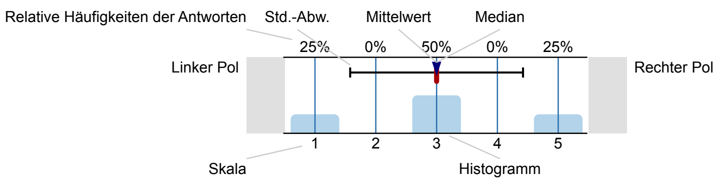
Globalwerte



Globalindikator	mw=2 s=0,8
2. Spaced Repetition	mw=2,3 s=1,2
3. ER-Maker	mw=2,1 s=0,9
4. Inhalte	mw=1,8 s=0,4
5. Mobile App	mw=2,3 s=1,2
6. Sonstiges	mw=1,5 s=0,6

Legende

Fragetext



n=Anzahl
mw=Mittelwert
md=Median
s=Std.-Abw.
E.=Enthaltung

1. Allgemeine Daten

1.1) Geschlecht

n=10

männlich	50%
weiblich	50%

1.2) Alter

- 18
- 19 (3 Nennungen)
- 20
- 21 (2 Nennungen)
- 22
- 23
- 27

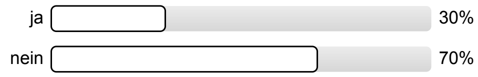
2. Spaced Repetition

2.1) War Ihnen das Konzept der "spaced repetitions" zuvor bereits bekannt?

n=10

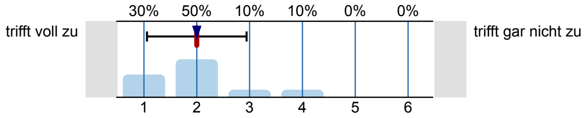
ja	30%
nein	70%

2.2) Haben Sie sich immer an die von der App vorgegebenen Intervalle gehalten?



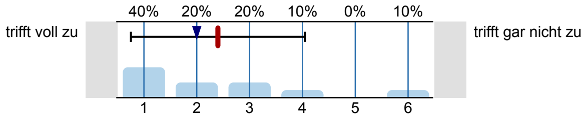
n=10

2.3) Haben Sie das Gefühl, dass Sie durch "spaced repetitions" effizienter gelernt haben?



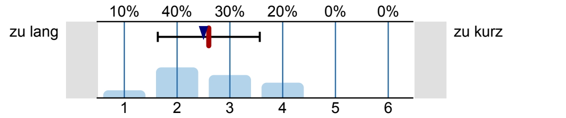
n=10
mw=2
md=2
s=0,9

2.4) Haben Sie das Gefühl, dass Sie durch "spaced repetitions" besser gelernt haben?



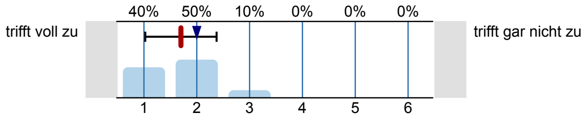
n=10
mw=2,4
md=2
s=1,6

2.5) Erschienen Ihnen die Intervalle zwischen den Wiederholungen der einzelnen Aufgaben...



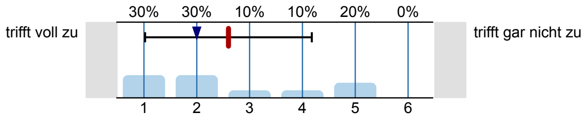
n=10
mw=2,6
md=2,5
s=1

2.6) Denken Sie, dass "spaced repetitions" (über einen längeren Zeitraum) das Langzeitgedächtnis fördern?



n=10
mw=1,7
md=2
s=0,7

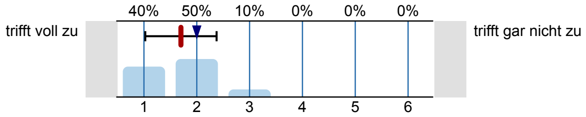
2.7) Halten Sie "spaced repetitions" allgemein für eine gute Form der Klausurvorbereitung?



n=10
mw=2,6
md=2
s=1,6

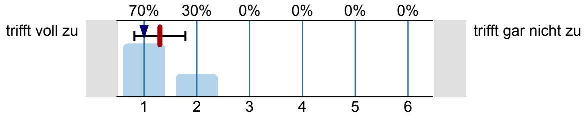
3. ER-Maker

3.1) Hat Ihnen das Spiel grundsätzlich Spaß gemacht?



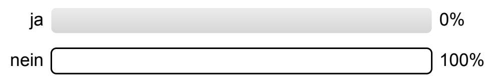
n=10
mw=1,7
md=2
s=0,7

3.2) War das Spiel intuitiv bedienbar?



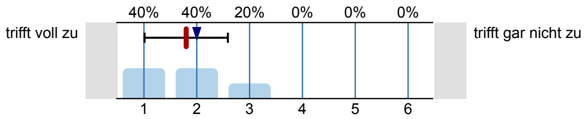
n=10
mw=1,3
md=1
s=0,5

3.3) Hatten Sie technische Schwierigkeiten mit dem Spiel?



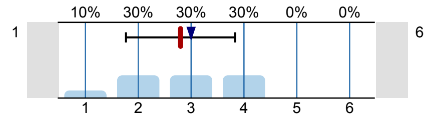
n=10

3.4) Empfinden Sie die Verwendung eines solchen Spiels grundsätzlich als motivationsfördernd?



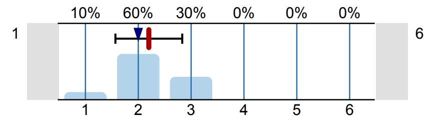
n=10
mw=1,8
md=2
s=0,8

3.5) Grafische Gestaltung



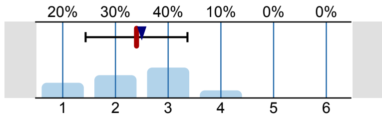
n=10
mw=2,8
md=3
s=1

3.6) Formulierung der Aufgaben



n=10
mw=2,2
md=2
s=0,6

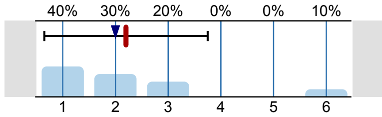
3.7) Hinweise zur Einhaltung der Intervalle

1


6

n=10
mw=2,4
md=2,5
s=1

3.8) Hat Ihnen das Spiel grundsätzlich dabei geholfen, die zugrundeliegenden Datenbankkonzepte besser zu verstehen?

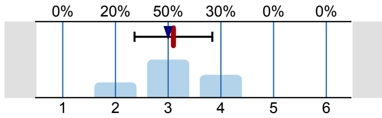
trifft voll zu


trifft gar nicht zu

n=10
mw=2,2
md=2
s=1,5

4. Inhalte

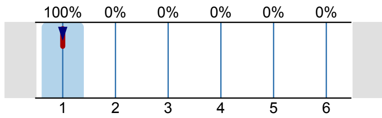
4.1) War die Variation der Inhalte innerhalb des Spiels für die Vertiefung der Konzepte ausreichend?

trifft voll zu


trifft gar nicht zu

n=10
mw=3,1
md=3
s=0,7

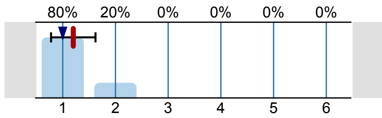
4.2) Wäre ein umfangreicherer Aufgabenkatalog hilfreicher beim Lernen der Konzepte?

trifft voll zu


trifft gar nicht zu

n=10
mw=1
md=1
s=0

4.3) Würde ein umfangreicherer Aufgabenkatalog die Motivation zur Nutzung der App steigern?

trifft voll zu


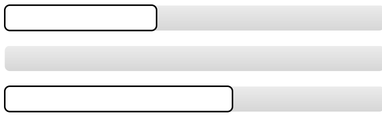
trifft gar nicht zu

n=10
mw=1,2
md=1
s=0,4

4.4) Waren die gestellten Aufgaben...

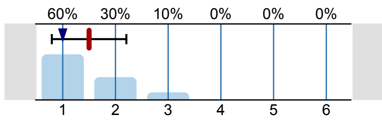
n=10

zu leicht 40%
zu schwierig 0%
genau richtig 60%



5. Mobile App

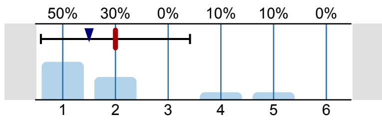
5.1) Halten Sie "spaced repetitions" für die Nutzung auf einem mobilen Endgerät (Tablet oder Smartphone) für sinnvoll?

trifft voll zu


trifft gar nicht zu

n=10
mw=1,5
md=1
s=0,7

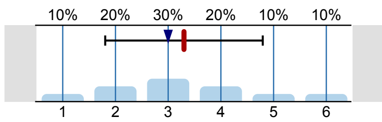
5.2) Halten Sie die Display-Benachrichtigungen bei Erreichen eines errechneten Lerntages für sinnvoll?

trifft voll zu


trifft gar nicht zu

n=10
mw=2
md=1,5
s=1,4

5.3) Haben Ihnen die Benachrichtigungen auf dem Bildschrim dabei geholfen, sich an die berechneten Lernintervalle zu halten?

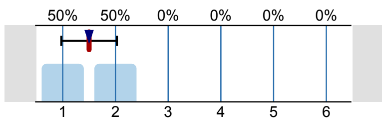
trifft voll zu


trifft gar nicht zu

n=10
mw=3,3
md=3
s=1,5

6. Sonstiges

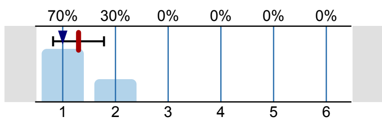
6.1) Gutes User Interface Design

trifft voll zu


trifft gar nicht zu

n=10
mw=1,5
md=1,5
s=0,5

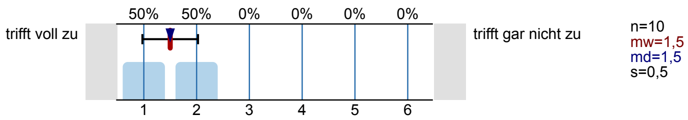
6.2) Potenzial zur Motivationssteigerung

trifft voll zu


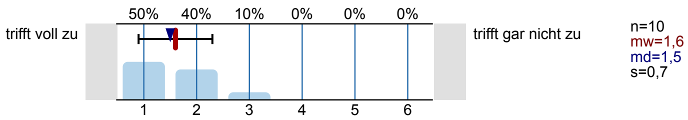
trifft gar nicht zu

n=10
mw=1,3
md=1
s=0,5

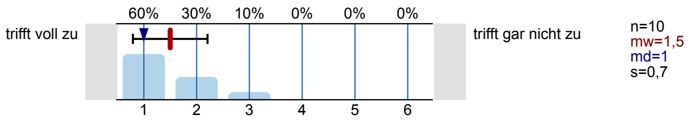
6.3) Gute Visualisierung



6.4) Einfache Navigation



6.5) Benutzerfreundlichkeit



6.6) Haben Sie abschließend noch Anregungen, Kritik, Verbesserungsvorschläge, mögliche weitere Inhalte, etc. für uns? Dann lassen Sie es uns wissen...

- Bei falsch beantworteten Fragen wäre das zeigen der richtigen Antwort mit einer Erklärung warum diese richtig ist noch sinnvoll. Dadurch lernt man gleich mehr und weiß genau was falsch war und weiß nicht irgendwann die Lösungen auswendig, sondern kann sogar sagen warum es richtig ist.

- Eine sehr gute Lösung um Wissen zu vermitteln und zu lernen. Die App hat mir sehr geholfen und ich hoffe, dass so eine App an Hochschulen zur Verfügung gestellt werden.

- Es wäre sehr hilfreich, dass wenn eine Antwort falsch ausgewählt wird, warum diese falsch ist. Denn beim antippen einer falschen Antwort wird dann zwar das richtige Ergebnis angezeigt, aber ich weiß nicht warum, was dann für mich schwierig wurde die richtige Antwort bzw meine falsche Antwort nachvollzuziehen.

  Ansonsten finde ich die App sehr gut - vorallem zur Unterstützung für die Klausurvorbereitung eigent sich so eine App.
  Vielen Dank :)

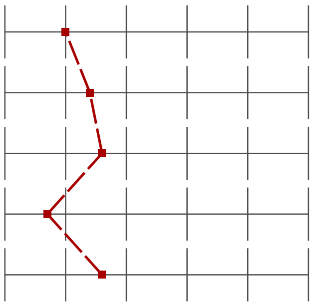
- Grundsätzlich eine sehr gute Variante der Wissensvermittlung. Leider waren die Aufgaben aus meiner Sicht zu einfach und zu wenig abwechslungsreich. Hier ist noch Luft nach oben. Ebenso, wie bei der grafischen Gestaltung. Ansonsten ein gutes Konzept, welches weiter verfolgt werden sollte.

Profillinie

Teilbereich: Tests/sonstige Umfragen_neu
Name der/des Lehrenden: Schimanke
Titel der Lehrveranstaltung: (Name der Umfrage) Evaluation ER-Maker

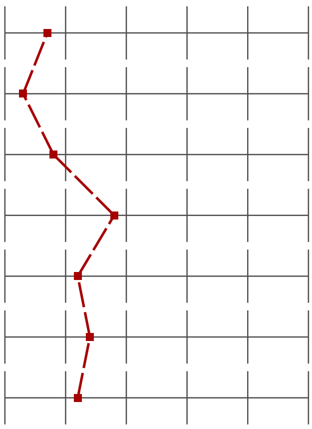
Verwendete Werte in der Profillinie: Mittelwert

2. Spaced Repetition



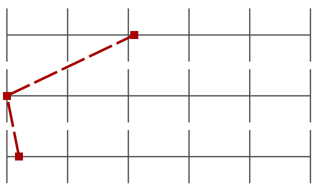
2.3)	Haben Sie das Gefühl, dass Sie durch "spaced repetitions" effizienter gelernt haben?	trifft voll zu	trifft gar nicht zu	n=10	mw=2,0	md=2,0	s=0,9
2.4)	Haben Sie das Gefühl, dass Sie durch "spaced repetitions" besser gelernt haben?	trifft voll zu	trifft gar nicht zu	n=10	mw=2,4	md=2,0	s=1,6
2.5)	Erschienen Ihnen die Intervalle zwischen den Wiederholungen der einzelnen Aufgaben...	zu lang	zu kurz	n=10	mw=2,6	md=2,5	s=1,0
2.6)	Denken Sie, dass "spaced repetitions" (über einen längeren Zeitraum) das Langzeitgedächtnis fördern?	trifft voll zu	trifft gar nicht zu	n=10	mw=1,7	md=2,0	s=0,7
2.7)	Halten Sie "spaced repetitions" allgemein für eine gute Form der Klausurvorbereitung?	trifft voll zu	trifft gar nicht zu	n=10	mw=2,6	md=2,0	s=1,6

3. ER-Maker



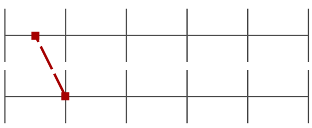
3.1)	Hat Ihnen das Spiel grundsätzlich Spaß gemacht?	trifft voll zu	trifft gar nicht zu	n=10	mw=1,7	md=2,0	s=0,7
3.2)	War das Spiel intuitiv bedienbar?	trifft voll zu	trifft gar nicht zu	n=10	mw=1,3	md=1,0	s=0,5
3.4)	Empfinden Sie die Verwendung eines solchen Spiels grundsätzlich als motivationsfördernd?	trifft voll zu	trifft gar nicht zu	n=10	mw=1,8	md=2,0	s=0,8
3.5)	Grafische Gestaltung	1	6	n=10	mw=2,8	md=3,0	s=1,0
3.6)	Formulierung der Aufgaben	1	6	n=10	mw=2,2	md=2,0	s=0,6
3.7)	Hinweise zur Einhaltung der Intervalle	1	6	n=10	mw=2,4	md=2,5	s=1,0
3.8)	Hat Ihnen das Spiel grundsätzlich dabei geholfen, die zugrundeliegenden Datenbankkonzepte besser zu verstehen?	trifft voll zu	trifft gar nicht zu	n=10	mw=2,2	md=2,0	s=1,5

4. Inhalte



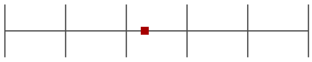
4.1)	War die Variation der Inhalte innerhalb des Spiels für die Vertiefung der Konzepte ausreichend?	trifft voll zu	trifft gar nicht zu	n=10	mw=3,1	md=3,0	s=0,7
4.2)	Wäre ein umfangreicherer Aufgabenkatalog hilfreicher beim Lernen der Konzepte?	trifft voll zu	trifft gar nicht zu	n=10	mw=1,0	md=1,0	s=0,0
4.3)	Würde ein umfangreicherer Aufgabenkatalog die Motivation zur Nutzung der App steigern?	trifft voll zu	trifft gar nicht zu	n=10	mw=1,2	md=1,0	s=0,4

5. Mobile App



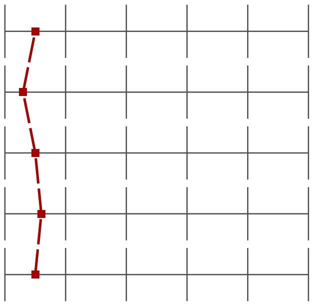
5.1)	Halten Sie "spaced repetitions" für die Nutzung auf einem mobilen Endgerät (Tablet oder Smartphone) für sinnvoll?	trifft voll zu	trifft gar nicht zu	n=10	mw=1,5	md=1,0	s=0,7
5.2)	Halten Sie die Display-Benachrichtigungen bei Erreichen eines errechneten Lerntages für sinnvoll?	trifft voll zu	trifft gar nicht zu	n=10	mw=2,0	md=1,5	s=1,4

5.3)	Haben Ihnen die Benachrichtigungen auf dem Bildschrim dabei geholfen, sich an die berechneten Lernintervalle zu halten?	trifft voll zu	trifft gar nicht zu	n=10	mw=3,3	md=3,0	s=1,5



6. Sonstiges

6.1)	Gutes User Interface Design	trifft voll zu	trifft gar nicht zu	n=10	mw=1,5	md=1,5	s=0,5
6.2)	Potenzial zur Motivationssteigerung	trifft voll zu	trifft gar nicht zu	n=10	mw=1,3	md=1,0	s=0,5
6.3)	Gute Visualisierung	trifft voll zu	trifft gar nicht zu	n=10	mw=1,5	md=1,5	s=0,5
6.4)	Einfache Navigation	trifft voll zu	trifft gar nicht zu	n=10	mw=1,6	md=1,5	s=0,7
6.5)	Benutzerfreundlichkeit	trifft voll zu	trifft gar nicht zu	n=10	mw=1,5	md=1,0	s=0,7



A.2 Learning strategies among HSW students

The following pages contain the full results from the questionnaire among HSW students about their learning strategies (Chapter 3.2.4 and paper "Spaced repetition in mobile learning games – A cure to bulimic learning?").

Evaluation zum Thema "Lernstrategien"

Schimanke ()
Erfasste Fragebögen = 242

Globalwerte

1. Allgemeine Daten

1.1) Geschlecht

n=240

männlich	71.7%
weiblich	28.3%

1.2) Alter

- 18 (7 Nennungen)
- 19 (29 Nennungen)
- 20 (84 Nennungen)
- 21 (46 Nennungen)
- 22 (22 Nennungen)
- 23 (16 Nennungen)
- 24 (10 Nennungen)
- 25 (5 Nennungen)
- 26 (8 Nennungen)
- 27 (2 Nennungen)
- 28 (2 Nennungen)
- 29
- 30
- 31
- 32 (2 Nennungen)
- 38

1.3) Aktuelles Semester an der HSW

n=242

1	0.8%
2	30.2%
3	2.5%
4	46.3%
5	2.9%
6	17.4%

2. Lernen

2.1) Wann lernen Sie?

Ich lerne meist das gesamte Semester über.	37.2%	n=242
Ich beginne mit dem Lernen meist ca. eine Woche vor einer Prüfung.	44.6%	
Ich beginne mit dem Lernen meist wenige Tage vor einer Prüfung.	18.2%	

2.2) Wie lernen Sie?

Ich lerne meist alleine.	55.8%	n=242
Ich lerne meist in Lerngruppen.	5.4%	
Ich lerne sowohl alleine, als auch in Lerngruppen.	38.8%	

2.3) Verändert sich Ihre Lernintensität während des Semsters?

Die Lernintensität ist gleichbleibend.	7.1%	n=241
Die Lernintensität wird zum Ende hin höher.	92.9%	

2.4) Womit lernen Sie?

Lernzettel	84.3%	n=242
Karteikarten	28.1%	
Lernspiele	1.2%	
Vorlesungsaufzeichnungen	43%	
Vorlesungsunterlagen	83.9%	
Andere Hilfsmittel	25.6%	

2.5) Machen Sie Pausen während des Lernens?

Ich lerne längere Zeit am Stück ohne Pausen.	19.1%	n=241
Ich mache während des Lernens öfter kurze Pausen.	55.6%	
Ich mache während des Lernens öfter längere Pausen.	25.3%	

2.6) Sind Sie mit bestimmten Lerntechniken vertraut (z.B. Spacing, Massing, Cramming, etc.)? Falls ja, nennen Sie diese bitte im folgenden Freitextfeld.

- "Dozentenansatz": Meinem Sofa die klausurrelevanten Inhalte so verständlich wie möglich darlegen.
  Für Auswendiglernen: Jede Seite einmal lesen, erst den ersten Stichpunkt, dann den erste+zweiten etc auswendig lernen.
- /
- 50-10-50-30 Methode
- Brechreizlernen
- Ich mach Youtubing :D
- Lernen mit mimik und gestik, mit gedächtnis von körper.
  assoziationen
- Loki-Methode
- Manchmal mache ich Lieder aus den Inhalten und kann sie so besser merken
- Nein (3 Nennungen)
- Pomodoro-Technik, Lerntagebuch und Lernjornal
- Spacing mit Karteikasten
- Verstehen; logische Brücken bauen und daraus ableiten
- Vorrangig Eselsbrücken

Und wiederholtes aufschreiben

- nein (2 Nennungen)

2.7) Nutzen Sie bewusst Lerntechniken?

Nein, ich versuche möglichst viel zu behalten, indem ich die Inhalte immer wieder	94.2%	n=242
Ja, ich lerne nach dem Spacing-Prinzip.	0.8%	
Ja, ich lerne nach dem Leitner-Konzept.	0.8%	
Ja, und zwar die folgende(n)	3.7%	

2.8)

- Eigene Fragen formulieren und versuchen diese zu beantworten.
- Eigene Lerntechniken
- Eigentlich nutze ich keine Lerntechnik.
  Die Antwort passt aber auch nicht zur ersten Auswahlmöglichkeit.

  Ich wiederhole nicht großartig, sondern gucke mir die Vorlesungsunterlagen an. Im Allgemeinen reicht das aus um das wichtigste zu behalten. Manchmal versuche ich noch Kausalitätsketten zu bilden oder die Dinge eben einfach zu verstehen. Alles was darüber hinaus geht ist meistens nur Detailwissen. Dieses ist unwichtig, da man es spätestens nach einer Woche sowieso vergisst. Darüber hinaus ist dieses Wissen in dieser Tiefe in den Klausuren zumeist nicht gefragt.
- Ich versuche alle Inhalte möglichst zu visualisieren. Vor jeder Prüfung wird auch immer mindestens eine Mind-Map angefertigt mit möglichst strukturierter Farbgebung.
- Loki-Methode
- Nach dem Elevator Pitch Verfahren. Sobald ich jemanden Fremden das Thema erklären kann, habe ich es verstanden.
- Pomodoro-Technik

  die Aufgabe schriftlich formulieren
  den Kurzzeitwecker auf 25 Minuten stellen
  die Aufgabe bearbeiten, bis der Wecker klingelt; mit einem X markieren
  kurze Pause machen (5 Minuten)
  alle vier pomodori eine längere Pause machen (15 Minuten)
- Singen, Geschichten aus Lernstoff machen um es zu verstehen, anderen (Unbeteiligten) beibringen, was ich lernen muss
- ich weiß nicht,wie diese alle techniken in psychologie heißen, ich zeichne viel, lerne mit assoziationen und mit gesten, klebe notizzettelchen in der küche . lernzettel mit Stichwörter und Pfeilen um zusammenhänge abzubilden. usw

3. Nachhaltigkeit des Lernens

3.1) Welches Ziel verfolgen Sie beim Lernen während des Semesters? (max. 2 Antworten möglich)

Gute Note erzielen.	81.8%	n=242
Modul bestehen.	47.9%	
Erwerb von nachhaltigem Wissen.	38.4%	

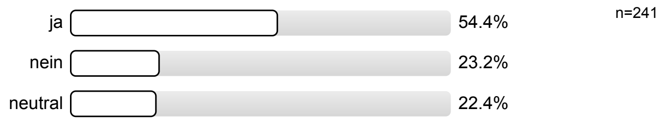
3.2) Wie schätzen Sie den Lerneffekt Ihrer Strategie ein?

Zielgerichtet auf eine Prüfung/Klauaur.	85.5%	n=241
Dauerhaft auch nach Abschluss des Moduls.	14.5%	

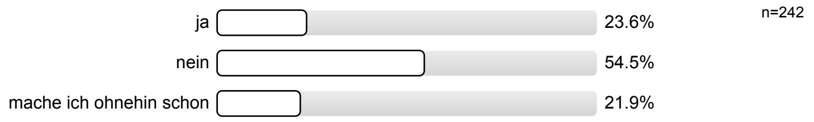
3.3) Denken Sie, dass Ihre im Studium verwendete Lernstrategie nachhaltig ist?

ja	20.4%	n=240
nein	35.8%	
neutral	43.8%	

3.4) Vermittelt Ihnen das Lernen am Abend vor einer Klausur ein gutes/beruhigendes Gefühl?



3.5) Würden Sie auf das Lernen am Abend vor einer Klausur verzichten, wenn Sie zuvor bereits längere Zeit für das Thema gelernt hätten?



A.3 Use of learning games - Random people

The following pages contain the full results from the questionnaire among random participants about their use of learning games (Chapter 3.3.4 and paper "Mobile Game-Based Learning in the App-Age – Where we are and where we want to be").

Nutzung mobiler Lernspiele

Ergebnisse

Umfrage 133457

Anzahl der Datensätze in dieser Abfrage:	314
Gesamtzahl der Datensätze dieser Umfrage:	314
Anteil in Prozent:	100.00%

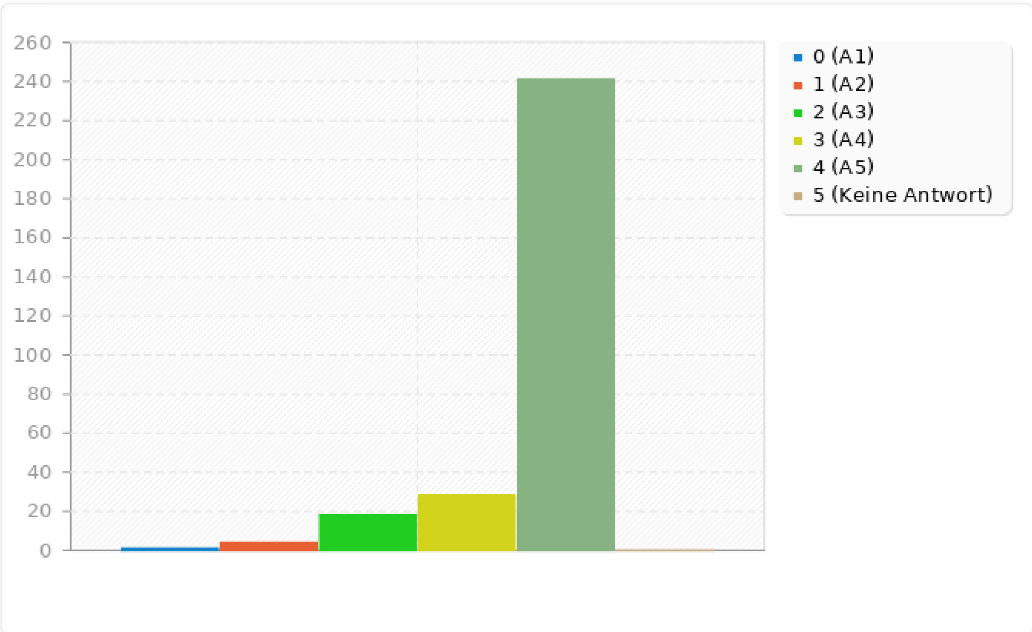
Zusammenfassung für Alter

Alter

Antwort	Anzahl	Prozent
8 bis 14 Jahre (A1)	1	0.34%
15 bis 18 Jahre (A2)	4	1.37%
19 bis 24 Jahre (A3)	18	6.16%
25 bis 30 Jahre (A4)	28	9.59%
älter als 30 Jahre (A5)	241	82.53%
Keine Antwort	0	0.00%

Zusammenfassung für Alter

Alter



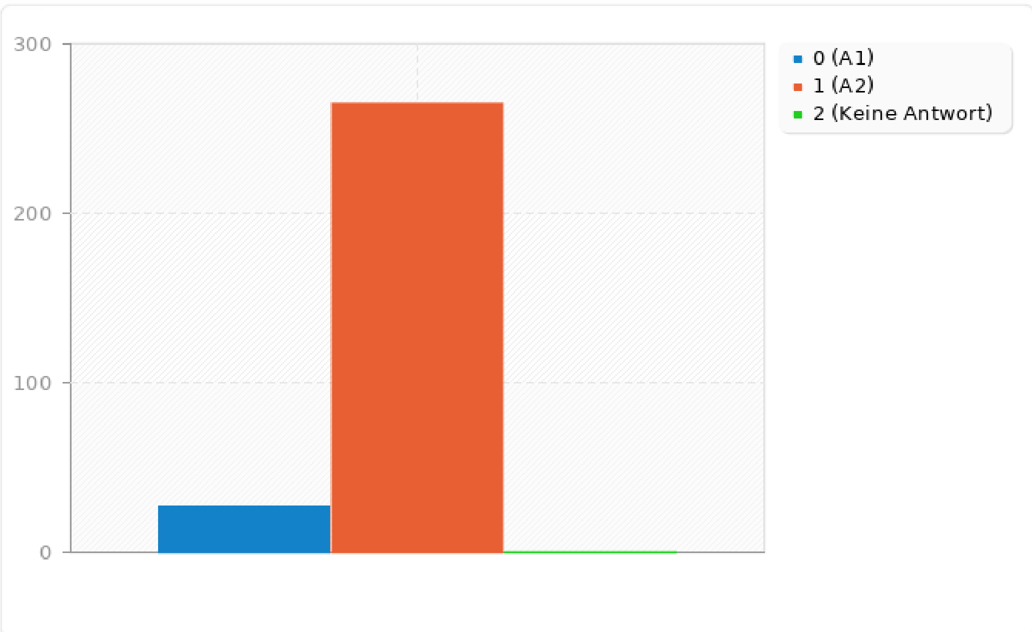
Zusammenfassung für Geschlecht

Geschlecht

Antwort	Anzahl	Prozent
weiblich (A1)	27	9.25%
männlich (A2)	265	90.75%
Keine Antwort	0	0.00%

Zusammenfassung für Geschlecht

Geschlecht



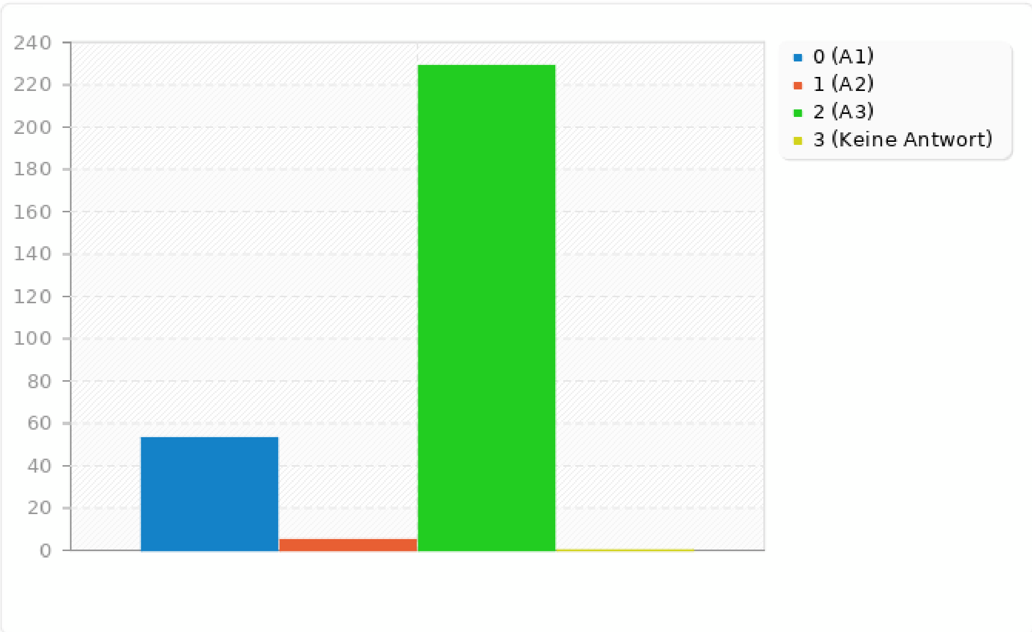
Zusammenfassung für Geraet

Welches mobile Gerät besitzen und benutzen Sie regelmäßig?

Antwort	Anzahl	Prozent
Smartphone (A1)	53	18.47%
Tablet (A2)	5	1.74%
beides (A3)	229	79.79%
Keine Antwort	0	0.00%

Zusammenfassung für Geraet

Welches mobile Gerät besitzen und benutzen Sie regelmäßig?



Zusammenfassung für OS

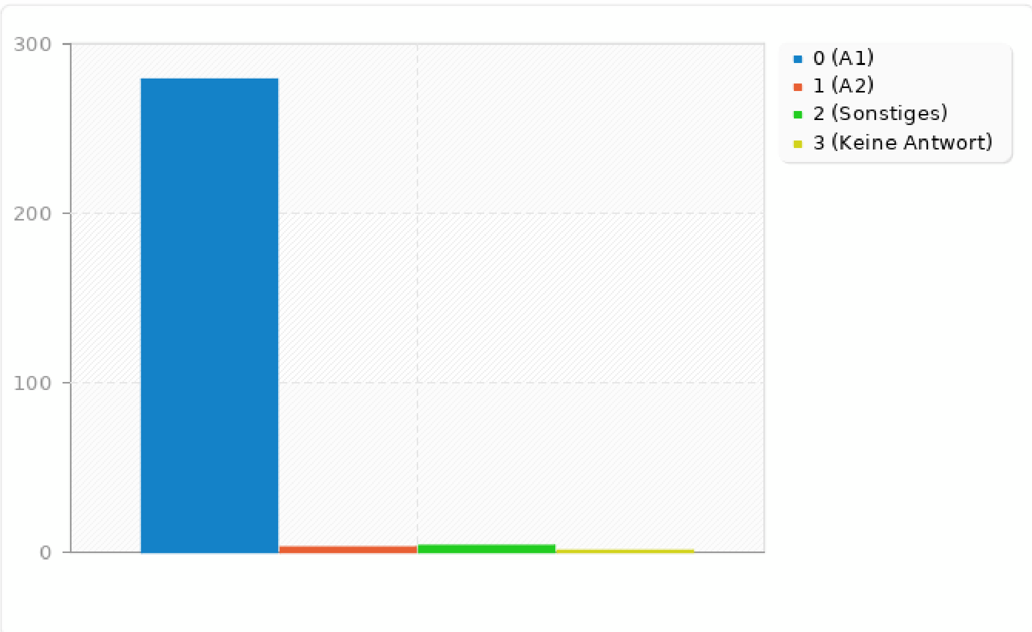
Welches mobile Betriebssystem nutzen Sie?

Antwort	Anzahl	Prozent
iOS (Apple) (A1)	279	97.21%
Android (A2)	3	1.05%
Sonstiges	4	1.39%
Keine Antwort	1	0.35%

ID	Antwort
88	beides
102	AmazonOS und iOS
112	Beides
218	Beides

Zusammenfassung für OS

Welches mobile Betriebssystem nutzen Sie?



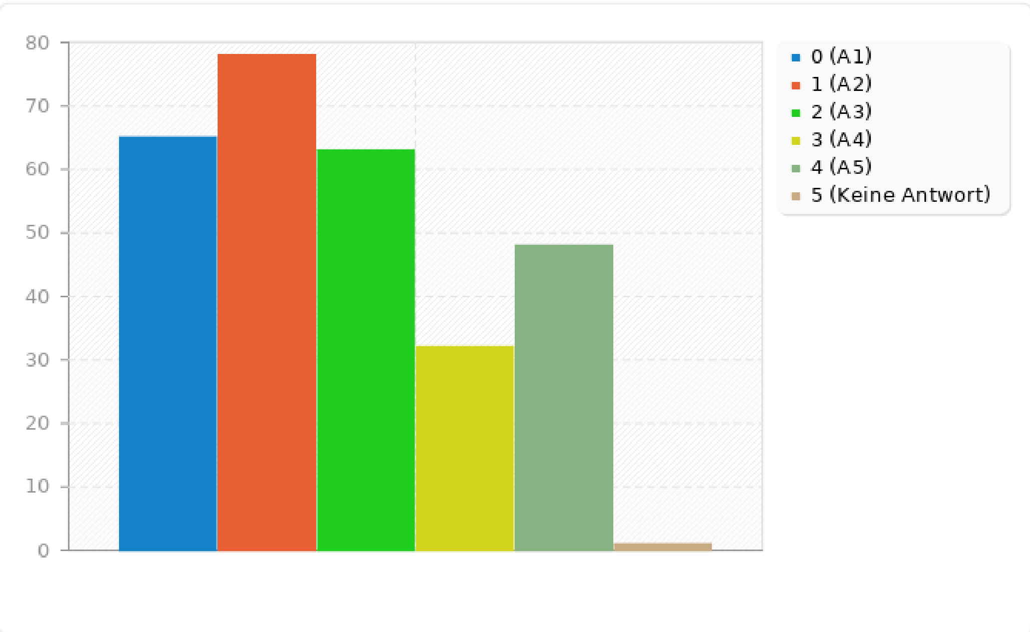
Zusammenfassung für Apps

Wie viele Apps befinden sich aktuell auf Ihrem Gerät?

Antwort	Anzahl	Prozent
1 bis 50 (A1)	65	22.65%
51 bis 99 (A2)	78	27.18%
100 bis 149 (A3)	63	21.95%
150 bis 199 (A4)	32	11.15%
200 oder mehr (A5)	48	16.72%
Keine Antwort	1	0.35%

Zusammenfassung für Apps

Wie viele Apps befinden sich aktuell auf Ihrem Gerät?



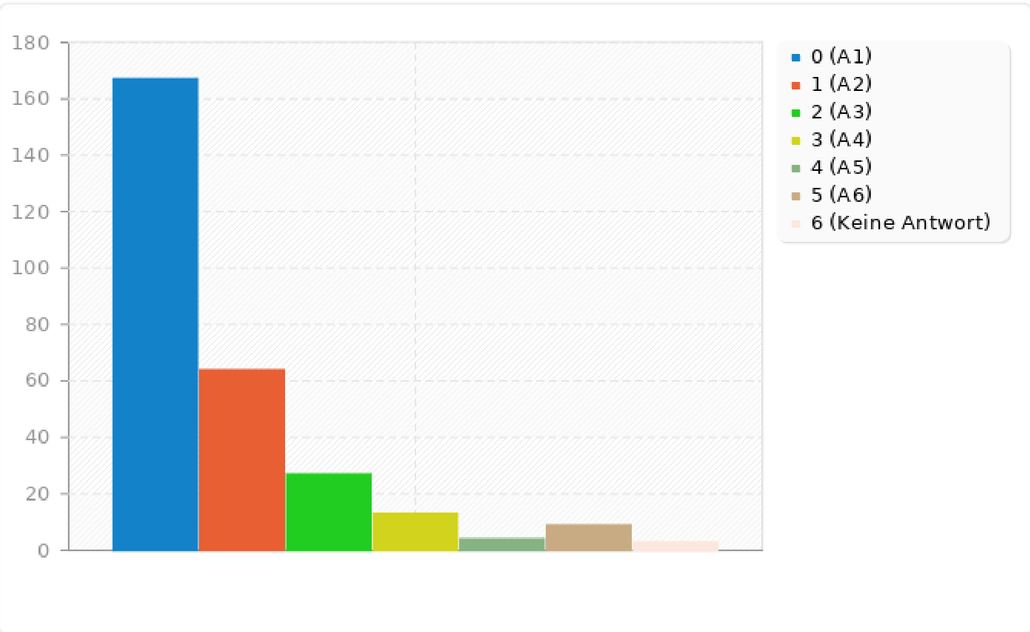
Zusammenfassung für Games

Wie viele der aktuell installierten Apps sind Spiele?

Antwort	Anzahl	Prozent
1 bis 9 (A1)	167	58.19%
10 bis 19 (A2)	64	22.30%
20 bis 29 (A3)	27	9.41%
30 bis 39 (A4)	13	4.53%
40 bis 49 (A5)	4	1.39%
50 und mehr (A6)	9	3.14%
Keine Antwort	3	1.05%

Zusammenfassung für Games

Wie viele der aktuell installierten Apps sind Spiele?



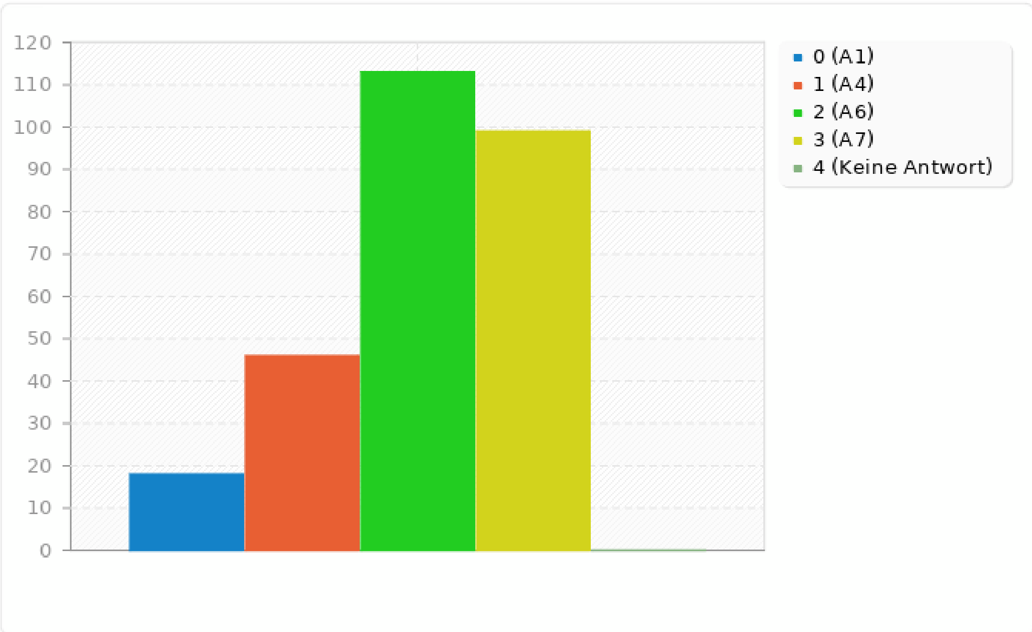
Zusammenfassung für AppNutzung

Wie viel Zeit verbringen Sie durchschnittlich am Tag mit der Nutzung von Apps auf Ihrem Smartphone oder Tablet?

Antwort	Anzahl	Prozent
weniger als eine Stunde (A1)	18	6.52%
mehr als drei Stunden (A4)	46	16.67%
ein bis zwei Stunden (A6)	113	40.94%
zwei bis drei Stunden (A7)	99	35.87%
Keine Antwort	0	0.00%

Zusammenfassung für AppNutzung

Wie viel Zeit verbringen Sie durchschnittlich am Tag mit der Nutzung von Apps auf Ihrem Smartphone oder Tablet?



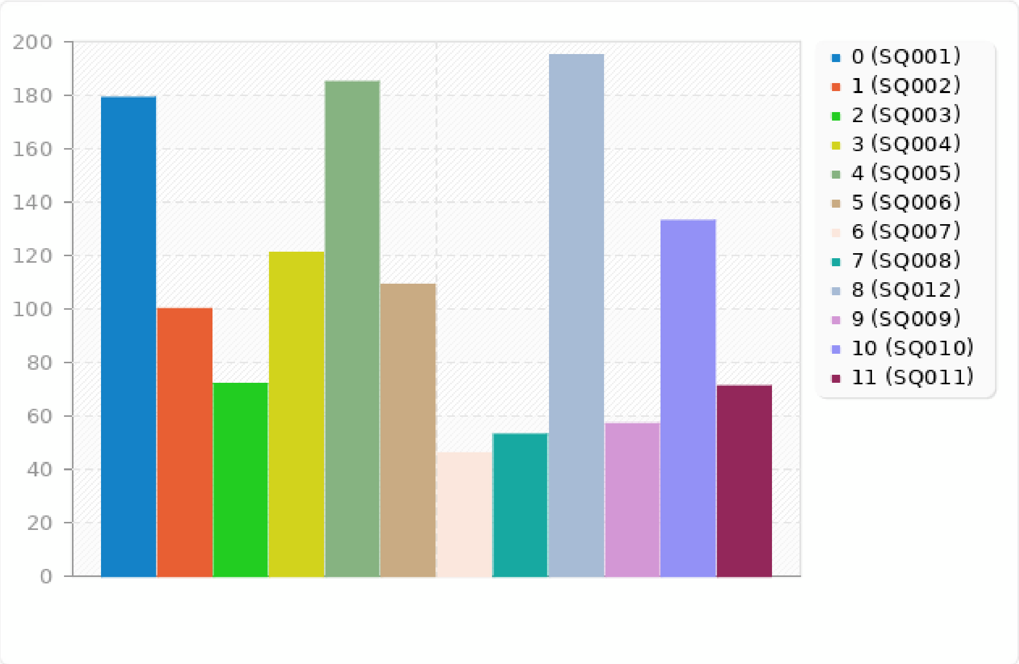
Zusammenfassung für AppType

Welche Art von mobilen Apps nutzen Sie am häufigsten?

Antwort	Anzahl	Prozent
Produktiv-Apps (SQ001)	179	64.86%
Spiele (SQ002)	100	36.23%
Bildung (SQ003)	72	26.09%
Foto und Video (SQ004)	121	43.84%
Messaging (SQ005)	185	67.03%
Soziale Netze (SQ006)	109	39.49%
Sport (SQ007)	46	16.67%
Unterhaltung (SQ008)	53	19.20%
News (SQ012)	195	70.65%
Zeitschriften und Bücher (SQ009)	57	20.65%
Musik (SQ010)	133	48.19%
Gesundheit und Fitness (SQ011)	71	25.72%

Zusammenfassung für AppType

Welche Art von mobilen Apps nutzen Sie am häufigsten?



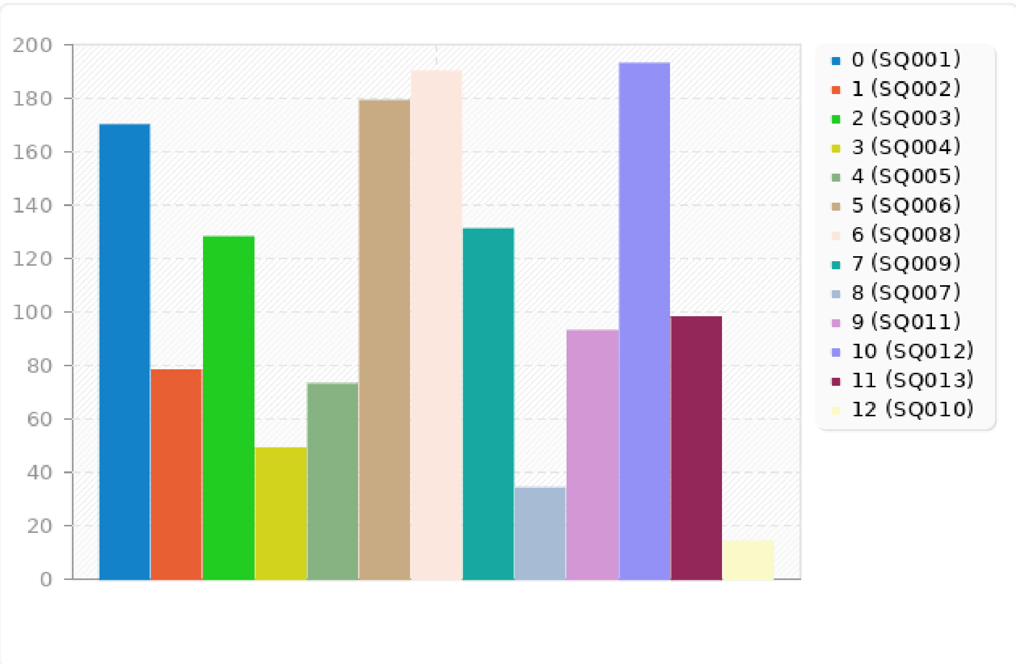
Zusammenfassung für NutzungsArt

Welche Art der App-Nutzung trifft am ehesten auf Sie zu?

Antwort	Anzahl	Prozent
Websurfen (SQ001)	170	61.59%
Videos schauen (SQ002)	78	28.26%
Musik hören (SQ003)	128	46.38%
E-Books lesen (SQ004)	49	17.75%
Spiele spielen (SQ005)	73	26.45%
E-Mail (SQ006)	179	64.86%
Messaging (WhatsApp und Co.) (SQ008)	190	68.84%
Kamera und Fotos (SQ009)	131	47.46%
Lernen (SQ007)	34	12.32%
Soziale Netzwerke (SQ011)	93	33.70%
Nachrichten und Informationen (SQ012)	193	69.93%
Shopping (SQ013)	98	35.51%
Sonstiges (SQ010)	14	5.07%

Zusammenfassung für NutzungsArt

Welche Art der App-Nutzung trifft am ehesten auf Sie zu?



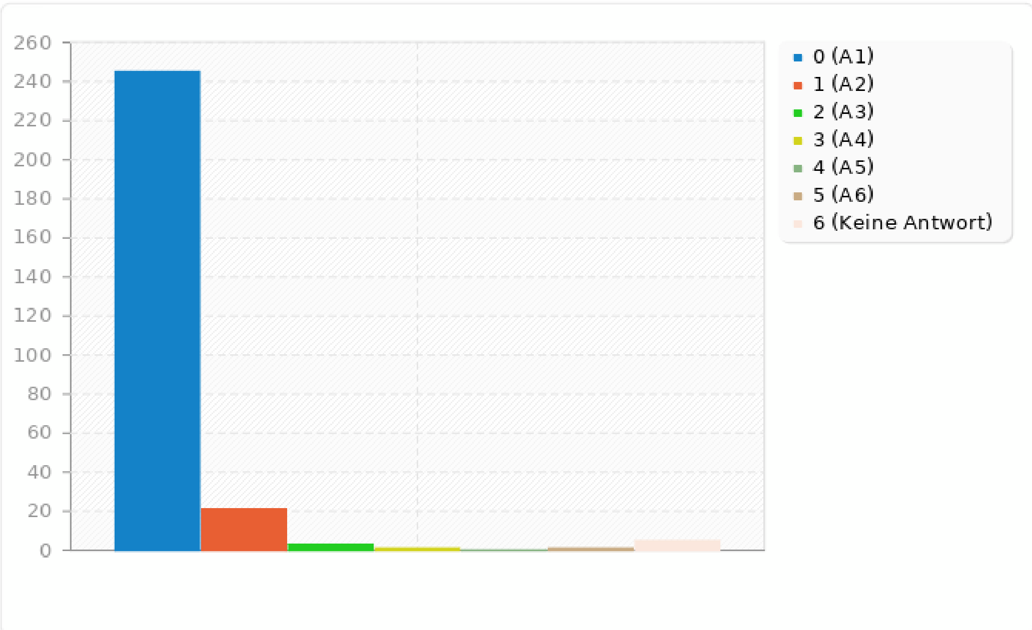
Zusammenfassung für Lernspiele

Wie viele der installierten Spiele würden Sie als Spiele bezeichnen, die sowohl vorder-, als auch hintergründig Wissen oder Fertigkeiten vermitteln (sogenannte Lernspiele)?

Antwort	Anzahl	Prozent
1 bis 9 (A1)	245	88.77%
10 bis 19 (A2)	21	7.61%
20 bis 29 (A3)	3	1.09%
30 bis 39 (A4)	1	0.36%
40 bis 49 (A5)	0	0.00%
50 oder mehr (A6)	1	0.36%
Keine Antwort	5	1.81%

Zusammenfassung für Lernspiele

Wie viele der installierten Spiele würden Sie als Spiele bezeichnen, die sowohl vorder-, als auch hintergründig Wissen oder Fertigkeiten vermitteln (sogenannte Lernspiele)?



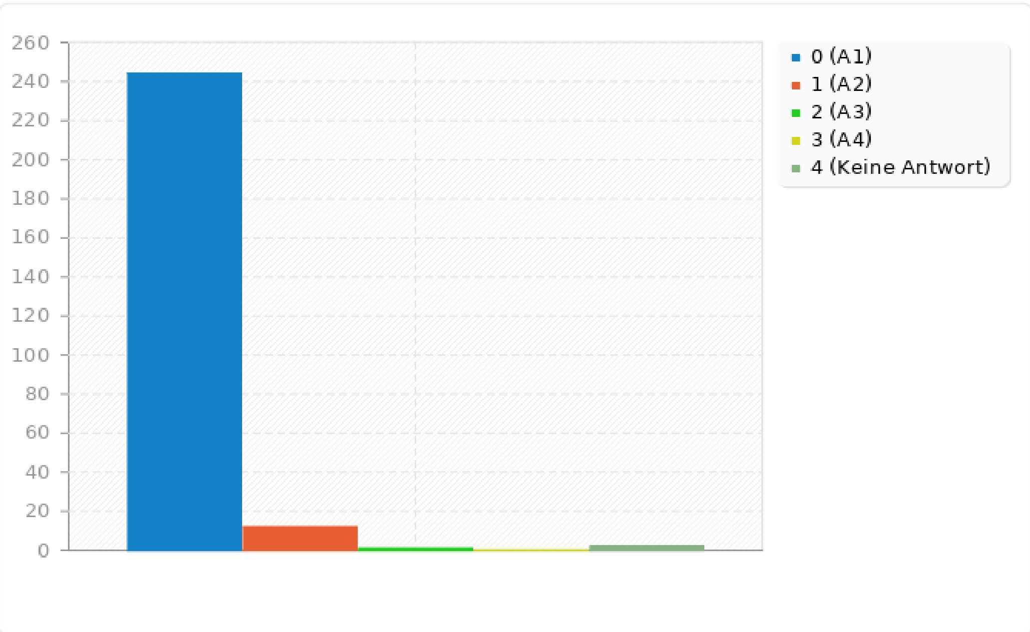
Zusammenfassung für AnzahlLernspiele

Wie viel Zeit verbringen Sie durchschnittlich am Tag mit Spielen, die Sie zuvor als Lernspiele klassifiziert haben?

Antwort	Anzahl	Prozent
weniger als eine Stunde (A1)	244	94.21%
ein bis zwei Stunden (A2)	12	4.63%
zwei bis drei Stunden (A3)	1	0.39%
mehr als drei Stunden (A4)	0	0.00%
Keine Antwort	2	0.77%

Zusammenfassung für AnzahlLernspiele

Wie viel Zeit verbringen Sie durchschnittlich am Tag mit Spielen, die Sie zuvor als Lernspiele klassifiziert haben?



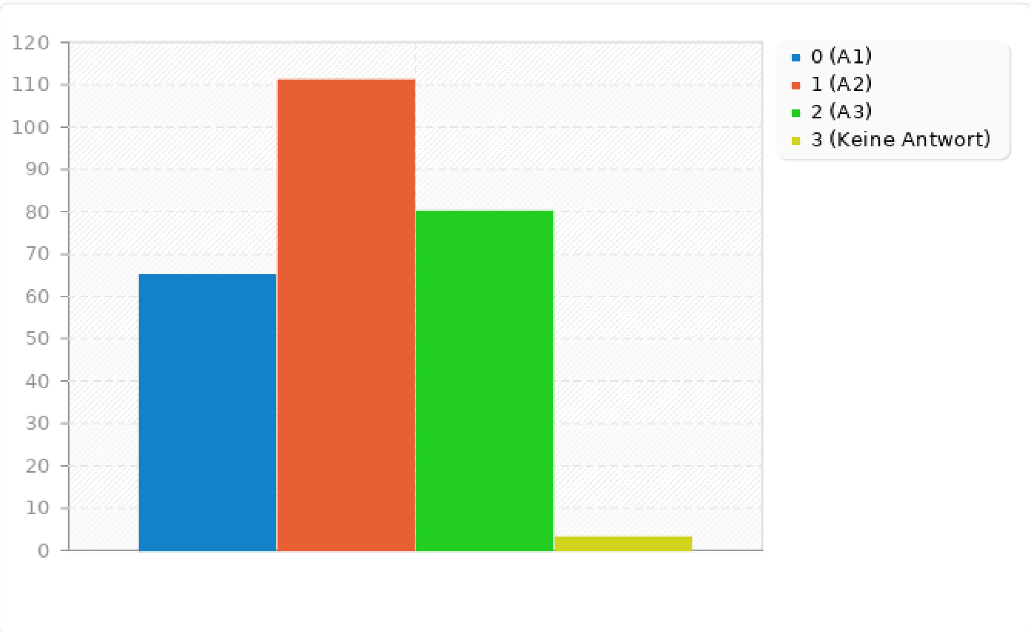
Zusammenfassung für LernspieleBewusst

Nutzen Sie Lernspiele bewusst zum Erlangen von bestimmtem Wissen oder Fertigkeiten?

Antwort	Anzahl	Prozent
ja (A1)	65	25.10%
nein (A2)	111	42.86%
teilweise (A3)	80	30.89%
Keine Antwort	3	1.16%

Zusammenfassung für LernspieleBewusst

Nutzen Sie Lernspiele bewusst zum Erlangen von bestimmtem Wissen oder Fertigkeiten?



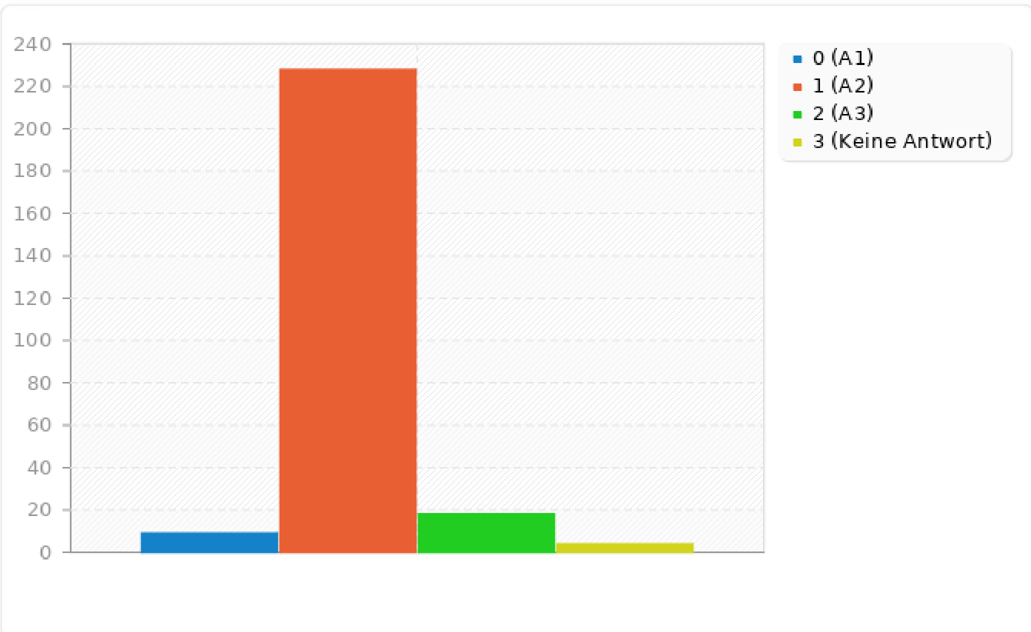
Zusammenfassung für LernspielePlan

Folgen Sie bei der Nutzung von Lernspielen einem bestimmten Zeitplan?

Antwort	Anzahl	Prozent
ja (A1)	9	3.47%
nein (A2)	228	88.03%
teilweise (A3)	18	6.95%
Keine Antwort	4	1.54%

Zusammenfassung für LernspielePlan

Folgen Sie bei der Nutzung von Lernspielen einem bestimmten Zeitplan?



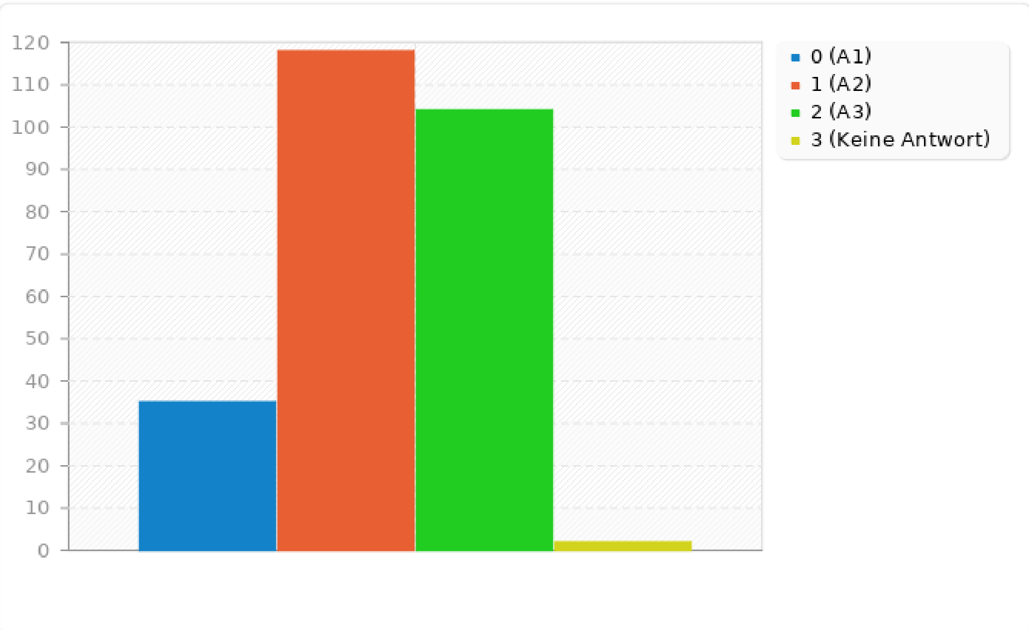
Zusammenfassung für Notifications

Würden Sie der Zeitplanung eines Lernspiels folgen, wenn dieses Sie per Benachrichtigung hieran erinnert?

Antwort	Anzahl	Prozent
ja (A1)	35	13.51%
nein (A2)	118	45.56%
vielleicht (A3)	104	40.15%
Keine Antwort	2	0.77%

Zusammenfassung für Notifications

Würden Sie der Zeitplanung eines Lernspiels folgen, wenn dieses Sie per Benachrichtigung hieran erinnert?



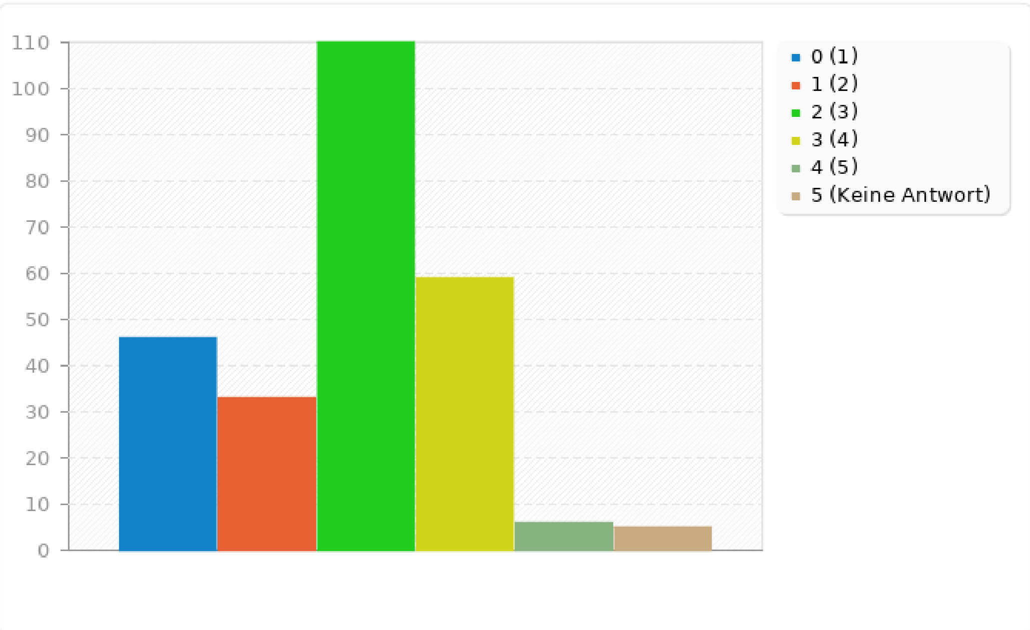
Zusammenfassung für Lernerfolg

Wie schätzen Sie den Lernerfolg bei der Nutzung der von Ihnen verwendeten Lernspiele ein?

Antwort	Anzahl	Prozent	Summe
1 (1)	46	14.89%	25.57%
2 (2)	33	10.68%	
3 (3)	110	35.60%	35.60%
4 (4)	59	19.09%	
5 (5)	6	1.94%	21.04%
Keine Antwort	5	1.59%	0.00%
Arithmetisches Mittel	2.79		
Standard Abweichung	1.07		
Summe (Antworten)	254	100.00%	100.00%
Anzahl Fälle		0%	

Zusammenfassung für Lernerfolg

Wie schätzen Sie den Lernerfolg bei der Nutzung der von Ihnen verwendeten Lernspiele ein?



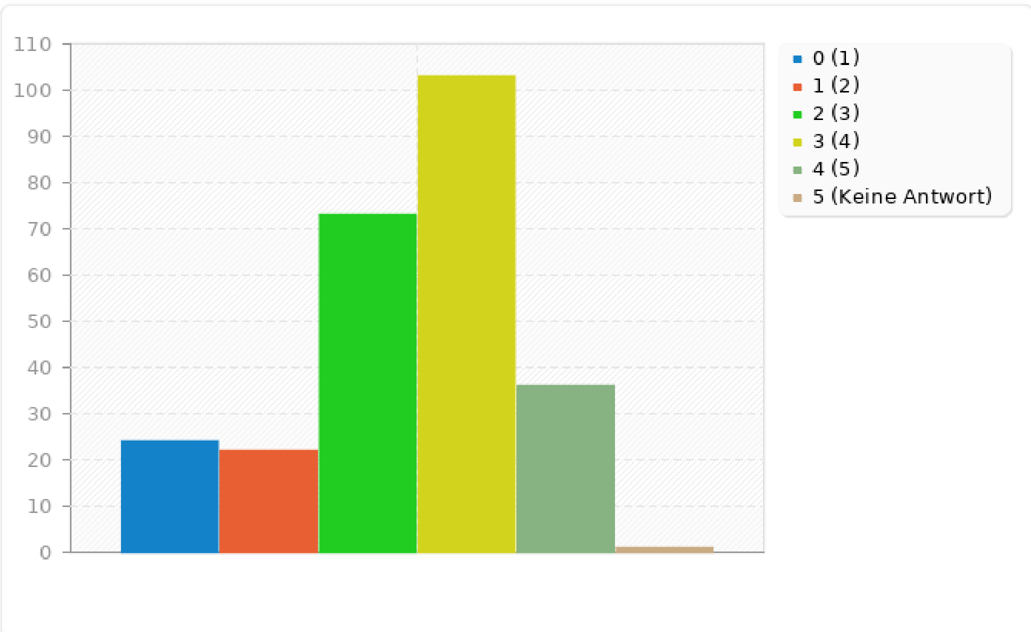
Zusammenfassung für Wissenschaft

Wie hoch ist die Wahrscheinlichkeit, dass Sie Lernspiele eher benutzen würden, wenn sie durch einen wissenschaftlichen Lernansatz gestützt wären?

Antwort	Anzahl	Prozent	Summe
1 (1)	24	7.67%	14.70%
2 (2)	22	7.03%	
3 (3)	73	23.32%	23.32%
4 (4)	103	32.91%	
5 (5)	36	11.50%	44.41%
Keine Antwort	1	0.32%	0.00%
Arithmetisches Mittel	3.41		
Standard Abweichung	1.12		
Summe (Antworten)	258	100.00%	100.00%
Anzahl Fälle		0%	

Zusammenfassung für Wissenschaft

Wie hoch ist die Wahrscheinlichkeit, dass Sie Lernspiele eher benutzen würden, wenn sie durch einen wissenschaftlichen Lernansatz gestützt wären?



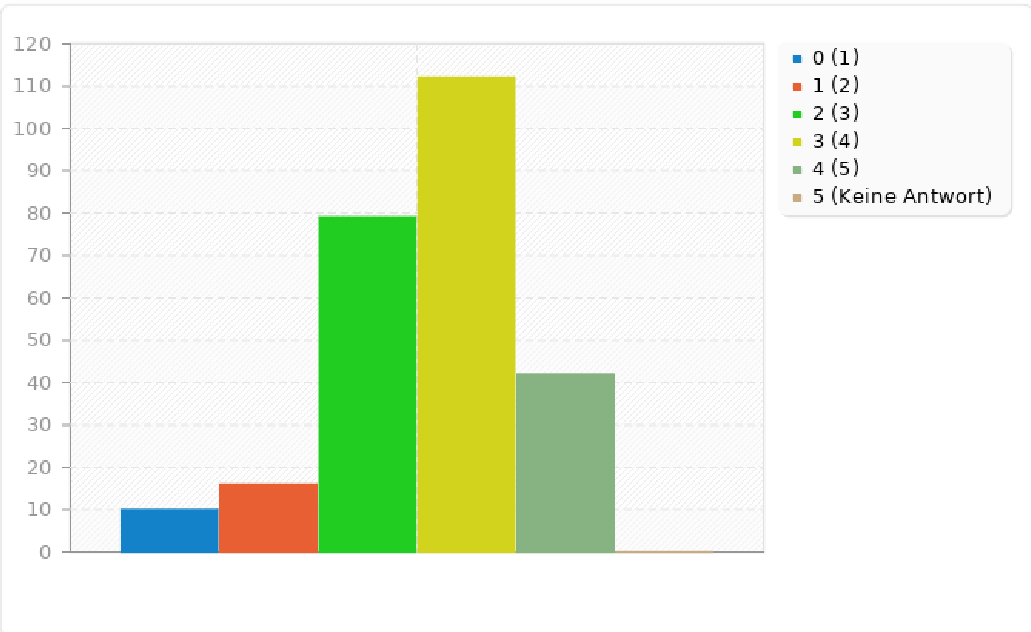
Zusammenfassung für BesseresErgebnis

Wie hoch beurteilen Sie die Wahrscheinlichkeit, dass Lernspiele mit einem wissenschaftlichen Lernansatz zu einem besseren Lernergebnis führen?

Antwort	Anzahl	Prozent	Summe
1 (1)	10	3.18%	8.28%
2 (2)	16	5.10%	
3 (3)	79	25.16%	25.16%
4 (4)	112	35.67%	
5 (5)	42	13.38%	49.04%
Keine Antwort	0	0.00%	0.00%
Arithmetisches Mittel	3.62		
Standard Abweichung	0.96		
Summe (Antworten)	259	100.00%	100.00%
Anzahl Fälle		0%	

Zusammenfassung für BesseresErgebnis

Wie hoch beurteilen Sie die Wahrscheinlichkeit, dass Lernspiele mit einem wissenschaftlichen Lernansatz zu einem besseren Lernergebnis führen?



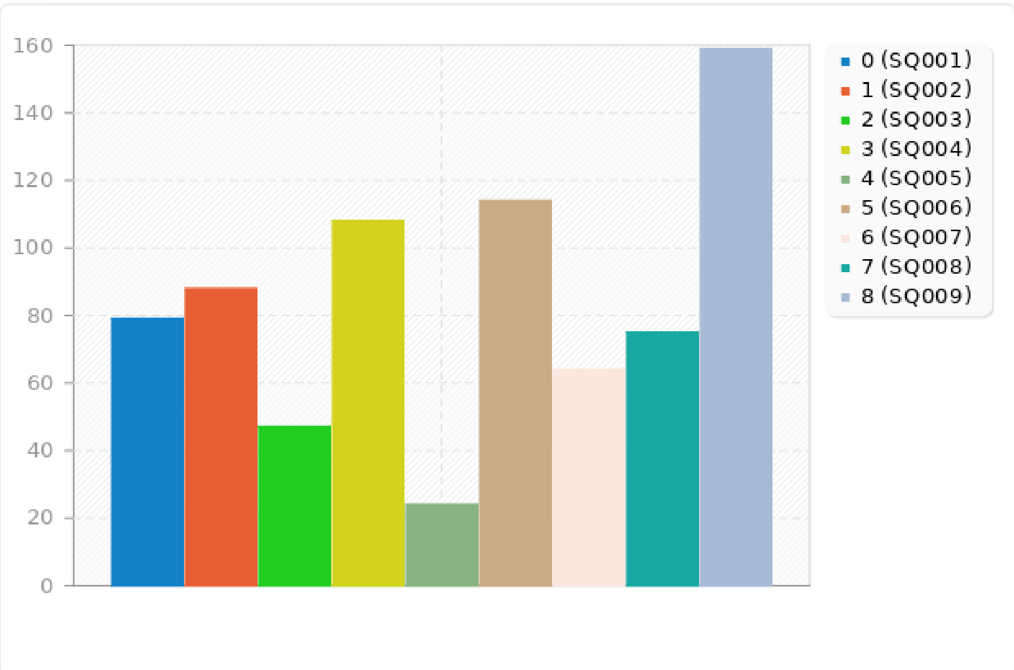
Zusammenfassung für Motivation

Welche Elemente würde Ihre Motivation, ein Lernspiel zu nutzen steigern?

Antwort	Anzahl	Prozent
Augmented Reality (SQ001)	79	30.50%
Ortsbezogene Funktionen (SQ002)	88	33.98%
Zeitbezogene Funktionen (SQ003)	47	18.15%
Multimediale Inhalte (SQ004)	108	41.70%
3D-Elemente (SQ005)	24	9.27%
herausragende grafische Gestaltung (SQ006)	114	44.02%
Ranglisten (SQ007)	64	24.71%
Benachrichtigungen / Erinnerungen (SQ008)	75	28.96%
Wissenschaftlicher Lernmethode (SQ009)	159	61.39%

Zusammenfassung für Motivation

Welche Elemente würde Ihre Motivation, ein Lernspiel zu nutzen steigern?



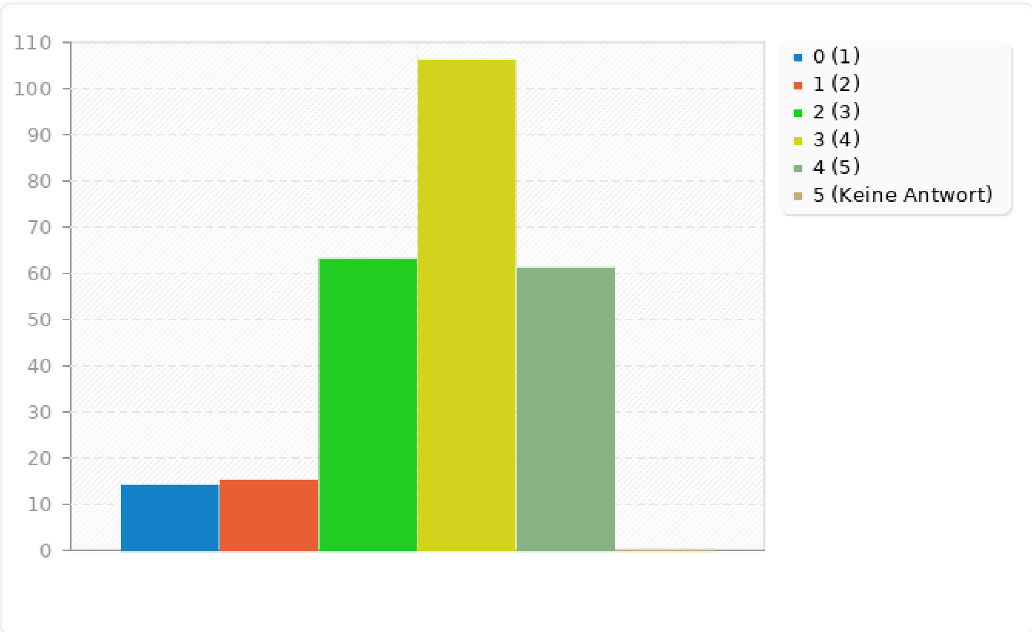
Zusammenfassung für Feedback

Wie wichtig ist Ihnen ein unmittelbares Feedback zu Ihrem aktuellen Lernstand bzw. Lernerfolg?

Antwort	Anzahl	Prozent	Summe
1 (1)	14	4.46%	9.24%
2 (2)	15	4.78%	
3 (3)	63	20.06%	20.06%
4 (4)	106	33.76%	
5 (5)	61	19.43%	53.18%
Keine Antwort	0	0.00%	0.00%
Arithmetisches Mittel	3.71		
Standard Abweichung	1.06		
Summe (Antworten)	259	100.00%	100.00%
Anzahl Fälle		0%	

Zusammenfassung für Feedback

Wie wichtig ist Ihnen ein unmittelbares Feedback zu Ihrem aktuellen Lernstand bzw. Lernerfolg?



A.4 Use of learning games - HSW students

The following pages contain the full results from the questionnaire among HSW students about their use of learning games (Chapter 3.3.4 and paper "Mobile Game-Based Learning in the App-Age – Where we are and where we want to be").

Nutzung mobiler Lernspiele durch HSW-Studierende

Ergebnisse

Umfrage 778312

Anzahl der Datensätze in dieser Abfrage:	126
Gesamtzahl der Datensätze dieser Umfrage:	126
Anteil in Prozent:	100.00%

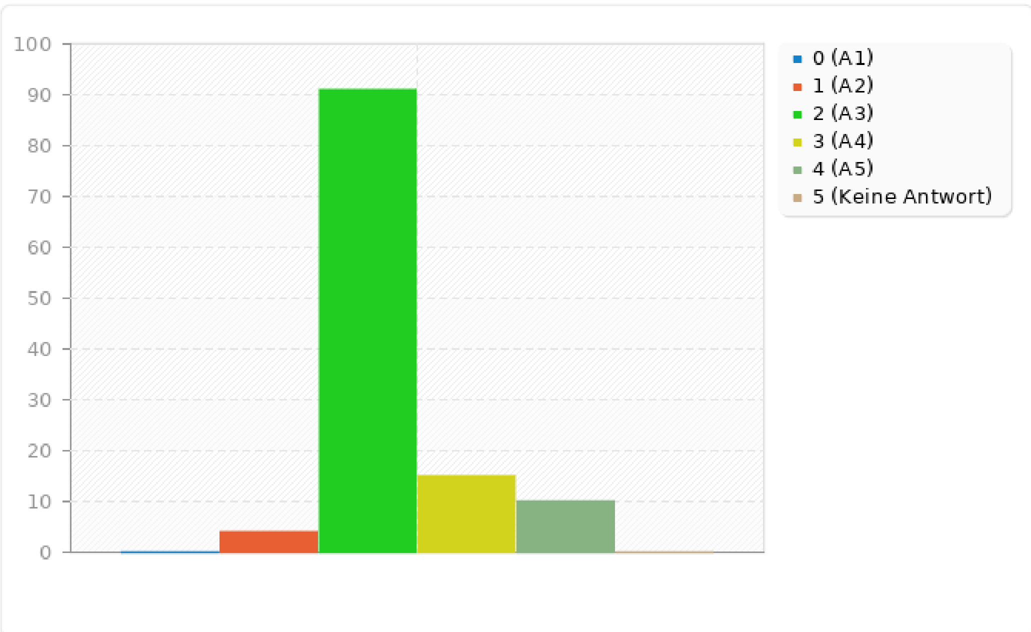
Zusammenfassung für Alter

Alter

Antwort	Anzahl	Prozent
8 bis 14 Jahre (A1)	0	0.00%
15 bis 18 Jahre (A2)	4	3.33%
19 bis 24 Jahre (A3)	91	75.83%
25 bis 30 Jahre (A4)	15	12.50%
älter als 30 Jahre (A5)	10	8.33%
Keine Antwort	0	0.00%

Zusammenfassung für Alter

Alter



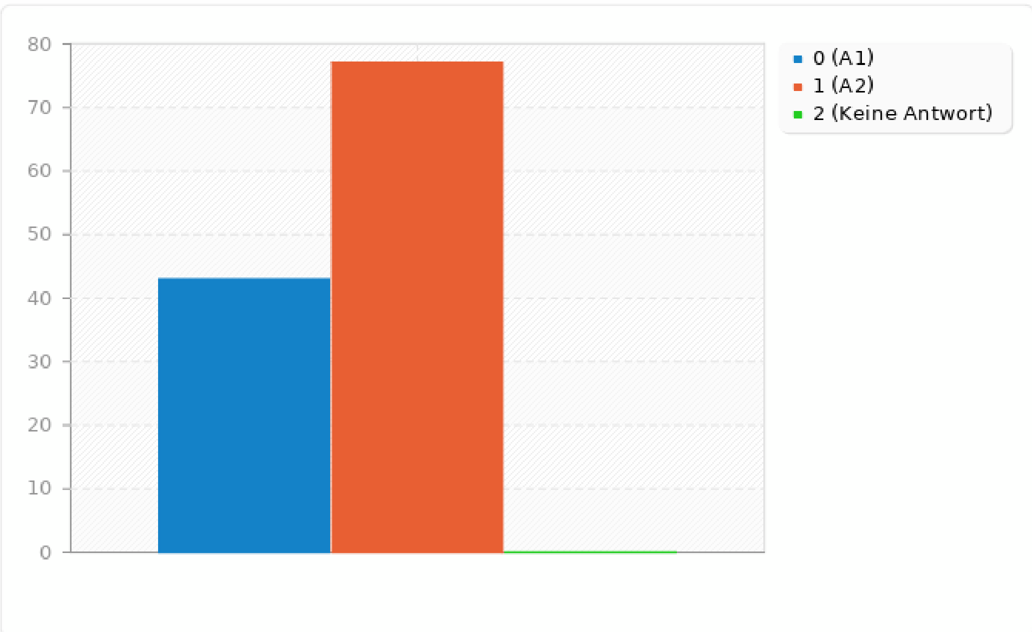
Zusammenfassung für Geschlecht

Geschlecht

Antwort	Anzahl	Prozent
weiblich (A1)	43	35.83%
männlich (A2)	77	64.17%
Keine Antwort	0	0.00%

Zusammenfassung für Geschlecht

Geschlecht



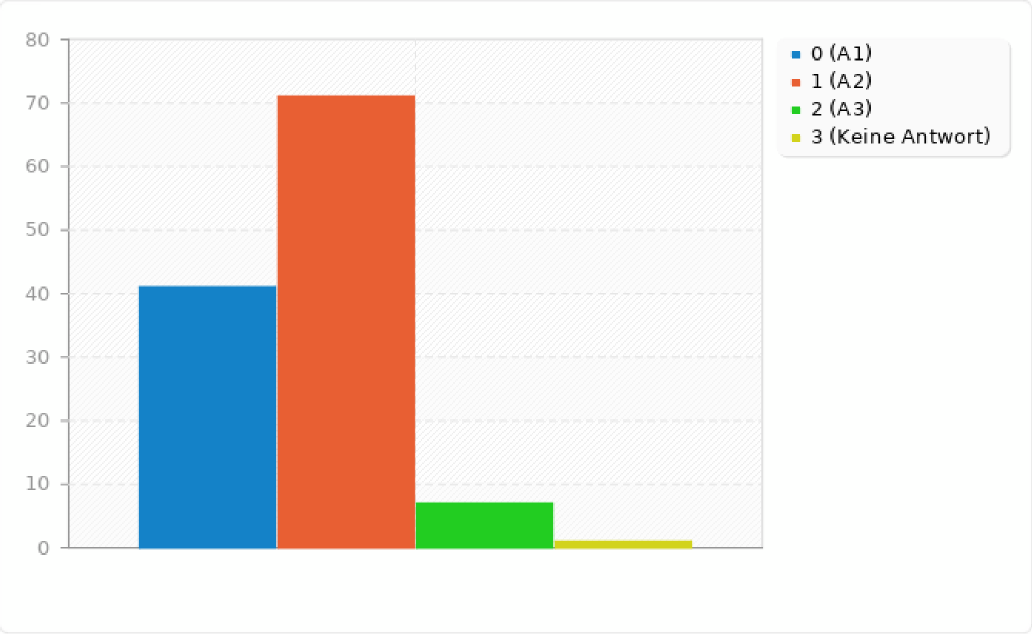
Zusammenfassung für Studiengang

In welchem Studiengang studieren Sie an der HSW?

Antwort	Anzahl	Prozent
Betriebswirtschaftslehre (A1)	41	34.17%
Wirtschaftsinformatik (A2)	71	59.17%
Wirtschaftsingenieur (A3)	7	5.83%
Keine Antwort	1	0.83%

Zusammenfassung für Studiengang

In welchem Studiengang studieren Sie an der HSW?



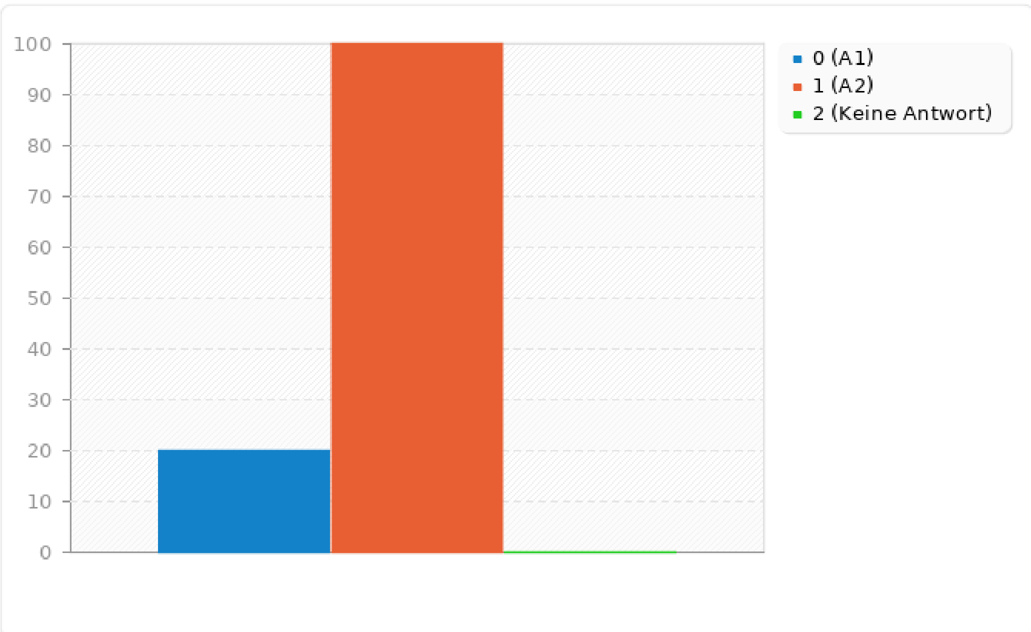
Zusammenfassung für Studienform

Studieren Sie dual oder berufsbegleitend an der HSW?

Antwort	Anzahl	Prozent
berufsbegleitend (A1)	20	16.67%
dual (A2)	100	83.33%
Keine Antwort	0	0.00%

Zusammenfassung für Studienform

Studieren Sie dual oder berufsbegleitend an der HSW?



Zusammenfassung für Semester

In welchem Semester studieren Sie an der HSW?

Antwort	Anzahl	Prozent
Antwort	119	99.17%
Keine Antwort	1	0.83%

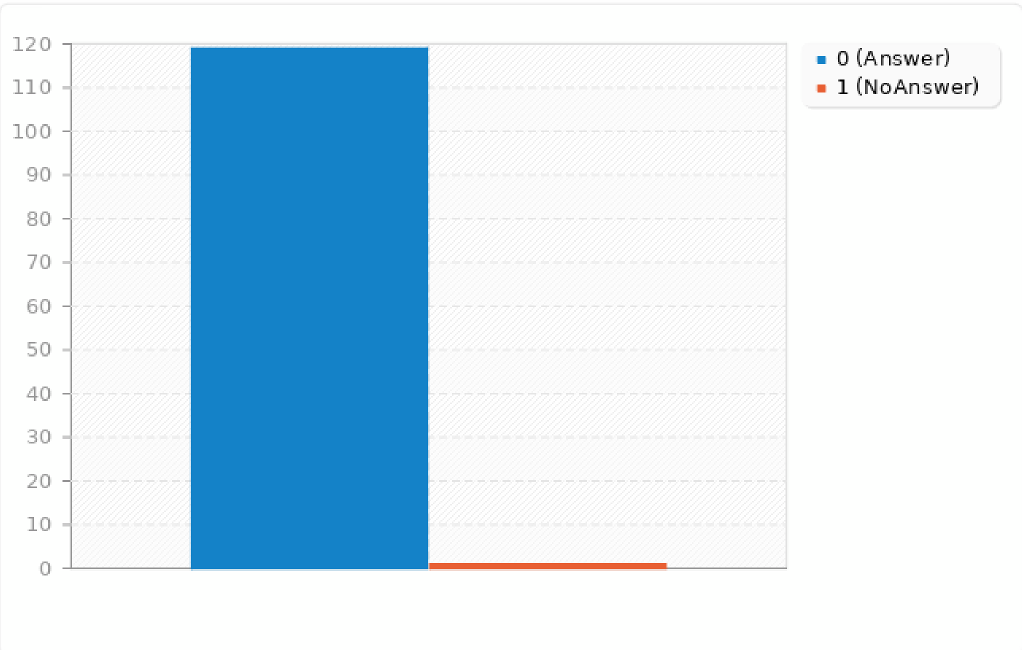
ID	Antwort
1	3
2	6
3	6
4	6. Semester
5	2
6	6
7	2
8	2.
10	4.
11	6.
12	2. Semester
13	6
14	6
15	2
16	2
17	2
18	5
19	2
20	6
21	2.
22	6
23	4.
24	6
25	2
26	2
27	6
28	4
29	1. - Masterstudium
30	6
31	4
33	4
34	4
35	5
36	2
37	2
38	6.
39	2
40	4
41	2
42	2
43	2
44	2. Semester
45	4
46	5
47	2
48	2
49	6
50	4
51	2
52	6. Semester
53	2
55	6.

56	4
57	4
58	4
59	BbBWL 02/15
60	2
61	2
62	4
63	2
64	6
65	6
66	4
67	4
68	2
69	6
70	4
71	2
72	2
73	2
74	4
75	4
76	6
77	4.
78	7
79	6
80	6
81	6
82	6
83	6
84	6
85	6.
86	6
87	2
88	2
89	4
91	4
92	6
93	4
94	2
95	4
96	2
98	6.
99	2
100	2
101	6
102	4
103	4
105	4
106	4
107	4
108	4
109	4.
110	4
111	6
112	4
113	6
114	6
115	4
116	2
117	2.
118	6
119	2
122	4
123	2
124	6
125	6. Semester

126	2
127	4

Zusammenfassung für Semester

In welchem Semester studieren Sie an der HSW?



Zusammenfassung für OS

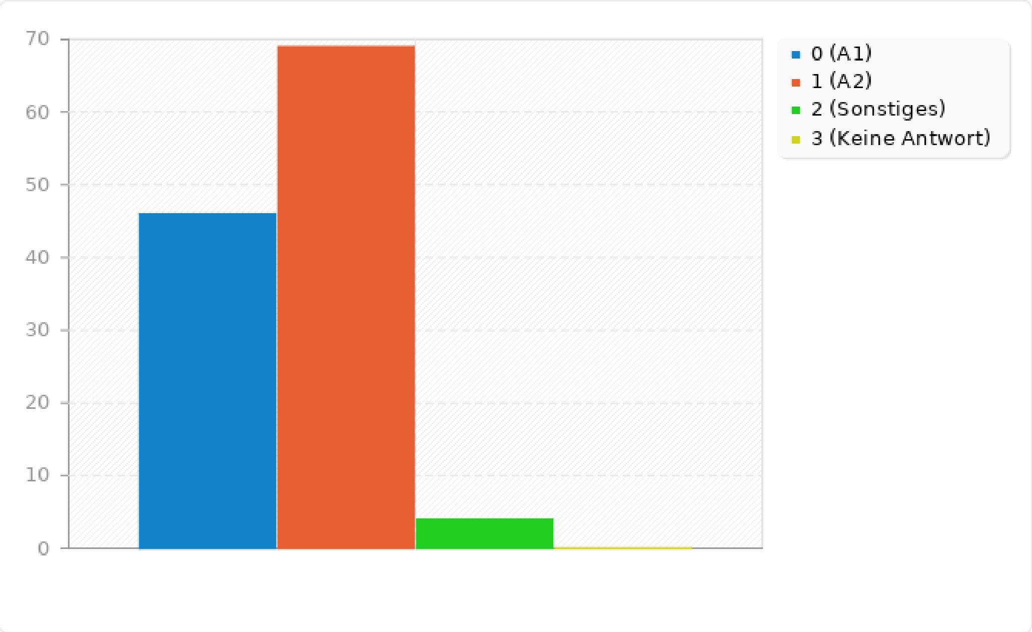
Welches mobile Betriebssystem nutzen Sie?

Antwort	Anzahl	Prozent
iOS (Apple) (A1)	46	38.66%
Android (A2)	69	57.98%
Sonstiges	4	3.36%
Keine Antwort	0	0.00%

ID	Antwort
60	Blackberry OS
67	Beide
75	Beide
114	Windows CE

Zusammenfassung für OS

Welches mobile Betriebssystem nutzen Sie?



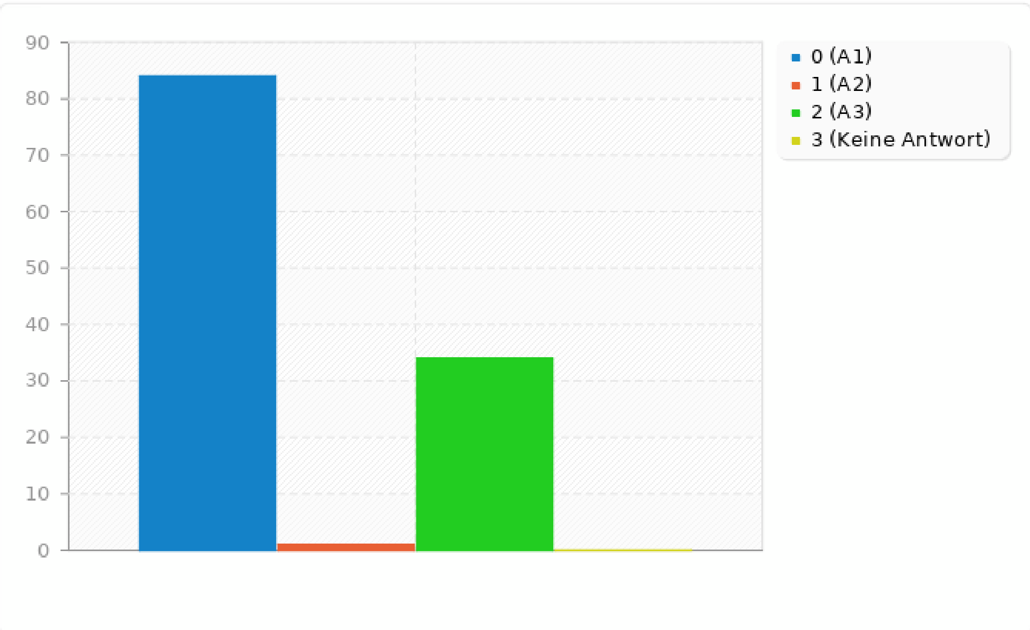
Zusammenfassung für Geraet

Welches mobile Gerät besitzen und benutzen Sie regelmäßig?

Antwort	Anzahl	Prozent
Smartphone (A1)	84	70.59%
Tablet (A2)	1	0.84%
beides (A3)	34	28.57%
Keine Antwort	0	0.00%

Zusammenfassung für Geraet

Welches mobile Gerät besitzen und benutzen Sie regelmäßig?



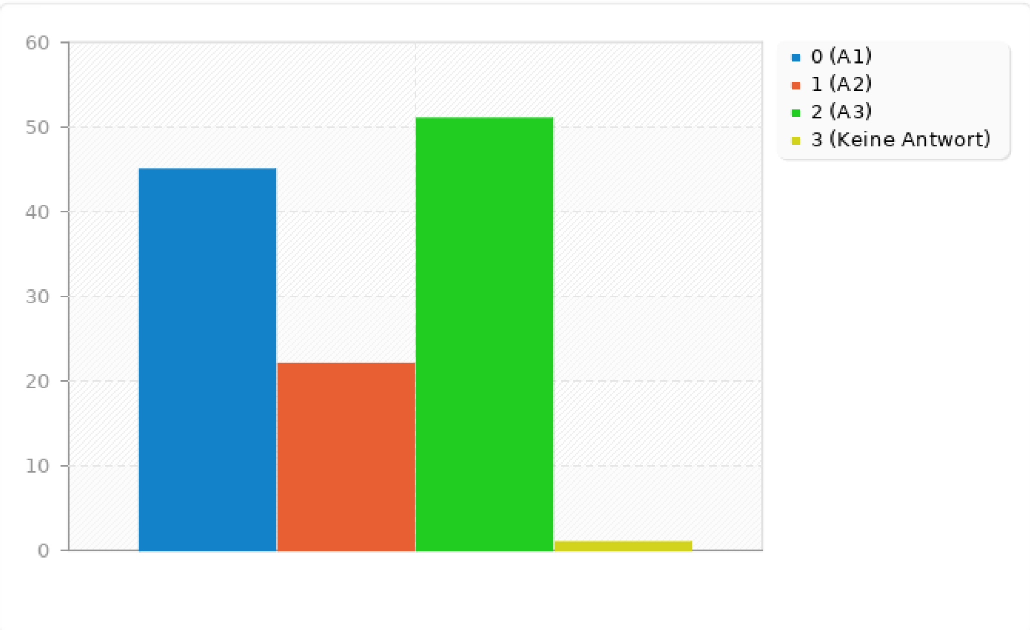
Zusammenfassung für Studium

Nutzen Sie ihr mobiles Gerät aktiv für Ihr Studium an der HSW?

Antwort	Anzahl	Prozent
ja (A1)	45	37.82%
nein (A2)	22	18.49%
nur manchmal (A3)	51	42.86%
Keine Antwort	1	0.84%

Zusammenfassung für Studium

Nutzen Sie ihr mobiles Gerät aktiv für Ihr Studium an der HSW?



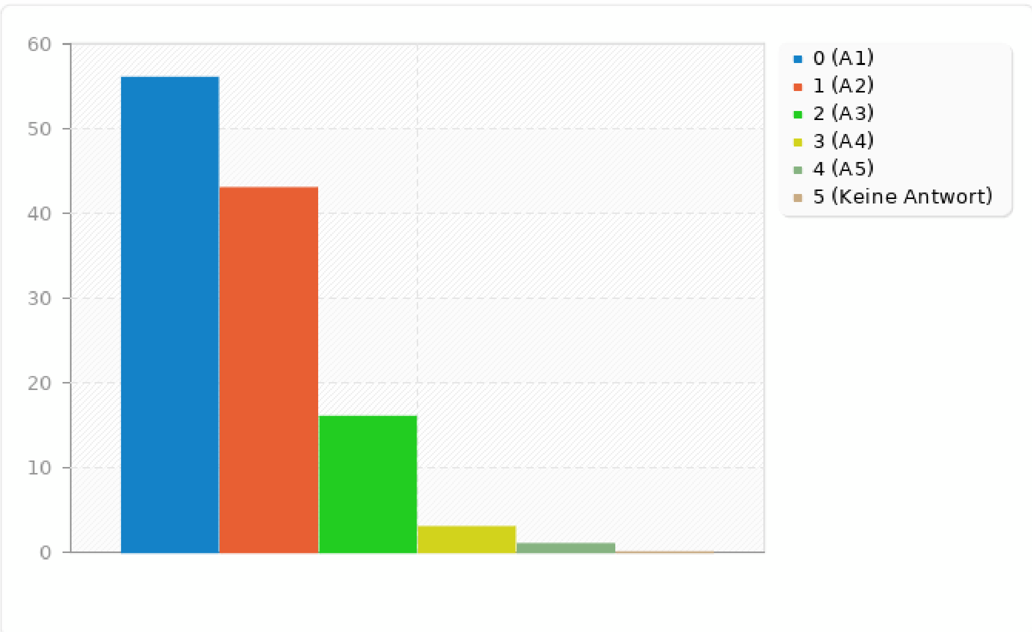
Zusammenfassung für Apps

Wie viele Apps befinden sich aktuell auf Ihrem Gerät?

Antwort	Anzahl	Prozent
1 bis 50 (A1)	56	47.06%
51 bis 99 (A2)	43	36.13%
100 bis 149 (A3)	16	13.45%
150 bis 199 (A4)	3	2.52%
200 oder mehr (A5)	1	0.84%
Keine Antwort	0	0.00%

Zusammenfassung für Apps

Wie viele Apps befinden sich aktuell auf Ihrem Gerät?



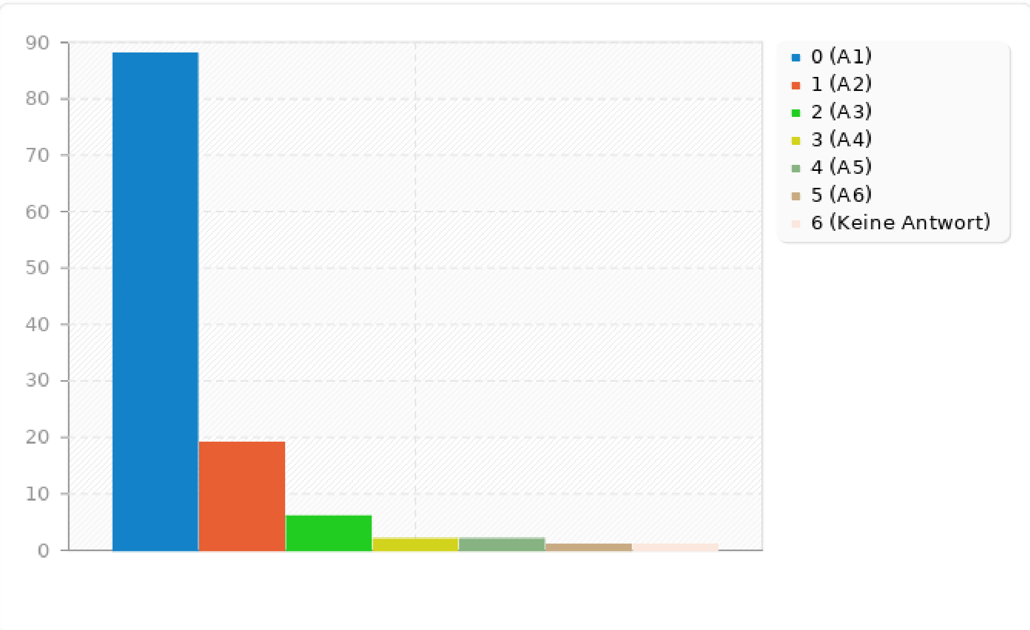
Zusammenfassung für Games

Wie viele der aktuell installierten Apps sind Spiele?

Antwort	Anzahl	Prozent
1 bis 9 (A1)	88	73.95%
10 bis 19 (A2)	19	15.97%
20 bis 29 (A3)	6	5.04%
30 bis 39 (A4)	2	1.68%
40 bis 49 (A5)	2	1.68%
50 und mehr (A6)	1	0.84%
Keine Antwort	1	0.84%

Zusammenfassung für Games

Wie viele der aktuell installierten Apps sind Spiele?



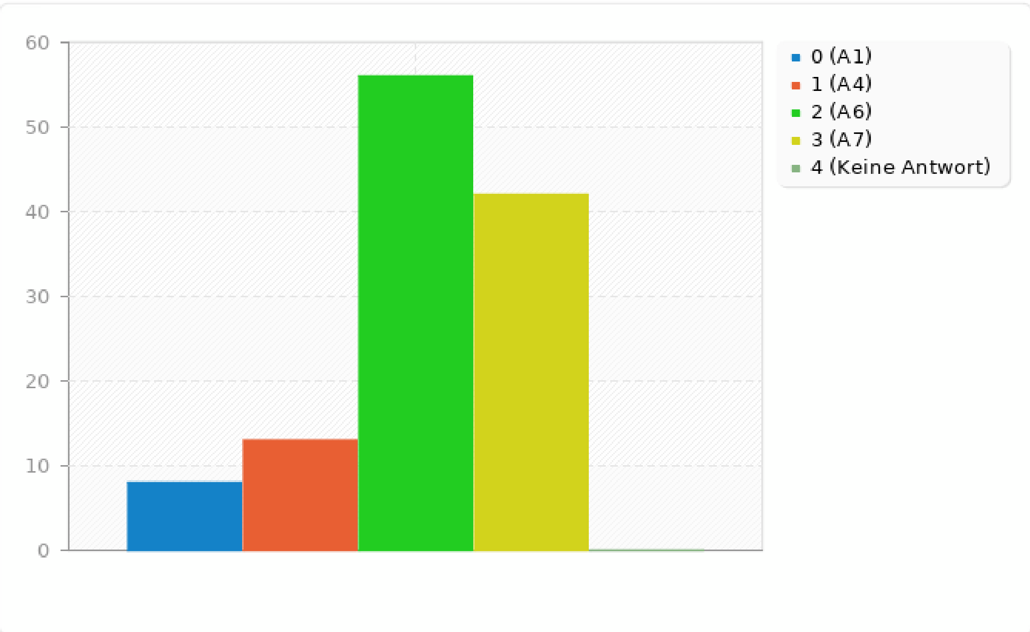
Zusammenfassung für AppNutzung

Wie viel Zeit verbringen Sie durchschnittlich am Tag mit der Nutzung von Apps auf Ihrem Smartphone oder Tablet?

Antwort	Anzahl	Prozent
weniger als eine Stunde (A1)	8	6.72%
mehr als drei Stunden (A4)	13	10.92%
ein bis zwei Stunden (A6)	56	47.06%
zwei bis drei Stunden (A7)	42	35.29%
Keine Antwort	0	0.00%

Zusammenfassung für AppNutzung

Wie viel Zeit verbringen Sie durchschnittlich am Tag mit der Nutzung von Apps auf Ihrem Smartphone oder Tablet?



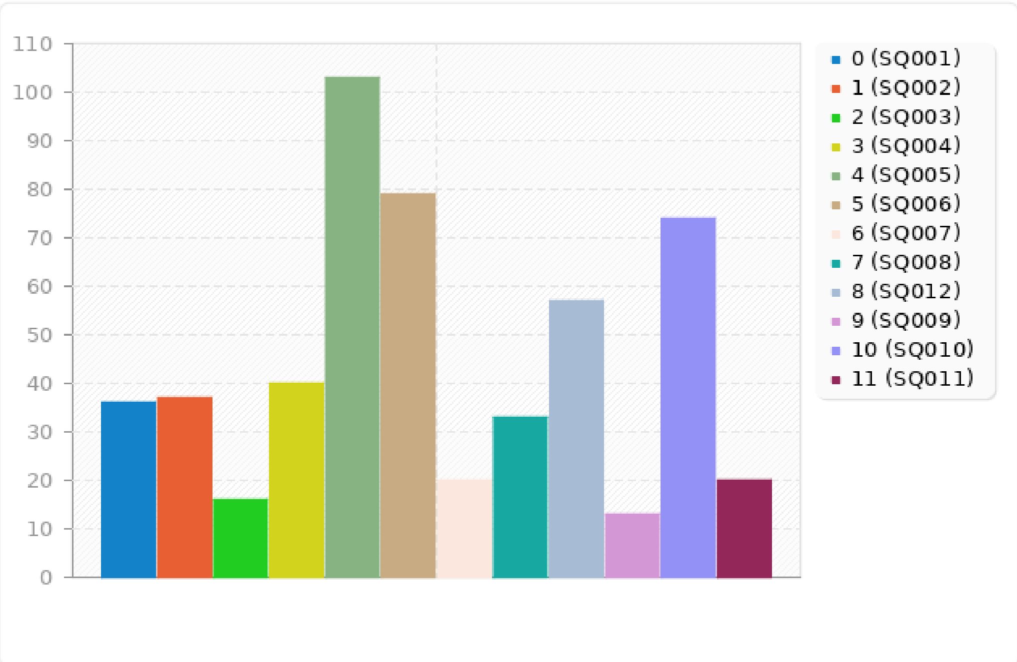
Zusammenfassung für AppType

Welche Art von mobilen Apps nutzen Sie am häufigsten?

Antwort	Anzahl	Prozent
Produktiv-Apps (SQ001)	36	30.25%
Spiele (SQ002)	37	31.09%
Bildung (SQ003)	16	13.45%
Foto und Video (SQ004)	40	33.61%
Messaging (SQ005)	103	86.55%
Soziale Netze (SQ006)	79	66.39%
Sport (SQ007)	20	16.81%
Unterhaltung (SQ008)	33	27.73%
News (SQ012)	57	47.90%
Zeitschriften und Bücher (SQ009)	13	10.92%
Musik (SQ010)	74	62.18%
Gesundheit und Fitness (SQ011)	20	16.81%

Zusammenfassung für AppType

Welche Art von mobilen Apps nutzen Sie am häufigsten?



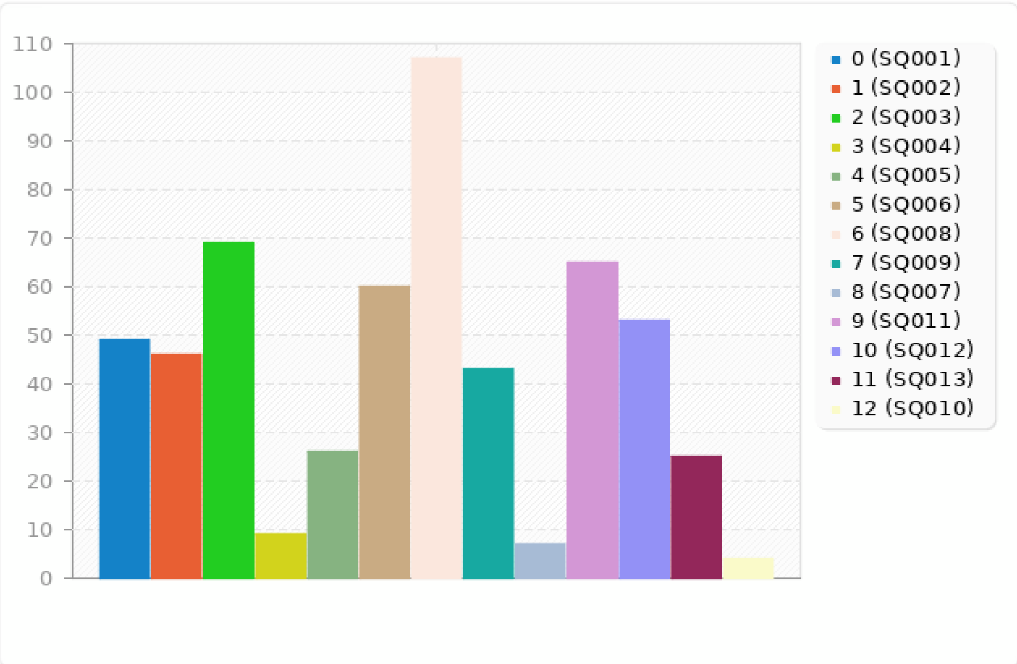
Zusammenfassung für NutzungsArt

Welche Art der App-Nutzung trifft am ehesten auf Sie zu?

Antwort	Anzahl	Prozent
Websurfen (SQ001)	49	41.18%
Videos schauen (SQ002)	46	38.66%
Musik hören (SQ003)	69	57.98%
E-Books lesen (SQ004)	9	7.56%
Spiele spielen (SQ005)	26	21.85%
E-Mail (SQ006)	60	50.42%
Messaging (WhatsApp und Co.) (SQ008)	107	89.92%
Kamera und Fotos (SQ009)	43	36.13%
Lernen (SQ007)	7	5.88%
Soziale Netzwerke (SQ011)	65	54.62%
Nachrichten und Informationen (SQ012)	53	44.54%
Shopping (SQ013)	25	21.01%
Sonstiges (SQ010)	4	3.36%

Zusammenfassung für NutzungsArt

Welche Art der App-Nutzung trifft am ehesten auf Sie zu?



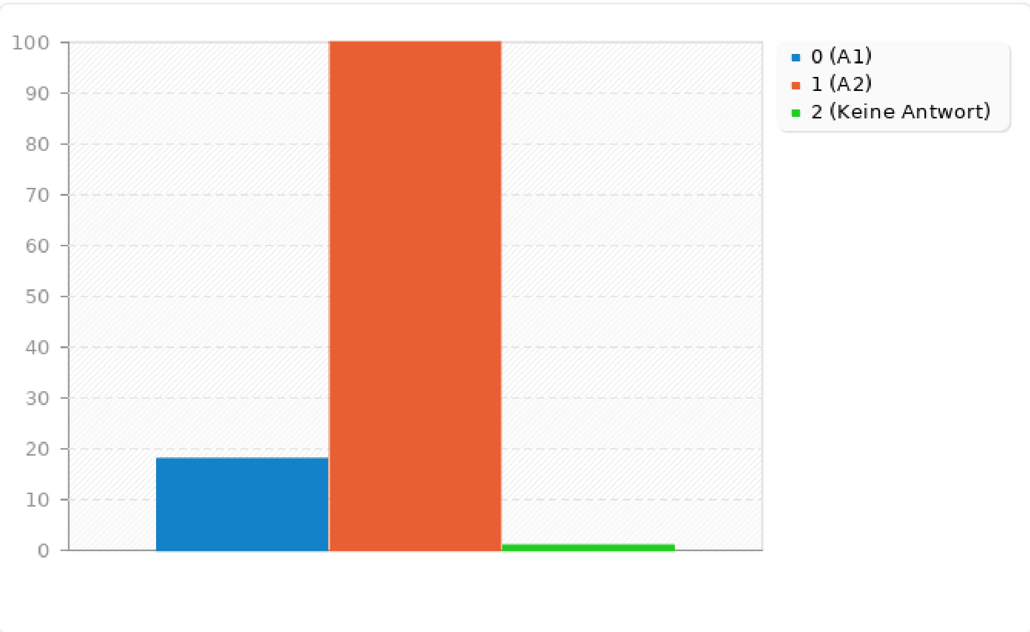
Zusammenfassung für Lernapps

Haben Sie spezielle Lern-Apps installiert, die Sie für Ihr Studium nutzen?

Antwort	Anzahl	Prozent
ja (A1)	18	15.13%
nein (A2)	100	84.03%
Keine Antwort	1	0.84%

Zusammenfassung für Lernapps

Haben Sie spezielle Lern-Apps installiert, die Sie für Ihr Studium nutzen?



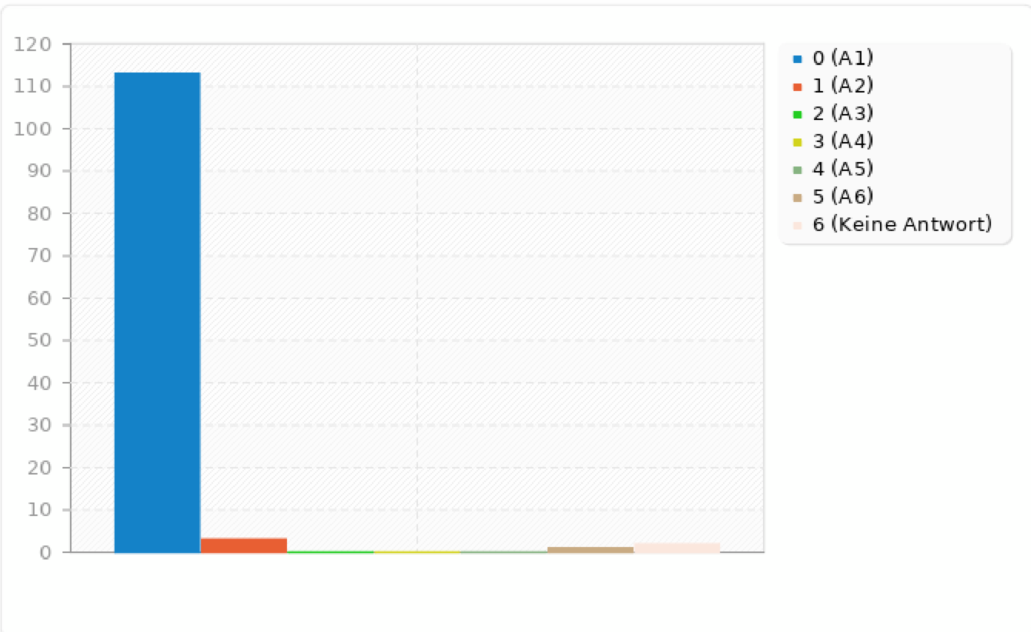
Zusammenfassung für Lernspiele

Wie viele der installierten Spiele würden Sie als Spiele bezeichnen, die sowohl vorder-, als auch hintergründig Wissen oder Fertigkeiten vermitteln (sogenannte Lernspiele)?

Antwort	Anzahl	Prozent
1 bis 9 (A1)	113	94.96%
10 bis 19 (A2)	3	2.52%
20 bis 29 (A3)	0	0.00%
30 bis 39 (A4)	0	0.00%
40 bis 49 (A5)	0	0.00%
50 oder mehr (A6)	1	0.84%
Keine Antwort	2	1.68%

Zusammenfassung für Lernspiele

Wie viele der installierten Spiele würden Sie als Spiele bezeichnen, die sowohl vorder-, als auch hintergründig Wissen oder Fertigkeiten vermitteln (sogenannte Lernspiele)?



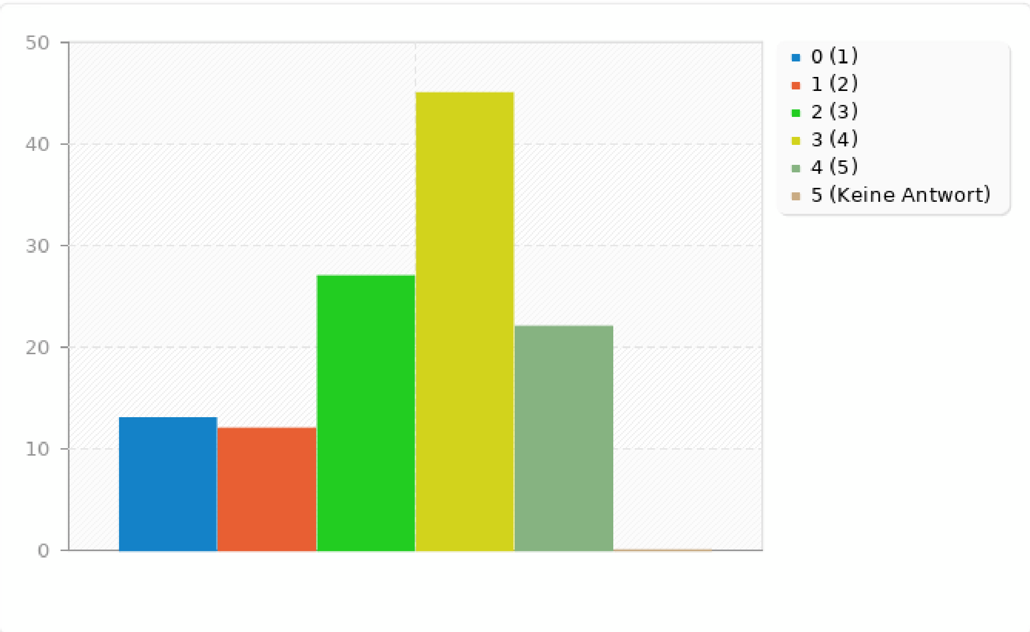
Zusammenfassung für Einbeziehung

Wie sehr würden Sie sich eine stärkere Einbeziehung von mobilen Lern-Apps in das Studium wünschen?

Antwort	Anzahl	Prozent	Summe
1 (1)	13	10.32%	19.84%
2 (2)	12	9.52%	
3 (3)	27	21.43%	21.43%
4 (4)	45	35.71%	
5 (5)	22	17.46%	53.17%
Keine Antwort	0	0.00%	0.00%
Arithmetisches Mittel	3.43		
Standard Abweichung	1.22		
Summe (Antworten)	119	100.00%	100.00%
Anzahl Fälle		0%	

Zusammenfassung für Einbeziehung

Wie sehr würden Sie sich eine stärkere Einbeziehung von mobilen Lern-Apps in das Studium wünschen?



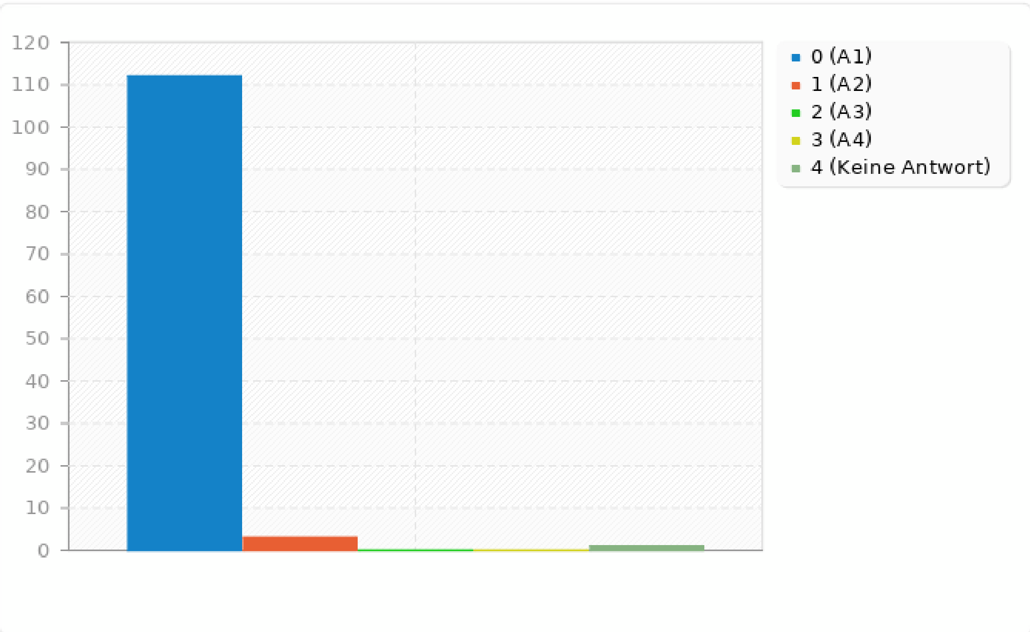
Zusammenfassung für AnzahlLernspiele

Wie viel Zeit verbringen Sie durchschnittlich am Tag mit Spielen, die Sie zuvor als Lernspiele klassifiziert haben?

Antwort	Anzahl	Prozent
weniger als eine Stunde (A1)	112	96.55%
ein bis zwei Stunden (A2)	3	2.59%
zwei bis drei Stunden (A3)	0	0.00%
mehr als drei Stunden (A4)	0	0.00%
Keine Antwort	1	0.86%

Zusammenfassung für AnzahlLernspiele

Wie viel Zeit verbringen Sie durchschnittlich am Tag mit Spielen, die Sie zuvor als Lernspiele klassifiziert haben?



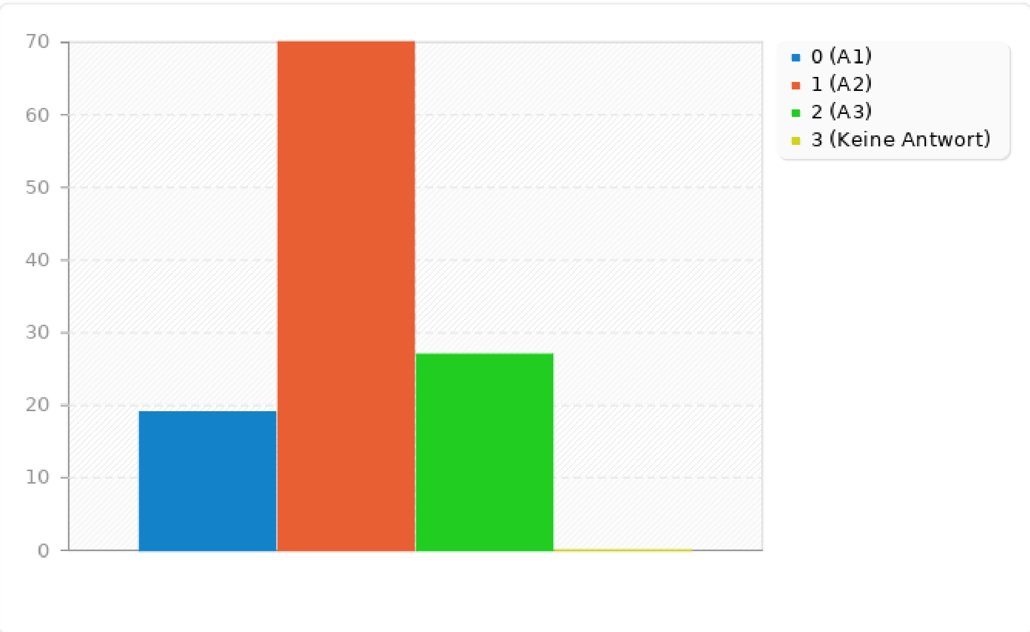
Zusammenfassung für LernspieleBewusst

Nutzen Sie Lernspiele bewusst zum Erlangen von bestimmtem Wissen oder Fertigkeiten?

Antwort	Anzahl	Prozent
ja (A1)	19	16.38%
nein (A2)	70	60.34%
teilweise (A3)	27	23.28%
Keine Antwort	0	0.00%

Zusammenfassung für LernspieleBewusst

Nutzen Sie Lernspiele bewusst zum Erlangen von bestimmtem Wissen oder Fertigkeiten?



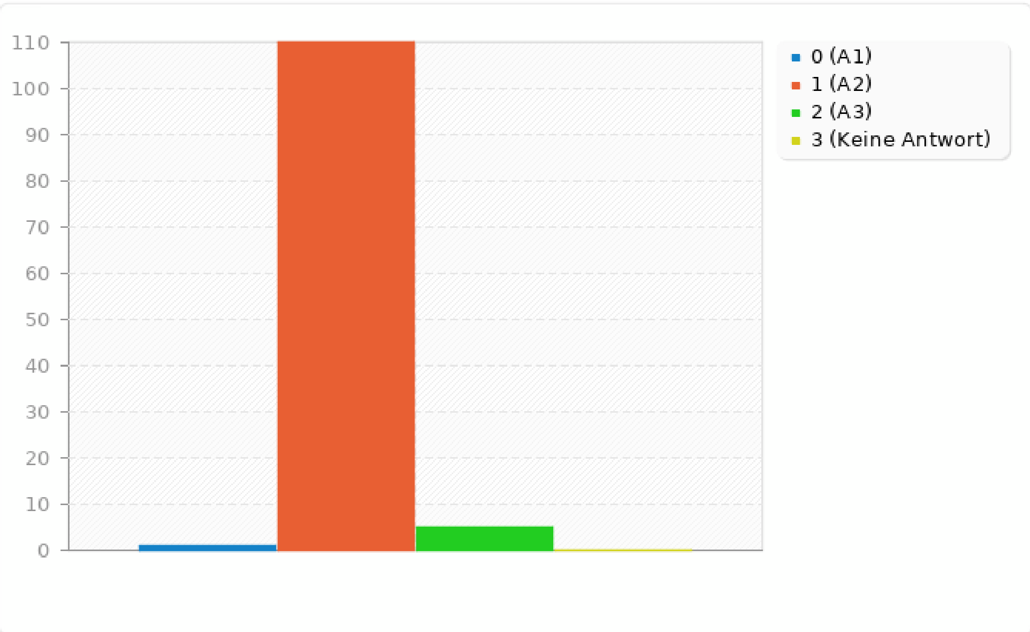
Zusammenfassung für LernspielePlan

Folgen Sie bei der Nutzung von Lernspielen einem bestimmten Zeitplan?

Antwort	Anzahl	Prozent
ja (A1)	1	0.86%
nein (A2)	110	94.83%
teilweise (A3)	5	4.31%
Keine Antwort	0	0.00%

Zusammenfassung für LernspielePlan

Folgen Sie bei der Nutzung von Lernspielen einem bestimmten Zeitplan?



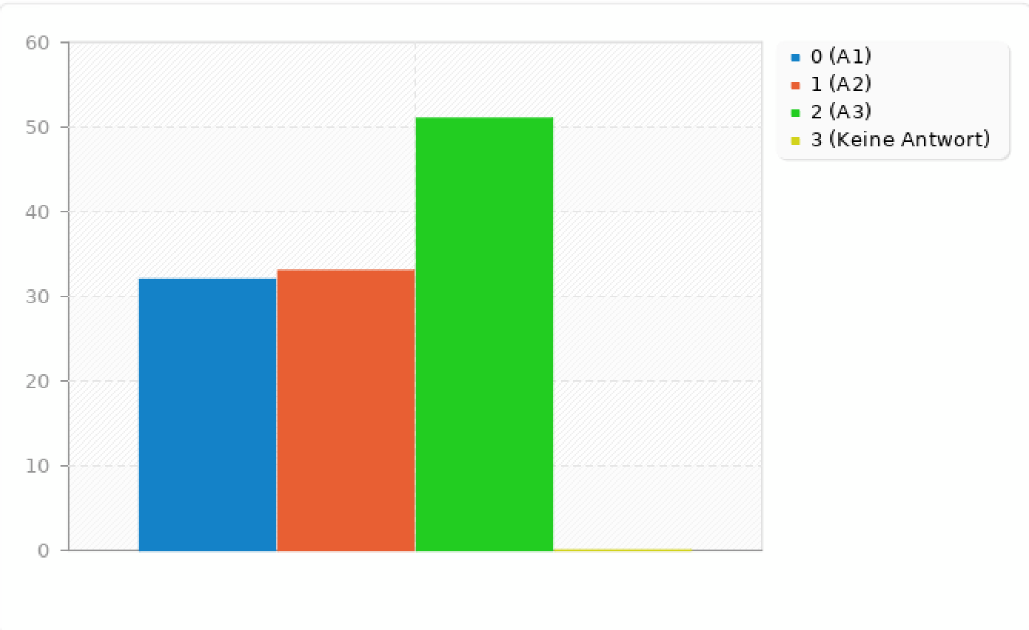
Zusammenfassung für Notifications

Würden Sie der Zeitplanung eines Lernspiels folgen, wenn dieses Sie per Benachrichtigung hieran erinnert?

Antwort	Anzahl	Prozent
ja (A1)	32	27.59%
nein (A2)	33	28.45%
vielleicht (A3)	51	43.97%
Keine Antwort	0	0.00%

Zusammenfassung für Notifications

Würden Sie der Zeitplanung eines Lernspiels folgen, wenn dieses Sie per Benachrichtigung hieran erinnert?



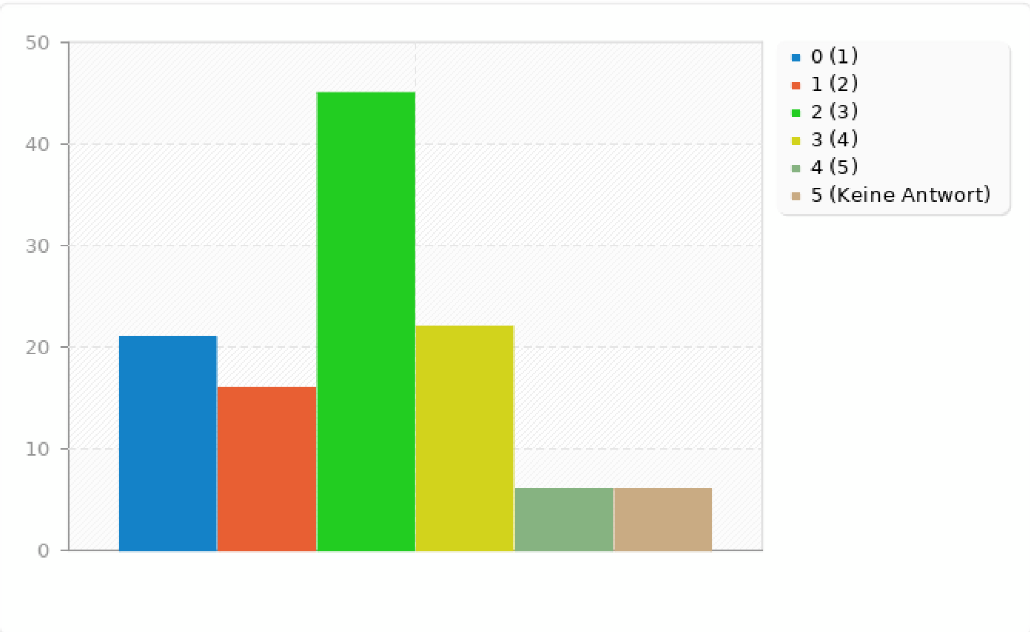
Zusammenfassung für Lernerfolg

Wie schätzen Sie den Lernerfolg bei der Nutzung der von Ihnen verwendeten Lernspiele ein?

Antwort	Anzahl	Prozent	Summe
1 (1)	21	17.50%	30.83%
2 (2)	16	13.33%	
3 (3)	45	37.50%	37.50%
4 (4)	22	18.33%	
5 (5)	6	5.00%	23.33%
Keine Antwort	6	4.76%	0.00%
Arithmetisches Mittel	2.78		
Standard Abweichung	1.14		
Summe (Antworten)	110	100.00%	100.00%
Anzahl Fälle		0%	

Zusammenfassung für Lernerfolg

Wie schätzen Sie den Lernerfolg bei der Nutzung der von Ihnen verwendeten Lernspiele ein?



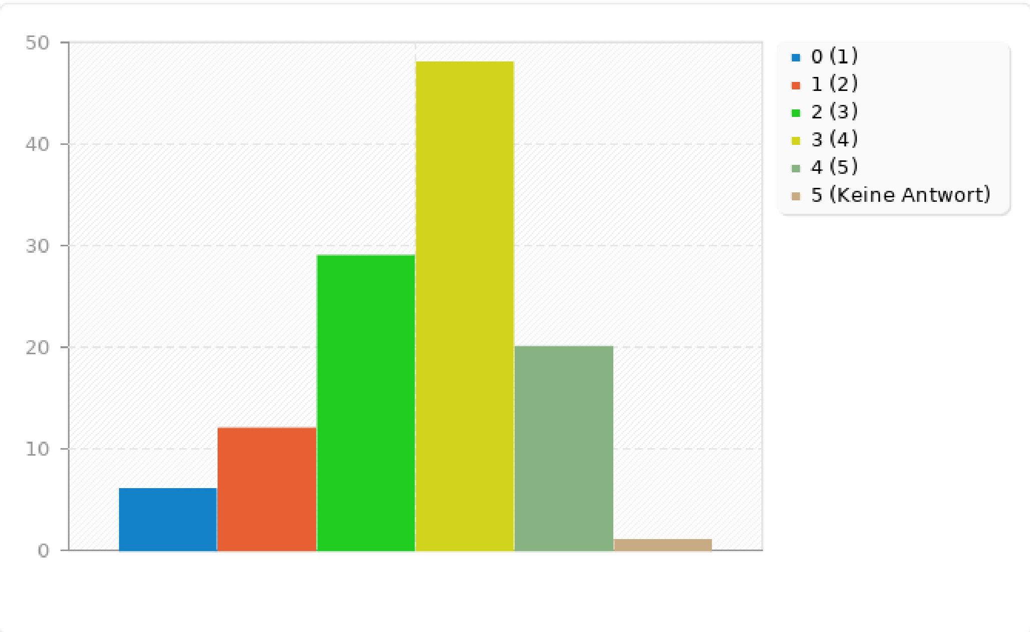
Zusammenfassung für Wissenschaft

Wie hoch ist die Wahrscheinlichkeit, dass Sie Lernspiele eher benutzen würden, wenn sie durch einen wissenschaftlichen Lernansatz gestützt wären?

Antwort	Anzahl	Prozent	Summe
1 (1)	6	4.80%	14.40%
2 (2)	12	9.60%	
3 (3)	29	23.20%	23.20%
4 (4)	48	38.40%	
5 (5)	20	16.00%	54.40%
Keine Antwort	1	0.79%	0.00%
Arithmetisches Mittel	3.56		
Standard Abweichung	1.06		
Summe (Antworten)	115	100.00%	100.00%
Anzahl Fälle		0%	

Zusammenfassung für Wissenschaft

Wie hoch ist die Wahrscheinlichkeit, dass Sie Lernspiele eher benutzen würden, wenn sie durch einen wissenschaftlichen Lernansatz gestützt wären?



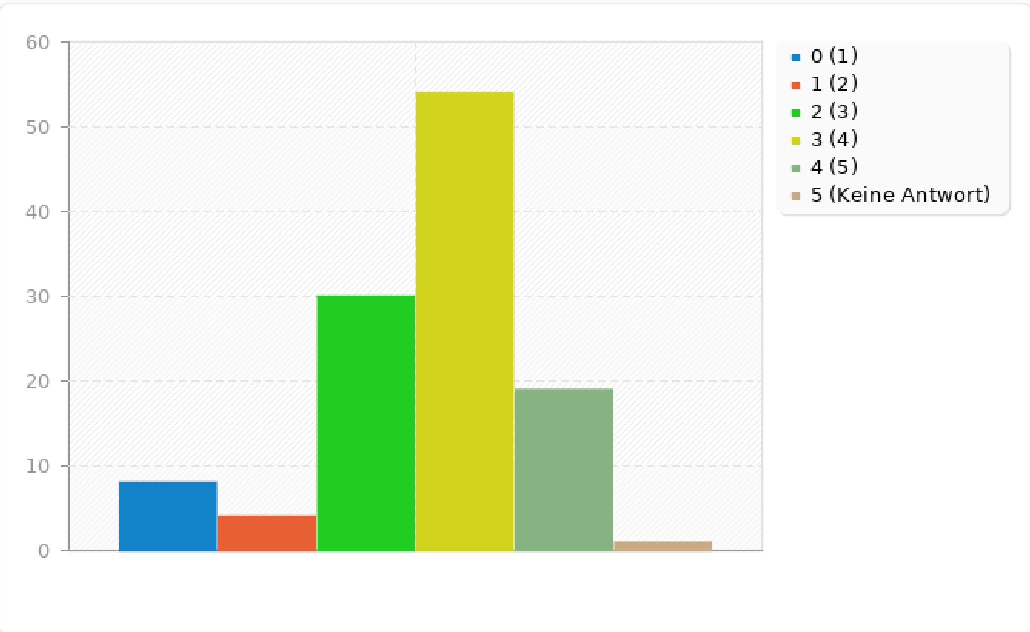
Zusammenfassung für BesseresErgebnis

Wie hoch beurteilen Sie die Wahrscheinlichkeit, dass Lernspiele mit einem wissenschaftlichen Lernansatz zu einem besseren Lernergebnis führen?

Antwort	Anzahl	Prozent	Summe
1 (1)	8	6.40%	9.60%
2 (2)	4	3.20%	
3 (3)	30	24.00%	24.00%
4 (4)	54	43.20%	
5 (5)	19	15.20%	58.40%
Keine Antwort	1	0.79%	0.00%
Arithmetisches Mittel	3.63		
Standard Abweichung	1.03		
Summe (Antworten)	115	100.00%	100.00%
Anzahl Fälle		0%	

Zusammenfassung für BesseresErgebnis

Wie hoch beurteilen Sie die Wahrscheinlichkeit, dass Lernspiele mit einem wissenschaftlichen Lernansatz zu einem besseren Lernergebnis führen?



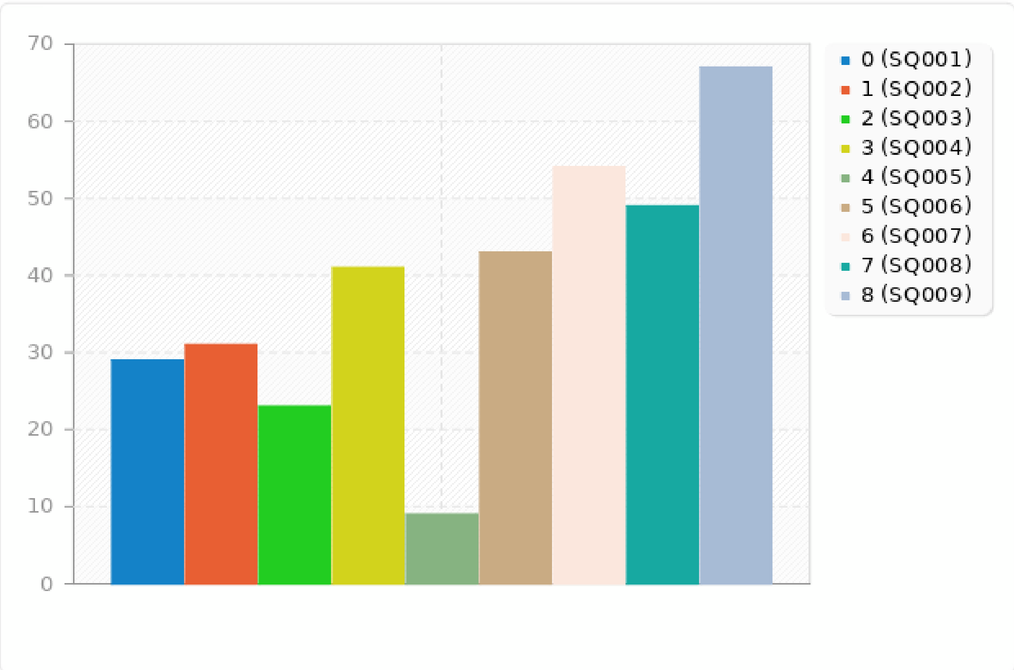
Zusammenfassung für Motivation

Welche Elemente würde Ihre Motivation, ein Lernspiel zu nutzen steigern?

Antwort	Anzahl	Prozent
Augmented Reality (SQ001)	29	25.00%
Ortsbezogene Funktionen (SQ002)	31	26.72%
Zeitbezogene Funktionen (SQ003)	23	19.83%
Multimediale Inhalte (SQ004)	41	35.34%
3D-Elemente (SQ005)	9	7.76%
herausragende grafische Gestaltung (SQ006)	43	37.07%
Ranglisten (SQ007)	54	46.55%
Benachrichtigungen / Erinnerungen (SQ008)	49	42.24%
Wissenschaftlicher Lernmethode (SQ009)	67	57.76%

Zusammenfassung für Motivation

Welche Elemente würde Ihre Motivation, ein Lernspiel zu nutzen steigern?



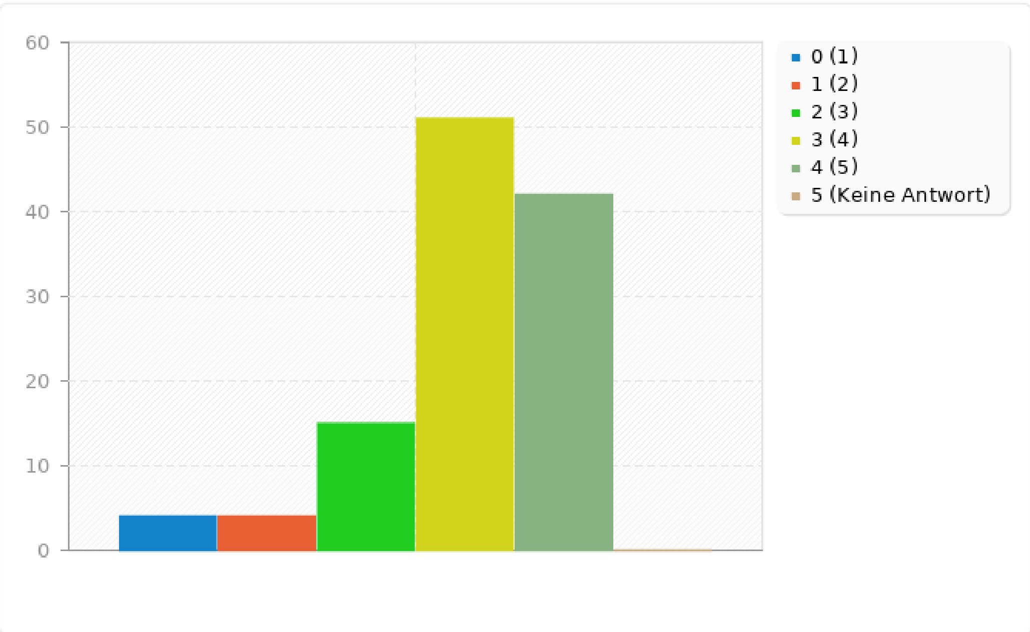
Zusammenfassung für Feedback

Wie wichtig ist Ihnen ein unmittelbares Feedback zu Ihrem aktuellen Lernstand bzw. Lernerfolg?

Antwort	Anzahl	Prozent	Summe
1 (1)	4	3.17%	6.35%
2 (2)	4	3.17%	
3 (3)	15	11.90%	11.90%
4 (4)	51	40.48%	
5 (5)	42	33.33%	73.81%
Keine Antwort	0	0.00%	0.00%
Arithmetisches Mittel	4.06		
Standard Abweichung	0.97		
Summe (Antworten)	116	100.00%	100.00%
Anzahl Fälle		0%	

Zusammenfassung für Feedback

Wie wichtig ist Ihnen ein unmittelbares Feedback zu Ihrem aktuellen Lernstand bzw. Lernerfolg?



Appendix B

Code and data used in the analyses

B.1 JSON data collection

The following string was sent from the "Where is that?" app to the data collection server for each played location. .


```
{       "secretKey": "23498z3bub374238423b42b1234",
        "dataPoints": [
                {"uniqueId": "0000002C-3E80-464F-BC14-927EAD1D4396",
                 "userId": "3A7AA4FE-1FFC-48ED-B6D4-7E91BB7B0DCE",
                 "locationId": "535e4a8f-1c7c-4525-a83b-74280ad00388",
                 "categoryId": "d6687910-81a8-11e3-9629-001851e6d0d1",
                 "correct": 1,
                 "box": 1,
                 "timestamp": 1234553223,
                 "sm2fs_mode": "sm2",
                 "sm2_repetitions": 2,
                 "sm2_q": 3,
                 "sm2_ef": 2.5,
                 "sm2_interval": 6,
                 "sm2_repetitionDate": 1534057200,
                 "fs_repetitions": 0,
                 "fs_rank": 2,
                 "fs_score": 0,
                 "fs_relevance": 0,
                 "fs_flag": 0
                 }
        ]
}
```


Figure B.1: JSON string for data collection (own representation)

The purposes of the different data fields in figure B.1 were as follows:

- **uniqueID:** ID to identify a unique data set
- **userID:** ID to identify a unique user
- **locationID:** ID to identify a location within a category
- **categoryID:** ID to identify a category
- **correct:** indicator if an answer was correct (1) or incorrect (0)
- **box:** indicator in which box of the Leitner System the location is moved
- **timestamp:** date and time when the data set was created
- **sm2fs_mode:** indicator whether the SM2 algorithm or the FS algorithm was used
- **sm2_repetitions:** number of repetitions in SM2 mode
- **sm2_q:** SM2 quality of response
- **sm2_ef:** SM2 easiness factor
- **sm2_interval:** calculated interval until the next repetition according to SM2 in days
- **sm2_repetitionDate:** next scheduled repetition according to SM2
- **fs_repetitions:** number of repetitions in FS mode
- **fs_rank:** rank of the played location according to FS calculation
- **fs_score:** score of the played location according to FS calculation
- **fs_relevance:** relevance of the played location according to FS calculation
- **fs_flag:** FS mode flag that the location was the latest played location

B.2 Code used in R

B.2.1 Code for boxplot creation

```
Testgruppe <- read_excel("/Testgruppe.xlsx")
Kontrollgruppe <- read_excel("/Kontrollgruppe.xlsx")

par(mfrow = c(1,2))

boxplot(Testgruppe $Mittelwert~Testgruppe $Spiel,
xlab="Number of repetitions", xlim=c(0,15),
ylab="Average number of correct answers", ylim=c(0,1.1),
main="Test Group")

boxplot(Kontrollgruppe $Mittelwert~Kontrollgruppe $Wiederholung,
xlab="Number of repetitions", xlim=c(0,15),
ylab="Average number of correct answers", ylim=c(0,1.1),
main="Control Group")

par(mfrow = c(1,1))
```

B.2.2 Code for scatterplot creation

```
colors= c("darkred","darkblue")
colors <- colors[as.numeric(alleSpieler$Gruppe)]
shapes = c(16,16)
shapes <- shapes[as.numeric(alleSpieler$Gruppe)]

plot(alleSpieler$‘Durchschnittsergebnis erste Spiele‘,
alleSpieler$‘Durschnittsergebnis Lernspiele‘,

xlim = c(0,1),ylim = c(0,1), col=colors, pch=shapes, xlab = "Mean Value
for First Game",

ylab = "Mean Value for Learning Games (3 to n)", main = "Mean Value
of Correct Answers",

cex.lab = 0.8, cex.main=1.5) abline(0,1)

legend("bottom", c("Test Group","Control Group"), pch=c(16,16),
col=c("darkred","darkblue"))
```

B.2.3 Code for dendrogram creation

```
average <- agnes(alleSpieler[,3:5],
stand = FALSE, method = "average")
plot(average, labels = alleSpieler$Gruppe, col=c("blue","red"))
```

B.2.4 Code for logistic regression

```
logreg_korrelation_lernspiele <- glm(alleSpieler$sm2_genutzt~
alleSpieler$‘Durschnittsergebnis Lernspiele‘, family = binomial)
summary(logreg_korrelation_lernspiele)

### Erstellung einer Klassifikationstabelle
Lernen_lernspiele_0.3 <- fitted(logreg_korrelation_lernspiele)>0.3
Lernen_lernspiele_0.5 <- fitted(logreg_korrelation_lernspiele)>0.5
Lernen_lernspiele_0.7 <- fitted(logreg_korrelation_lernspiele)>0.7

table(alleSpieler$sm2_genutzt,Lernen_lernspiele_0.3)
table(alleSpieler$sm2_genutzt,Lernen_lernspiele_0.5)
table(alleSpieler$sm2_genutzt,Lernen_lernspiele_0.7)
```

B.3 Data used for analysis

B.3.1 Pre-processed data used for the analysis

The pre-processed data that was used for the analysis within the presented research can be accessed through the following links as SQL dumps:

- Usable Datasets: `https://flo.vc/preprocessed` (3 GB)
- Test Group: `https://flo.vc/testgroup` (29 MB)
- Control Group: `https://flo.vc/controlgroup` (51 MB)

B.3.2 Access to complete database

The complete database containing the data that was used for the analysis within the presented research can be accessed through the following link as an SQL dump:

- Complete Database: `https://flo.vc/complete` (6 GB)