



Fachhochschule Nordwestschweiz  
Hochschule für Angewandte Psychologie

# MASTERARBEIT

Interface Ästhetik und Design von digitalen Lernmedien  
Konstruktion eines Fragebogens zur Erfassung des ästhetischen Empfindens

2017

Autor

Beat Röthlisberger

Betreuende Person

Prof. Dr. Carmen Zahn

Praxispartner

Hochschule für Technik, FHNW Brugg

Institut für 4D-Technologien

**Abstract**

In Switzerland interest in the MINT subjects (mathematics, IT, natural sciences and technology) needs to be encouraged at all levels of education. In order to do that, extracurricular places of learning are offering unique opportunities. New communication methods used by museums and digital media enable positive experiences to be gained in exploring technical phenomena and objects and support interest in learning. An important factor with regard to this is the question as to what influence the interface aesthetics of interactive digital media has on learning about natural sciences. For that reason, the research question for this study is: How can the aesthetic sensibility of young people towards digital learning media (applications) be addressed? In order to do that, a questionnaire was designed in order to address aesthetic sensibility that was checked with two different apps on tablet computers. The sample included 160 young people ( $m = 91$ ,  $f = 69$ , Age:  $M = 13.7$ ,  $SD = 1.04$ ). The statistical validation (PAF, Promax) confirmed the dimensions of "Cognitive Stimulation", "Expressive Aesthetics", "Emotion" and "Usability" and resulted in the adjustment of the questionnaire.

*Keywords*

*Aesthetics, interface aesthetics, applications, aesthetic sensibility, multimedia learning, questionnaire development*

## **Zusammenfassung**

Aufgrund des Mangels an Fachkräften in der Schweiz muss das Interesse an den MINT-Fächern (Mathematik, Informatik, Naturwissenschaften, Technik) auf allen Bildungsstufen gefördert werden. Hierfür bieten ausserschulische Lernangebote wie Naturkunde- und Technikmuseen einzigartige Chancen. Neue museale Vermittlungsmethoden und digitale Medien ermöglichen positive Erfahrungen mit technischen Phänomenen und Objekten und unterstützen das interessierte Lernen. Eine wichtige Rolle spielt dabei die Frage, welchen Einfluss die Interface-Ästhetik interaktiver digitaler Medien auf das Lernen naturwissenschaftlicher Inhalte hat. Die Forschungsfrage dieser Arbeit lautet deshalb: Wie kann das ästhetische Empfinden von Jugendlichen bei digitalen Lernmedien (Applikationen) erfasst werden? Dazu wurde ein Fragebogen zur Erfassung des ästhetischen Empfindens konstruiert, der anhand eines Tests mit zwei verschiedenen Apps auf Tablet-Computern überprüft wurde. Die Stichprobe umfasste 160 Jugendliche ( $m = 91$ ,  $w = 69$ ,  $MW$  Alter = 13.7,  $SD = 1.04$ ). Die statistische Validierung (Hauptachsenanalyse, Promax) bestätigt die Dimensionen „Kognitive Stimulation“, Expressive Ästhetik“, „Emotion“ und „Usability“ und führt zu einer Anpassung des Fragebogens.

## *Schlüsselwörter*

*Ästhetik, Interface-Ästhetik, Applikationen, ästhetisches Empfinden, Multimedia-Lernen, Fragebogenentwicklung*

Zeichen im Bericht, gerundet: 182'000 (mit Leerzeichen, exklusive Anhang)

## **Danksagung**

An dieser Stelle danke ich allen Personen, die mich bei der Erstellung dieser Masterarbeit fachlich wie persönlich unterstützt haben.

Ein grosses Dankeschön gilt meiner Betreuerin Prof. Dr. Carmen Zahn und Magdalena Mateescu, die mich mit wertvollen Hinweisen während des Arbeitsprozess unterstützt und begleitet haben.

Weiter bedanke ich mich herzlich bei den Lehrpersonen, die sich mit ihren Klassen für den Lern-App Test zur Verfügung gestellt und durch ihre Unterstützung zum Gelingen der empirischen Untersuchung beigetragen haben. Besonders bedanke ich mich bei allen Schülerinnen und Schülern für das interessierte und engagierte Ausfüllen des Fragebogens.

## **Abkürzungsverzeichnis**

App	Application Software (Applikation)
MINT	Mathematik, Informatik, Naturwissenschaft und Technik
MW	Mittelwert
SD	Standardabweichung
Solar Walk	SW
Physics at school Lite	PHL

## Inhaltsverzeichnis

1. Einleitung.....	1
2. Theorie .....	4
2.1. Ästhetische Erfahrung.....	5
2.2. Ästhetik von Websites .....	10
2.2.1. Definitionen.....	10
2.2.2. Einfluss der Betrachtungsdauer auf wahrgenommene Ästhetik.....	14
2.2.3. Personenvariablen.....	15
2.2.4. Usability und User Experience .....	20
2.2.5. Zentrale Variablen der Website-Rezeption .....	24
2.3. Lernkontext.....	25
2.3.1. Emotion .....	26
2.3.2. Lernmotivation .....	29
2.3.3. Emotional Design .....	31
2.4. Arbeitsmodell.....	33
3. Methoden .....	36
3.1. Untersuchungs-Design.....	36
3.2. Methodische Grundlagen zur Fragebogenkonstruktion.....	37
3.2.1. Itemformulierung.....	37
3.2.2. Itempolung.....	37
3.2.3. Skalenniveau.....	38
3.2.4. Testgütekriterien.....	39
3.2.5. Objektivität .....	39
3.2.6. Reliabilität .....	40
3.2.7. Validität .....	41
3.3. Fragebogenkonstruktion .....	43
3.4. Testkonstruktion .....	48
3.5. Stichprobe .....	50
3.6. Testdurchführung.....	51
3.6.1. Methodische Grundlagen zur Datenauswertung.....	53
3.6.2. Itemanalyse.....	53
3.6.3. Rohwertverteilung .....	53
3.6.4. Itemschwierigkeit .....	54

---

3.6.5.	Trennschärfe .....	54
3.6.6.	Homogenität .....	55
3.6.7.	Dimensionalität.....	55
3.6.8.	Dateneignung.....	56
3.6.9.	Methoden für exploratorische Faktorenanalyse .....	58
3.6.10.	Extraktionskriterien für Faktoren.....	58
3.6.11.	Wahl der Rotationsmethode.....	59
3.6.12.	Weitere Werte der Faktorenanalyse.....	59
3.6.13.	Zahlenwerte für Items und Skalen .....	59
3.7.	Vorgehen bei der statistischen Auswertung.....	60
3.8.	Bezugnahme zu Gütekriterien.....	62
4.	Ergebnisse .....	64
4.1.	Deskriptive Statistik.....	64
4.2.	Faktorenanalyse .....	69
4.3.	Reliabilitätsanalyse .....	78
4.4.	Mittelwertvergleiche .....	81
4.5.	Korrelationen .....	83
4.6.	Erläuterungen zu den Ergebnissen.....	85
4.6.1.	Itemauswahl für die Faktorenanalyse .....	85
4.6.2.	Benennung der Faktoren als Dimensionen.....	86
4.7.	Zusammenfassung der Ergebnisse .....	86
4.8.	Anpassung des Fragebogen nach der statistischen Überprüfung.....	89
5.	Diskussion.....	91
5.1.	Theorie .....	91
5.2.	Methoden .....	92
5.3.	Ergebnisse .....	94
5.4.	Implikationen für die Praxis und Ausblick .....	99
6.	Literaturverzeichnis .....	103
7.	Anhang.....	107

## 1. Einleitung

Das Verständnis naturwissenschaftlicher und technischer Vorgänge ist eine Voraussetzung dafür, Phänomene und Prozesse des täglichen Lebens zu begreifen. Dadurch wird bewusstes und verantwortungsvolles Handeln möglich (Fachhochschule Nordwestschweiz, 2017).

Der wirtschaftliche Erfolg und damit der hohe Lebensstandard in der Schweiz hängen zu einem erheblichen Teil von naturwissenschaftlich-technischen Innovationen ab. Die Innovationskraft ist jedoch durch einen Mangel an Fachkräften gefährdet. Deshalb muss das Interesse an den MINT-Fächern (Mathematik, Informatik, Naturwissenschaften, Technik) auf allen Bildungsstufen gefördert werden (Fachhochschule Nordwestschweiz, 2017).

Diese Masterarbeit wird im Rahmen des Forschungsprojekts „Einfluss der Interface-Ästhetik digitaler Medien auf das Lernen von MINT-Themen in Ausstellungen“ durchgeführt. Am Projekt sind drei Institute der Fachhochschule Nordwestschweiz (FHNW) beteiligt: Das Institut für Kooperationsforschung der Hochschule für Angewandte Psychologie, das Institut für 4D-Technologien der Hochschule für Technik (FHNW) sowie das Institut Ästhetische Praxis und Theorie der Hochschule für Gestaltung und Kunst. Praxispartner ist das Institut für 4D-Technologien in Brugg.

Bei der Förderung des Interesses an MINT-Themen sind individuelle Lernbiografien von Bedeutung. Ein wichtiger Aspekt bilden dabei nach Zahn und Agotai (2015) positive Affekte und Begegnungen mit technischen Phänomenen und Objekten. Dafür bieten außerschulische Lernangebote wie zum Beispiel Naturkunde- oder Technikmuseen sowie Technikausstellungen einzigartige Chancen. Sie können mit der anschaulichen Vermittlung von Phänomenen den Praxisbezug herstellen, Jugendliche mithilfe von interaktiven Ausstellungsstücken aktiv einbinden und somit den Spass an der Technik fördern (acatech, 2011, zitiert nach Zahn & Agotai, 2015). Neue museale Vermittlungsmethoden und digitale Medien (z. B. interaktive Multitouch



Walls, Tablets und Apps) können bei Kindern und Jugendlichen positive Erfahrungen mit technischen Phänomenen und Objekten sowie interessiertes Lernen unterstützen (Tanner et al., 2014, zitiert nach Zahn & Agotai, 2015).

Dabei stellt sich nach Zahn und Agotai (2015) die Frage, welche Rolle die Interface-Ästhetik in Bezug auf positive Affekte und das Interesse an MINT-Themen spielt.

Um die Interface-Ästhetik digitaler Lernmedien beurteilen zu können, braucht es ein geeignetes Messinstrument. In diesem Zusammenhang leistet diese Masterarbeit einen Beitrag, indem ein Instrument (Fragebogen) zur Erfassung des ästhetischen Empfindens bei Jugendlichen konzipiert wird. Die Fragestellung lautet diesbezüglich:

Wie kann das ästhetische Empfinden von Jugendlichen bei der Verwendung digitaler Lernmedien (Applikation) erfasst werden?

Die Fragestellung weist darauf hin, dass das ästhetische Empfinden in einem klar umrissenen Kontext untersucht wird. Beim Interface handelt es sich um den Bildschirm eines Tablet-Computers (Pad) auf dem eine Applikation (App) beurteilt wird. Die Untersuchung befasst sich mit Jugendlichen in einem Lernkontext und nicht etwa in der Freizeit. Inhaltlich bewegen sich die Apps im Themenbereich MINT.

Für die Messung von Ästhetik oder ästhetischem Erleben gibt es bereits eine Vielzahl von validierten Messinstrumenten. Allerdings berücksichtigt keines vollumfänglich den oben beschriebenen Kontext oder die zu untersuchende Zielgruppe. Einerseits befassen sie sich mit anderen Objekten wie zum Beispiel Kunstgegenständen, Alltagsprodukten oder Webseiten, andererseits sind die Fragebogen in der Regel für Erwachsene konzipiert und werden häufig mit Studenten validiert. In diesem Sinne ist das Ziel dieser Masterarbeit, einen Fragebogen zu entwickeln, der auf die Kombination von Interface-Ästhetik, Applikation, Lernkontext und Jugendliche zuge-

schnitten ist. Weiter erfolgt eine erste Validierung des Fragebogens mit Jugendlichen, die in der Schweiz zur Schule gehen.

Grundlage eines jeden Fragebogens sind Theorien und daraus abgeleitete Konstrukte, die das Messinstrument erfassen soll. Hierzu wird ein Arbeitsmodell zu „ästhetischem Empfinden“ entwickelt, das die relevanten Konstrukte und Dimensionen in dem spezifischen Kontext beschreibt. Daneben findet in oben erwähntem Projekt eine fundierte Experten-Diskussion im über das Konstrukt "Interface-Ästhetik" statt, wobei derzeit das Konstrukt selbst theoretisch weiterentwickelt wird.

Die Bezeichnung „ästhetisches Empfinden“ wird deshalb gewählt, weil das Konstrukt im spezifischen Kontext dieser Masterarbeit verwendet wird. Es soll sich somit klar von Ausdrücken anderer Autorinnen und Autoren wie „ästhetisches Erleben“ oder „ästhetische Erfahrung“ unterscheiden.

Im zweiten Kapitel sind die theoretischen Grundlagen erläutert, die als Grundlage des Fragebogens das Arbeitsmodell zu „ästhetischem Empfinden“ beschreiben. Kapitel drei befasst sich mit den Methoden zur Fragebogen- und Testkonstruktion, zur empirischen Untersuchung im Feld und zu den Auswertungsmethoden. In Kapitel vier sind die Ergebnisse sowie die Anpassung des Fragebogens dargestellt. Kapitel fünf diskutiert Theorie, Methode und Ergebnisse und leitet aus den Erkenntnissen Implikationen für die Praxis ab.

## 2. Theorie

Um die Fragestellung „Wie kann das ästhetische Empfinden von Jugendlichen bei der Verwendung digitaler Lernmedien (Applikation) erfasst werden?“ zu beantworten, braucht es zunächst eine Definition, welche Konstrukte bei der Beurteilung der Interface-Ästhetik von Bedeutung sind. Hierzu wird ein Arbeitsmodell entwickelt, welches in diesem Kapitel vorgestellt wird.

Ästhetik und die Wahrnehmung von Ästhetik sind Gegenstände vieler Forschungsbereiche, insofern werden hier nur jene vorgestellt, die für die Fragestellung relevant sind. Aus dem Bereich der Wahrnehmungspsychologie stammt das Modell von Leder, Belke, Oeberst und Augustin (2004) zum ästhetischen Erleben. Hier spielen bottom-up und top-down-Prozesse bei der kognitiven Verarbeitung eine wesentliche Rolle. Im Bereich der Mensch-Computer-Interaktion (HCI) sind Ästhetik und Design der Schnittstelle wesentliche Faktoren, um die Nutzung zu erleichtern, positive Emotionen auszulösen und entsprechend Ziele zu erreichen. Dabei können Konzepte und Messinstrumente aus dem Website-Design wie zum Beispiel von Thielsch (2008) sowie Lavie und Tractinsky (2004) und dem Bereich User Experience (UX) wie zum Beispiel von O'Brian und Toms (2010), Minge und Riedel (2013), sowie Sauro (2015) einen Beitrag leisten. Da es noch wenige spezifische Arbeiten zu Applikationen gibt, werden Webseiten und Apps als gleichwertig angesehen.

Um den Lernkontext (Lern-Emotionen, Multimedia-Learning) zu berücksichtigen, fließen unter anderem Arbeiten von Pekrun und Perry (2014), Um, Plass, Hayward und Homer (2012) sowie Mayer (2009) in das Arbeitsmodell ein. In folgenden Abschnitten werden die Konzepte vorgestellt, die dann in das Arbeitsmodell der ästhetischen Empfindung integriert werden.

## 2.1. Ästhetische Erfahrung

Leder et al. (2004) entwickelten basierend auf den Analysen moderner Kunst ein prozessorientiertes Modell der ästhetischen Erfahrung (Abbildung 1).

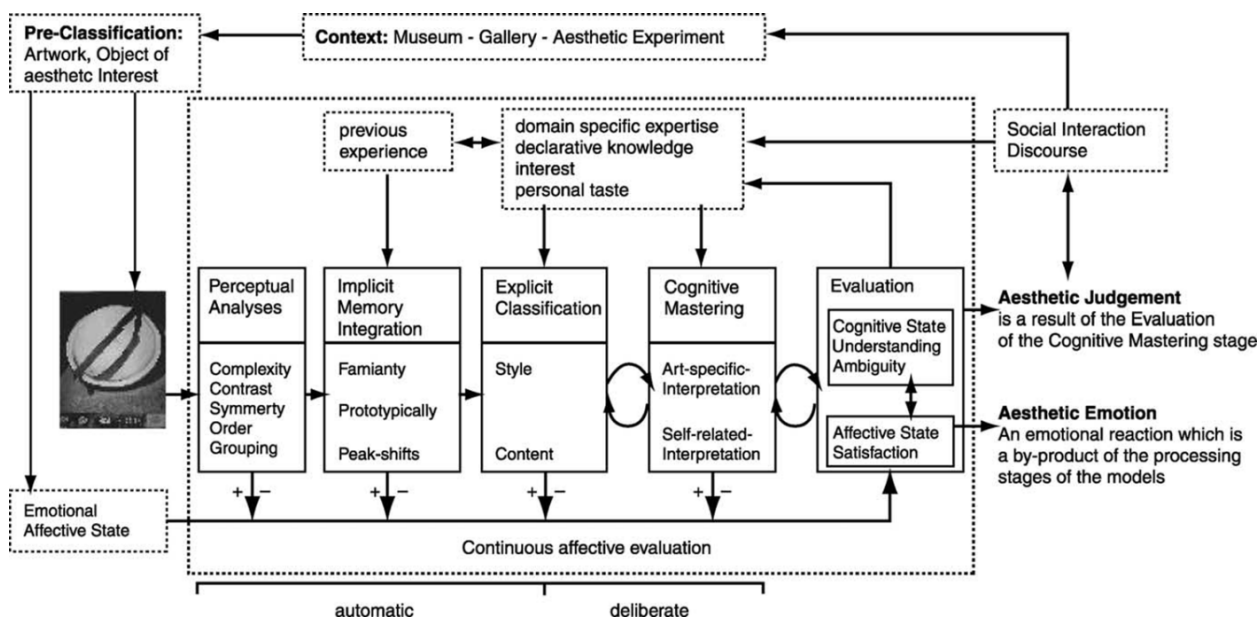


Abbildung 1: Modell der ästhetischen Erfahrung nach Leder et al. (2004, S. 492)

Belke und Leder (2006) schreiben, dass sich unter den Quellen, die für Entwicklung dieses Modells wichtig waren, den Ansatz von Kreidler (1984, zitiert nach Belke & Leder, 2006) zur „Kognitiven Orientierung“ und die Befunde der Wahrnehmungspsychologie über Bedeutungszuweisungsprozesse in der visuellen Wahrnehmung finden lassen. Zudem ist es ein Versuch, die vorliegenden Forschungsbefunde seit den Ursprüngen der Ästhetikforschung mit Fechner (1871, zitiert nach Belke & Leder, 2006) und Wundt (1974, zitiert nach Belke & Leder, 2006) zu integrieren. Das Modell ist rezeptionsästhetisch ausgerichtet. Es versucht zu beschreiben, welche zugrunde liegenden universellen mentalen Prozesse des Rezipienten das ästhetische Erleben bedingen. Weiter diskutiert es Variablen des Kontextes und des Rezipienten, die den Prozess modulieren können und bildet somit Rahmenbedingungen für die empirische Ästhetikforschung.

Das Modell der ästhetischen Erfahrung beschäftigt sich nach Belke und Leder (2006) vor allem mit visueller Kunst, insbesondere auf die Verarbeitung von Merkmalen der Gegenwartskunst, wenngleich ein Transfer auf andere Formen ästhetischer Erfahrung denkbar ist.

Das Modell beschreibt unter anderem verschiedene Stadien, die im Prozess der Verarbeitung beteiligt sind: „Perzeptuelle Analyse“, „Implizite Gedächtnisintegration“, „Explizite Klassifikation“, „Kognitive Bewältigung“ und „Evaluation“. Dabei handelt es sich nicht um eine strikt serielle Abfolge von Phasen, sondern um einen hierarchischen Prozess. Dieser kann im Verlauf auf frühere Phasen zurückfallen und die Elemente sind teilweise durch Feedbackschlaufen miteinander verbunden (Belke & Leder 2006).

Input des Modells ist ein Kunstwerk in einem situativen Kontext (Kino, Galerie, Museum), wobei der kontextuelle Bezugsrahmen beim Rezipienten bereits eine spezifische Erwartungshaltung provoziert. Nebenbei stellt der affektive Ausgangszustand des Rezipienten einen weiteren wesentlichen Kontextaspekt dar. Im Folgenden werden die einzelnen Verarbeitungsstufen gemäss Belke und Leder (2006) zusammenfassend erläutert.

### **Perzeptuelle Analyse**

In der ersten Stufe der kognitiven Verarbeitung, der perzeptuellen Analyse, sind grundlegende okzipitale (die rückwärtige Grosshirnrinde betreffende) Verarbeitungsprozesse beteiligt, die bei jeder Form der visuellen Wahrnehmung und Objektidentifizierung eine Rolle spielen. Beteiligte Variablen in dieser Phase sind Visuelle Komplexität, Symmetrie, Kontrastverteilung und Grouping, Klarheit der Repräsentation und Farbigkeit. Auf dieser Stufe kann Grouping (Zusammenfügen von Information) als grundlegender Mechanismus angenommen werden. Die Verarbeitung von perzeptuellen Merkmalen geschieht sehr schnell und braucht keine willkürliche Aufmerksamkeit oder Bewusstheit („automatic“).

### **Implizite Gedächtnisintegration**

In der impliziten Gedächtnisintegration werden die Ergebnisse der Wahrnehmungsanalyse in Relation zur individuellen Vorerfahrung des Rezipienten gesetzt. Die Effekte müssen nicht bewusst sein, um die ästhetische Wahrnehmung zu beeinflussen. Prototypikalität (Repräsentativität eines Objektes für eine zugehörige Oberklasse) kann dabei eine wesentliche Determinante zur ästhetischen Wertschätzung sein. Eine andere Variable, die ästhetische Wertschätzung in diesem Stadium beeinflusst, ist Vertrautheit. Durch das „mere-exposure“-Paradigma können Präferenzen alleine durch die wiederholte Darbietung erhöht werden. Mit dem „peak-shift“-Phänomen ist ein künstlerisches Prinzip gemeint, ein Objekt nicht nur hinsichtlich wesentlicher Eigenheiten (Form, Farben, Bewegung) darzustellen, sondern diese Merkmale, in Entsprechung zur Karikatur, so zu verstärken, dass Mechanismen der Wahrnehmung optimal angesprochen werden.

### **Explizite Klassifikation**

Auf der Stufe der expliziten Klassifikation erfolgt eine willkürliche (bewusstseinsfähige, „deliberate“) Auseinandersetzung mit dem ästhetischen Objekt, deren Inhalte auch verbalisiert werden können. Dieses Verarbeitungsstadium setzt entweder explizite Repräsentationen des Inhalts oder über den Stil des Objekts voraus. Welches dieser Aspekte im Vordergrund steht, hängt vom Kunstwerk oder der Expertise des Betrachters ab. Der Erwerb von Stilrepräsentation ist zwar auch ein Prozess des impliziten Lernen und als Fertigkeit schon im Kindesalter möglich, allerdings scheint explizites Wissen förderlich für das Lernen und Unterscheiden von Stil zu sein. Im Modell wird vermutet, dass die erfolgreiche Klassifikation von Stil und dargestelltem Inhalt an sich eine selbstbelohnende, affektiv positive Erfahrung ist, die eine ästhetische Evaluation des Objekts positiv beeinflussen kann. Die Ergebnisse und Inhalte dieser Klassifikationsstufe werden einer weitergehenden Reflexion unterzogen.

## **Kognitive Bewältigung**

Auf dieser Verarbeitungsstufe kommt es zu stark konstruktivistischer Verarbeitung und neuen Konzeptbildungen, in dem der Rezipient versucht, das ästhetische Gemeinte zu rekonstruieren. Es handelt sich hier um intentionale, top-down geleitete Verarbeitungsprozesse, in denen der Betrachter beispielsweise Hypothesen über das ästhetische Objekt generiert und überprüft. Eine Expertise in der Kunst macht es wahrscheinlich, dass zur Interpretation und Bedeutungsfindung eher kunstbezogene Bereiche des Wissens zur Verfügung stehen, während für ungeübte Betrachter selbst wahrgenommene Wissensbestände mit starkem Selbstbezug assoziiert werden. Aus diesem Grund nimmt das Modell eine Unterscheidung von zwei unterschiedlichen Reflexionsformen vor: eine selbstbezogene, personale Reflexionsweise sowie eine kunstspezifischere, stärker objektbezogene Reflexionsform.

## **Evaluation**

Auf der Stufe der Evaluation wird der Erfolg der Bedeutungsfindung der kognitiven Bewältigungsphase beurteilt. Dies führt zu weiterer Informationsverarbeitung, wenn der derzeitige Stand des Verstehens unbefriedigend oder unklar ist. Wesentlich für dieses Stadium ist die Integration von kognitiven (Umfang des Verständnisses) und emotionalen Aspekten (Grad der Spannung). Die Stufe „Evaluation“ ist über eine Feedbackschleife mit der Stufe „Kognitive Bewältigung“ verbunden, mit dem Ziel, die erlebte Ambiguität (Vieldeutigkeit) wesentlich zu reduzieren. Die vollständige Auflösung von Ambiguität ist auch bei erheblicher Expertise nicht garantiert und bei ästhetischen Objekten allenfalls gar nicht möglich. Möglicherweise ist es sogar ein konstituierendes Merkmal der ästhetischen Erfahrung, dass eine eindeutige Auflösung nicht möglich ist und vom Betrachter auch akzeptiert wird.

Das Modell unterscheidet nach Belke und Leder (2006) zwei ästhetische Outputformen, die „ästhetische Emotion“ und das „ästhetische Urteil“. Erstere beruht auf dem subjektiven Erfolg der Verarbeitung des ästhetischen Objektes und dessen affektiven Begleiterscheinungen. Sie

wird als aktueller affektiver Zustand auf der Stufe der Evaluation verstanden. Die „ästhetische Emotion“ bildet sich durch einen kontinuierlichen affektiven Evaluationsprozess, in dem der subjektive Erfolg der Verarbeitung innerhalb der einzelnen Phasen entscheidend ist. Demzufolge kann jede Stufe einen positiven oder negativen Effekt auf die ästhetische Lust des Betrachters haben. Die affektiven Begleitreaktionen können sich akkumulieren, positive affektive Konsequenzen können durch negative überlagert werden und umgekehrt.

Das „ästhetische Urteil“ ist sowohl von den Resultaten der kognitiven Bewältigung als auch vom Erfolg der Verarbeitung und dessen affektiven Konsequenzen abhängig. Welcher der beiden Outputs wichtiger ist, wird durch die Natur des ästhetischen Urteils bestimmt. Das Urteil des Gefallens bezieht sich eher auf den affektiven Zustand, wobei ein Urteil über die ästhetische Qualität eher als kognitives Resultat der ästhetischen Verarbeitung interpretiert werden kann. Die „ästhetische Emotion“ und das „ästhetische Urteil“ gehen wahrscheinlich meistens in die gleiche Richtung, sie können aber auch im Fall einer positiven ästhetischen Emotion voneinander abweichen. Ein Experte kann beispielsweise zum Schluss kommen, dass es sich um ein Werk von geringer Qualität handelt, dennoch den Prozess der erfolgreichen Bewältigung als affektiv positiv wahrgenommen haben. Ein Auseinandergehen von Affekt und Urteil ist aufgrund der ausgedehnteren Evaluationsphase bei Experten wahrscheinlicher als bei naiven Betrachtern, für die generell eine stärker lustbasierte Bewertung des ästhetischen Objekts angenommen werden kann.

Das Modell beinhaltet eine Variable, in der Effekte von domänenspezifischem und deklarativem Weltwissen auf den Verarbeitungsprozess zusammengefasst wird. Diese Expertise-Variablen sind besonders bedeutend in der experimentellen Forschung. Sie vermitteln zwischen den Prozessen der „expliziten Klassifikation“ und der „kognitiven Bewältigung“. Der Prozess der ästhetischen Erfahrung wird unter Umständen nicht immer vollständig durchlaufen. Ein frühzeitiger Abbruch ist beispielsweise durch stereotype Reaktionen, bedingt durch



einen festgelegten persönlichen Geschmack, denkbar. Insbesondere bei naiven Betrachtern ist davon auszugehen, dass ästhetische Urteile sehr schnell auch aufgrund spontaner, visueller Eindrücke, wie der Farbigkeit oder Typikalität der Darstellung, gefällt werden und eine umfangreiche Evaluation gar nicht stattfindet. Da der affektive Zustand jedoch jederzeit abrufbar ist, sollte das ästhetische Urteil in diesem Fall stark von der ästhetischen Emotion abhängen (Belke & Leder, 2006).

### **Messinstrumente**

In Bezug auf das Modell von Leder et al. (2004) entwickelten Hager, Hagemann, Danner und Schankin (2012) den Ars Reception Survey (ARS). Der Fragebogen zur Wahrnehmung von Kunst beinhaltet fünf Dimensionen. In Klammern ist der Wert von Cronbachs Alpha ( $\alpha$ ) und die Anzahl der Items (I) angegeben. Kognitive Stimulation ( $\alpha=.90$ , I=5), negative Emotionalität ( $\alpha=.85$ , I=5), Expertise ( $\alpha=.85$ , I=5), Selbstbezug ( $\alpha=.85$ , I=4), Künstlerische Qualität ( $\alpha=.85$ , I=5) und positive Attraktivität ( $\alpha=.83$ , I=5). Die Testpersonen beurteilen die Aussagen auf einem 5-stufigem Antwortformat (vgl. Kapitel 3.3, Tabelle 4).

## 2.2. Ästhetik von Websites

In diesem Kapitel wird die Definition von Thielsch (2008) bezüglich Website-Ästhetik erläutert sowie auf weitere Konstrukte und Variablen, die bei der Rezeption von Websites eine Rolle spielen, eingegangen.

### 2.2.1. Definitionen

Gemäss Thielsch (2008) ist die Anwendbarkeit des Leder-Modells auf die Wahrnehmung von Websites unter leichten Änderungen gut denkbar. Dabei ist anzunehmen, dass die ersten drei Stufen direkt auf Websites übertragbar sind. Expertise bezüglich Kunst kann durch Expertise der Webnutzer punkto World Wide Web ausgetauscht werden. Da Websites meistens weniger Interpretationsspielraum zulassen als Kunstwerke, sind die vierte und fünfte Stufe diesbezüglich allgemeiner zu fassen. Dennoch sind ästhetische Urteile und ästhetische Gefühlszustände

gut denkbar. Andere Kontextfaktoren wie Museum, sind durch Faktoren im Kontext des Webnutzers zu ersetzen. Durch die höhere Komplexität von Kunstwerken werden die Begriffe wahrnehmen, erleben und urteilen getrennt und mit differenzierender Bedeutung behandelt. Bei Websites werden zumeist Mehrdeutigkeiten oder Interpretationen jenseits des direkt wahrnehmbaren vermieden. Zudem ist die Kenntnis weiterer Kunstwerke (Oeuvre) eines Webdesigners nicht nötig.

Für seine Studien fasst Thielsch (2008) Grundannahmen und Ergebnisse verschiedener Arbeiten und empirischer Untersuchungen zur Ästhetik in folgenden zwölf Punkten zusammen:

1. Die Ästhetik von Websites stellt einen substantiellen Faktor in der Wahrnehmung des Betrachters dar und ist empirisch messbar. Dabei wird dem wissenschaftlichen Ansatz Fechners, 1876, zitiert nach Thielsch, 2008) gefolgt.
2. Ästhetik wird demnach im Sinne Fechners als Wohlgefallen definiert.
3. Ästhetische Reize müssen, um wahrgenommen zu werden, in ihrer Intensität eine gewisse Schwelle überschreiten. Die Wahrnehmung dieser Reize kann bewusst, vorbewusst oder unbewusst erfolgen. Hierbei spielen top-down- und bottom-up-Prozesse eine Rolle (Leder et al., 2004, zitiert nach Thielsch, 2008).
4. Die ästhetische Wirkung einer Website formiert sich aus mehreren Variablen; das Konstrukt ist voraussichtlich mehrdimensional zu betrachten (Lavie & Tractinsky, 2004; Moshagen, 2005, zitiert nach Thielsch, 2008). Zurzeit sind weder alle beteiligten Variablen schlüssig identifiziert, noch deren Wirkweise ausreichend erforscht. Folgt man Fechners zweitem Prinzip und den Annahmen der Gestaltpsychologie, kann bei diesen Reizkomponenten eine Nichtsummativität vermutet werden: Die Gesamtwirkung der Reizkomposition stimmt nicht zwingend mit der Summe der Teile überein.

5. Erinnerungen können entsprechend dem Assoziationsprinzip in Form von top-down-Prozessen zu ästhetischen Wahrnehmungen beitragen (Fechner, 1876; Leder et al., 2004; Norman, 2004, zitiert nach Thielsch, 2008).
6. Die ästhetische Wahrnehmung ist subjektiv, jedoch statistisch über verschiedene Rezipienten hinweg vergleichbar.
7. Die Betrachter sind in der Lage, ihre Wahrnehmung passend einzuschätzen und zu verbalisieren (dieser Annahme folgen prinzipiell alle empirischen Untersuchungen zur Website-Ästhetik).
8. Das ästhetische Urteil erfolgt sofort (Leder et al., 2004 zitiert nach Thielsch, 2008). Diese sehr schnell getroffenen Ästhetik-Einschätzungen sind vermutlich stabil, könnten aber auf einzelnen eher holistischen Variablen beruhen.
9. Die Website-Ästhetik beschränkt sich nicht nur auf eine Sinnesmodalität. Im Moment ist vor allem die visuelle Wahrnehmung von Websites bedeutsam, da meistens die anderen Sinneskanäle nicht angesprochen werden.
10. Die Ästhetik einer Website beeinflusst die Präferenz der Rezipienten (Schenkman & Jönsson, 2000; van der Heijden, 2003, zitiert nach Thielsch, 2008). Dazu gibt es noch keinen direkten experimentellen Beweis, aber empirische Hinweise bei Haig (2002, zitiert nach Thielsch, 2008) und Hall und Hanna (2004, zitiert nach Thielsch, 2008).
11. Eventuell ist die Wahrnehmung von Website-Ästhetik durch Personenvariablen beeinflusst, möglicherweise spielt hier die Variable ästhetische Sensitivität eine Rolle.
12. Den wichtigsten Aspekt für den Betrachter stellt meistens der Inhalt einer Website dar, Ästhetik ist aber gleichwohl bedeutsam (z. B. Tarasewich et al., 2001, zitiert nach Thielsch, 2008). Für den Rezipienten ergibt sich die visuelle Attraktivität einer Website abgesehen vom Inhalt aus der Wahrnehmung von Ästhetik und Usability (Hassenzahl, 2001; van der Heijden,

2003, zitiert nach Thielsch, 2008). Beide Konstrukte stehen in einer Beziehung zueinander, der empirisch noch nicht geklärt ist. Erste Ergebnisse sprechen dafür, dass Ästhetik nicht im Widerspruch zur Usability steht, unter Umständen sogar Usability-Probleme kompensieren kann (Moshagen, 2007, zitiert nach Thielsch, 2008). Ferner kann ein positiver Einfluss von Ästhetik auf die Wahrnehmung von Glaubwürdigkeit und Vertrauenswürdigkeit von Websites angenommen werden.

Als Folge seiner Studien definiert Thielsch (2008, S. 256) die Ästhetik von Websites wie folgt:

„Visuelle Ästhetik bezeichnet das subjektiv empfundene Wohlgefallen an einer Website. Die ästhetische Wahrnehmung einer Website zeichnet sich sowohl durch einen positiven emotionalen Ausdruck sowie eine positive kognitive Bewertung aus. Die affektiven und kognitiven Prozesse werden durch einzelne Gestaltungsmerkmale einer Website ausgelöst.“

### **Messinstrumente**

Thielsch und Moshagen (2010) entwickelten mit dem VisAWI (Visual Aesthetics of Websites Inventory) ein Messinstrument zur Erfassung visueller Ästhetik. Der VisAWI basiert auf dem Modell, dass Nutzer einen übergeordneten Generalfaktor Ästhetik wahrnehmen, der aus vier zugrundeliegenden Facetten besteht, Einfachheit ( $\alpha=.89$ ,  $I=5$ ), Vielfalt ( $\alpha=.87$ ,  $I=5$ ), Farbigkeit ( $\alpha=.90$ ,  $I=4$ ) und Kunstfertigkeit ( $\alpha=.85$ ,  $I=4$ ). In Klammern ist der Wert von Cronbachs Alpha ( $\alpha$ ) und die Anzahl der Items ( $I$ ) angegeben. Die Testpersonen beurteilen die Aussagen auf einer 7-stufigen Likert-Skala. Zusätzlich existiert eine Kurzversion (VisAWI-S) mit vier Items.

Lavie und Tractinsky (2004) identifizieren fünf Faktoren bei der wahrgenommen, visuellen Ästhetik von Websites. Dies sind: Klassische Ästhetik ( $\alpha=.85$ ,  $I=5$ ), Expressive Ästhetik ( $\alpha=.87$ ,  $I=5$ ), Usability ( $\alpha=.95$ ,  $I=5$ ), Angenehme Interaktion ( $\alpha=.86$ ,  $I=3$ ) und Service Qualität ( $\alpha=.86$ ,  $I=3$ ). In Klammern ist der Wert von Cronbachs Alpha ( $\alpha$ ) und die Anzahl der

Items (I) angegeben. Die Testpersonen beurteilen die Aussagen auf einem 7-stufigem Antwortformat (vgl. Kapitel 3.3, Tabelle 4).

### 2.2.2. Einfluss der Betrachtungsdauer auf wahrgenommene Ästhetik

Gemäss Thielsch (2008) finden sich mehrere Experimente zum Einfluss der Betrachtungsdauer einer Website auf das ästhetische Urteil (Fernandes et al. 2003; Lindgaard et al., 2006, Tractinsky et al., 2004; 2006, zitiert nach Thielsch, 2008). Die Autoren interpretieren die Ergebnisse dahingehend, dass ästhetische Urteile sehr schnell von Probanden gefällt werden und diese Urteile konstant bleiben. Bei diesen Studien ist vor allem die mangelnde Interaktion der Probanden mit den Websites kritisch zu hinterfragen. Wünschenswert wäre ein dritter Messzeitpunkt mit einer längeren Darbietung und Interaktion gewesen, der das reale Verhalten von Websurfern ausserhalb des Testlabors geboten hätte.

Die Experimente geben zwar erste Hinweise bezüglich der möglichen Ästhetik-Wahrnehmung. Die Versuchsaufbauten ähneln aber stark der aufeinanderfolgenden kurzen Darstellung von Bildern. Dabei könnte vermutet werden, dass sich die Probanden an sehr einfachen, holistischen Reizmustern orientieren und beispielsweise konsequent Screenshots wählen, die ihre jeweilige Lieblingsfarbe enthalten. Tractinsky et al. (2006, zitiert nach Thielsch, 2008) dagegen argumentieren, dass die Experimente nicht die Bedeutung weiterer Variablen ausschliessen, sondern die Bedeutung des Ersteindrucks betonen.

Tuch, Presslauer, Stöcklin, Opwis und Bargas-Avila (2012) geben in ihrer Arbeit eine Übersicht auf Studien, die nach Lindgaard, Fernandes, Dudek und Brown (2006) im Bereich der Betrachtungsdauer publiziert wurden. Sie schreiben, dass Benutzer in einer sehr kurzen Zeit verlässliche Urteile bezüglich Attraktivität und Vertrauen bilden und dass diese Urteile von Faktoren wie zum Beispiel Kontext oder visueller Komplexität abhängen. Weiter untersuchten sie die Rolle von visueller Komplexität und Prototypikalität als Design-Faktoren von Websites,

in dem der erste Eindruck anhand von Mittelwerten zweier Studien ( $N=59$ ) verglichen wurde. Als Grundlage diente das Modell von Leder et al. (2004).

Die Resultate zeigen, dass Visuelle Komplexität (VC) und Prototypikalität (PT) die ästhetische Wahrnehmung sogar innerhalb 17ms beeinflussen, wobei der Effekt von PT weniger ausgeprägt ist als von VC. Mit erhöhter Präsentationszeit wird der Effekt von PT gleich beeinflussend wie von VC. Diese Befunde unterstützen die Argumentation des kognitiven Verarbeitungsmodells nach Leder et al. (2004), in der visuelle Komplexität in einem früheren Stadium verarbeitet wird als Prototypikalität. Websites mit geringer visueller Komplexität und höherer Prototypikalität werden alles in allem äusserst attraktiv wahrgenommen.

Harrison, Reinecke und Chang (2015) führten eine Untersuchung mit 1278 Testpersonen durch, bei denen Informationsgrafiken nach einer Betrachtungsdauer von 500ms beurteilt wurden. Die Resultate zeigen drei Erkenntnisse: 1. Die Versuchspersonen bilden ein verlässliches Urteil bezüglich der Attraktivität der Infografik, basierend auf dem „mere-exposure-Effekt“. 2. Der erste Eindruck basiert hauptsächlich auf der Farbigkeit und der visuellen Komplexität. 3. Alter, Geschlecht und Bildungsniveau beeinflusst den bevorzugten Grad von Farbigkeit und Komplexität.

### 2.2.3. Personenvariablen

Gemäss Thielsch (2008) behandeln viele Studien zur Ästhetik von Websites Eigenschaften der Rezipienten nicht über die Darstellung der Demographie hinaus. In einzelnen Untersuchungen finden sich Prüfungen, inwieweit demographische Merkmale das ästhetische Urteil beeinflussen. Bei diesen Studien zeigt sich kein Einfluss demographischer (Parizotto-Ribeiro & Hammond, 2005; Moshagen et al., 2007; Zhang & von Dran, 2000, zitiert nach Thielsch, 2008) oder kultureller (Tractinsky, 1997, zitiert nach Thielsch, 2008) Merkmale auf das ästhetische Urteil. Ein Geschlechtseffekt bezüglich farblicher Gestaltung findet sich in einer Studie von

Mundorf, Westin und Dholakia (1993, zitiert nach Thielsch, 2008). Frauen bewerten hier ein sehr schlichtes Interface (ähnlich einem DOS-Interface) ohne Farbe hedonischer (vgl. Hassenzahl, 2001) als Männer. Innerhalb der Bedingung „Interface mit Farbe“ zeigt sich aber kein Unterschied.

Ästhetische Sensibilität als Eigenschaft des Betrachters formuliert bereits Eysenck (1940, zitiert nach Thielsch, 2008). Sie bezeichnet einen generellen Persönlichkeitsfaktor, welcher die Fähigkeit eines Individuums zu ästhetischer Wahrnehmung und ästhetischem Erleben beschreibt. In diesem Bereich existieren in der experimentellen Ästhetik bereits Ansätze, die ästhetische Sensitivität einer Person erfassen, zum Beispiel der VAST (visual aesthetic sensitivity test) von Götz, Borisy, Lynn und Eysenck (1979; Götz, 1985, zitiert nach Thielsch, 2008).

Bloch et al. (2003, zitiert nach Thielsch, 2008) konstruierten eine reliable und valide Skala zur Erfassung der ästhetischen Sensitivität CVPA (Scale for Centrality of Visual Product Aesthetics). Sie gehen von einer generellen, vom ästhetischen Objekt unabhängigen Persönlichkeitseigenschaft mit drei basalen Dimensionen aus: value, acumen und response.

Verschiedentlich findet sich ein Effekt hinsichtlich der Expertise im Bereich Webdesign (Haig, 2002; Haig & Whitfield, 2001; Hartmann et al., 2007, zitiert nach Thielsch, 2008). Karvonen (2000, zitiert nach Thielsch, 2008) regt an, dass Experten bei ihrer Bewertung möglicherweise auf andere Merkmale eines Systems achten. Ben-Bassat et al. (2006, zitiert nach Thielsch, 2008) und Moshagen (2007, zitiert nach Thielsch, 2008) finden keinen Expertise-Effekt.

Wang (2014) untersuchte die ästhetische Präferenz von Kindern bei Lernwebseiten aufgrund von Berlynes Theorie (1971, 1974, zitiert nach Wang, 2014). Dieser Ansatz besagt, dass Personen einen mittleren Grad an Reizen einem hohen oder tiefen Grad an Reizen vorziehen. 90 Kinder (45m,  $MW=13.6$ ,  $SD=0.54$ ) bewerteten 12 Websites in Bezug auf ihre ästhetische Vorliebe. Die Websites wurden gemäss ihres Aussehens in die Gruppen von tiefer, mittlerer und hoher visueller Komplexität eingeteilt. Die Resultate zeigen, dass Websites mit mittlerer visu-

eller Komplexität den Websites mit hoher oder tiefer visueller Komplexität vorgezogen werden, was Berlynes Theorie bestätigt. In Bezug auf das Geschlecht zeigt sich, dass Knaben hohe visuelle Komplexität und Mädchen mittlere oder tiefe visuelle Komplexität bevorzugen.

Loranger und Nielsen (2013) formulierten aus ihren Studien sechs Leitsätze, die bei der Gestaltung von Websites für die verschiedenen Altersgruppen berücksichtigt werden sollten. Ihre Untersuchungen basieren auf Stichproben von 38 respektive 46 Jugendlichen zwischen 13 und 17 Jahren. Dabei verwendeten Autorin und Autor folgende Methoden: Usability-Test, Feldstudien (Beobachtungen) sowie Interviews und Fokusgruppen.

Loranger und Nielsen (2013) fanden unter anderem heraus, dass Jugendliche (Teens) in Vergleich zu Erwachsenen bei einer Suchaufgabe weniger erfolgreich abschneiden. Sie führen es auf drei Gründe zurück: Ungenügende Fähigkeiten im Lesen, weniger ausgeklügelte Suchstrategie und dramatisch kleinere Ausprägung an Geduld.

Im Folgenden werden die erwähnten Gestaltungs-Leitsätze zusammengefasst (Loranger & Nielsen, 2013).

### **„Schreibe gut“**

Schreibe für ungeduldige User. Nichts schreckt die junge Zielgruppe mehr ab als überladene Bildschirme voll Text. Jugendliche können schnell gelangweilt, verwirrt und frustriert werden. Teens mögen es nicht, auf dem Web viel zu lesen. Die Lesefähigkeit von Teens, vor allem von jüngeren ist meistens nicht ideal. Stelle den Inhalt in kleinen, sinnvollen Einheiten mit genügend weissem Raum dar. Verwende Wörter, die die Teens verstehen. Schreibe in kleinen Sätzen und Abschnitten und mit einer Schriftgrösse 6 oder grösser. Jugendliche wie Erwachsene mögen keine kleinen Schriften.



### **„Vermeide langweiligen Inhalt – und eine Überlastung an Unterhaltung“**

Uninteressanter Inhalt ist tödlich, wenn Teens auf deiner Website behalten werden sollen. Aber nicht alles muss interaktiv und fancy sein. Obwohl Teens eine hohe Wertschätzung für Ästhetik haben, verabscheuen sie Sites, die überhäuft sind und sinnlose Multimedia enthalten. Multimedia kann Jugendliche abhängig von ihrer Nützlichkeit fesseln aber auch erzürnen. Die besten online-Erfahrungen machen Teens bei Inhalten, bei denen sie etwas Neues lernen oder bei denen sie auf ihrem Ziel fokussiert bleiben. Folgende interaktive Funktionen sind gut, weil sie Teens mehr tun lassen, als zu lesen und zu sitzen: Online-Quiz, Formen mit Feedback und Fragenstellen, Online-voting, Spiele, Sharing-Formen mit Bildern und Geschichten, Message boards, Foren mit angebotenen oder erhaltenen Anleitungen, Features für Website-Kreationen oder das Hinzufügen von Inhalt.

### **„Mach es schnell“**

Was immer du machst, vergewissere dich, dass deine Website schnell lädt. Langsame und schwerfällige Sites sind für jeden frustrierend, vor allem für junge offensive Zielgruppen, die sofortige Befriedigung erwarten. Überlege zweimal, ob du ein super cooles Widget entwickelst, da Teens häufig an älteren oder second-hand Computern mit teilweisen langsamen Internetverbindungen arbeiten. Jugendliche machen gerne Dinge auf dem Web. Sie missbilligen Sites, die zwar fancy aussehen, aber langsam und schwerfällig sind.

### **„Rede Teens nicht klein“**

Vermeide alles, was herablassend oder kindisch tönt. Jugendliche beziehen sich auf Inhalte, die von Jugendlichen gemacht wurden, also bereichere deinen Inhalt mit echten Geschichten, Bildern und Beispielen von andern Jugendlichen. Das Wort „Kinder“ stösst Jugendliche ab. Sie sind auf ihren neu gewonnenen Status stolz und wollen keinen kindischen Inhalt. Ein Grund mehr sich mit deftigen Animationen und auffälligen Farbschemen zurückzuhalten, die bei einem jüngeren Publikum funktionieren.

**„Lass Teens die sozialen Aspekte kontrollieren“**

Erleichtere den Austausch, aber forciere ihn nicht. Jugendliche bauen auf Technologien für soziale Kommunikation, aber sie wollen nicht die ganze Zeit „sozial“ sein. Sie möchten kontrollieren, was sie teilen und mit wem sie es teilen. Wenn du Sharing-Optionen anbietest, Sorge dafür, dass du E-Mail einschliesst. Im Gegensatz zu College-Studenten ziehen Jugendliche das E-Mail vor.

**„Designe für kleinere Bildschirme und limitierte Ergonomie“**

Viele Studenten surfen, in ungelungen Positionen sitzend, mit portablen Geräten und kleinen Bildschirmen (Laptops, Tablets und Smartphones) auf dem Netz. Diese Geräte erfordern, dass die Gestaltung der Website die Usability nicht beeinträchtigt. Obwohl die Bildschirme für Geschäftsnutzer immer grösser werden, erhalten Teens sehr selten diese High-End-Geräte. Jugendliche arbeiten oft mit Trackpads. Sie erschweren die Ausübung präziser Interaktionen, die bei drop-down Menüs, drag-n-drop und kleinen Buttons erforderlich sind. Weitere Design-Elemente wie kleine Klick-Zonen und roll over Effekte sind auch problematisch. Kleiner und dichter Text macht das Lesen schwierig.

In Tabelle 1 sind die hauptsächlichsten Ähnlichkeiten und Unterschiede verschiedener Altersgruppen in Bezug auf die Einstellung zum Webdesign zusammengefasst.

Tabelle 1: Web-Design für Altersgruppe, nach Loranger & Nielsen (2013, S. 5)

	Jagd auf Dinge zum klicken	Browsen mit Registerkarten	Scrollen	Suche	Geduld	Animationen und Sound Effekte
Kinder (3-12)	☺	☹	☹	☹	☹	☺
Teens (3-17)	☺	☺	☺	☹	☹	☺
College Studenten (18-24)	☹	☺	☺	☺	☹	☹
Erwachsene (25-64)	☹	☺	☺	☺	☺	☹

Anmerkungen. Originalsprache Englisch.

☺=Unterhaltsam, interessant, und ansprechend, oder User stellen sich leicht darauf ein.

☹=User schätzen es bis zu einem gewissen Grad, aber die Übertreibung kann problematisch sein.

☹=Users mögen es nicht, oder finden es schwierig zu handhaben. Mach es nicht.

#### 2.2.4. Usability und User Experience

##### Usability

Die Industrienorm zur Ergonomie der Mensch-System-Interaktion-Teil 210 definiert Usability als das: „Ausmass, in dem ein System ein Produkt oder eine Dienstleistung durch bestimmte Benutzer in einem bestimmten Nutzungskontext genutzt werden kann, um festgelegte Ziele, effektiv, effizient und zufriedenstellend zu erreichen“ (Deutsches Institut für Industrienorm e. V. EN ISO 9241-210 [DIN EN ISO 9241-210], 2010, S. 7).

Die Effektivität bezeichnet die „Genauigkeit und Vollständigkeit“ mit der Benutzer ein bestimmtes Ziel erreichen. „Effizienz wird mit „im Verhältnis zu Genauigkeit und Vollständigkeit eingesetzter Aufwand, mit dem Benutzer ein bestimmtes Ziel erreichen“ beschrieben. Die Norm definiert Zufriedenheit als „Freiheit von Beeinträchtigungen und positiven Einstellungen gegenüber der Nutzung des Produkts.“

Nach Thielsch (2008) soll den Benutzern eine positive Einstellung gegenüber der Nutzung des Systems sowie Beeinträchtigungen durch das System vermieden werden. Entsprechend zeichnet sich eine Website mit guter Usability beispielsweise durch eine einfache Navigation, übersichtliche Struktur und ihrem Zweck optimal angepasste Gestaltung aus.

Die DIN EN ISO 9241-110 (2008) beschreibt sieben Grundsätze zur Dialoggestaltung. Mithilfe dieser Kriterien lässt sich die Usability von interaktiven Systemen bewerten. Das sind: Aufgabenangemessenheit, Selbstbeschreibungsfähigkeit, Erwartungskonformität, Steuerbarkeit, Fehlertoleranz, Individualisierbarkeit und Lernförderlichkeit. Usability ist nach Thielsch (2008) demnach stark abhängig von der Interaktion zwischen Nutzer und System. Die DIN EN ISO 9241-115 (2008) geht näher auf die spezifischen Anforderungen im Rahmen der Gestaltung von Websites ein. Mit einem allgemeinen Referenzmodell, das Konzeption, Gestaltung und Evaluation von Websites beinhaltet, werden Leitlinien für den Schaffungsprozess einer Website insbesondere auf die Erstellung, Strukturierung und Darstellung von Inhalten vorgegeben.

Gemäss Thielsch (2008) sind Forschungen zu Usability sehr umfangreich (Barnum, 2002; Shneiderman & Plaisant, 2005, zitiert nach Thielsch, 2008). Dabei folgen nicht alle dem Forscher den Begrifflichkeiten und Definitionen der ISO-Norm. In den verschiedenen Definitionen werden aber meistens die Aspekte der Zielerreichung, der Fehlerfreiheit und Passung zwischen Nutzer und Aufgabe aufgegriffen.

### **User Experience (UX)**

Die DIN EN ISO 9241-210 (2010, S. 7) definiert UX als „Wahrnehmungen und Reaktionen einer Person, die aus tatsächlichen und/oder der erwarteten Benutzung eines Produkts, eines Systems oder einer Dienstleistung resultieren“. Dazu wird angemerkt, dass UX „sämtliche Emotionen, Vorstellungen, Vorlieben, Wahrnehmungen, physiologischen Reaktionen, Verhaltensweisen und Leistungen, die sich vor, während und nach der Nutzung ergeben“ umfasst.

Hassenzahl und Tractinsky (2006, S. 95) betonen die emotional affektiven Aspekte stärker.

Gemäss ihnen ist UX:

„a consequence of a user's internal state (predispositions, expectations, needs, motivation, mood etc.), the characteristics of the designed system (eg. complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g organisational/social setting, meaningfulness of the activity, voluntariness of use, etc.)“

Eine Übersicht über gemessene UX-Konstrukte geben Bargas-Avila und Hornbaek (2011) in ihrer Review, in der sie 51 Publikationen überprüften. Dabei stiessen sie auf folgende Konstrukte: Generic User Experience, Affect/Emotion, Enjoyment/Fun, Aesthetics/Appeal, Hedonic Quality, Engagement/Flow, Motivation, Enchantment, Frustration, Other constructs.

Die Ergebnisse zeigen auf, dass die oft genannten Schlüsselfaktoren wie Benutzungskontext und Haltung selten erforscht wurden. Am häufigsten untersucht wurden die Dimensionen Emotionen, Freude und Ästhetik.

In weiteren Studien im Bereich Website-Ästhetik und User Experience wurden unter anderem Farbe und Komplexität untersucht, wie zum Beispiel von Bonnardel und Bigot (2011) sowie Tuch, Bargas-Avila, Opwis und Wilhelm (2009). Ferner erforschten Tuch, Roth, Hornbaek, Opwis und Bargas-Avila (2012) die Beziehung zwischen Usability, Ästhetik und Affekt in Mensch-Computer-Interaktionen.

### **Messinstrumente**

Zur Erfassung von User Experience entwickelte Sauro (2015) den Standardized User Experience Percentile Rank Questionnaire (SUPR-Q). Er besteht aus insgesamt acht Items, die auf die vier Dimensionen verteilt sind: Aussehen ( $\alpha=.78$ ,  $I=2$ ), Loyalität ( $\alpha=.64$ ,  $I=2$ ), Usability ( $\alpha=.88$ ,  $I=2$ ), Vertrauen ( $\alpha=.85$ ,  $I=2$ ) verteilt sind. In Klammern ist der Wert von Cronbachs

Alpha ( $\alpha$ ) und die Anzahl der Items (I) angegeben. Die Testpersonen beurteilen die Aussagen auf einer 5-stufigen Skala, ausser Item „How likely are you recommend this webiste to a friend or colleague?“ wird auf einer 11-stufigen Skala beantwortet (vgl. Kapitel 3.3, Tabelle 4).

Minge und Riedel (2013) konzipierten auf der Basis des CUE-Modells von Thüring und Malke (2007, zitiert nach Minge & Riedel, 2013) den modularen Fragebogen zur Erfassung des Nutzungserlebens (meCUE). Der Fragebogen setzt sich aus drei Modulen „Produktwahrnehmung“, „Nutzeremotionen“ und „Konsequenzen“ zusammen. „Produktwahrnehmung“ umfasst Nützlichkeit ( $\alpha=.83$ , I=3), Benutzbarkeit ( $\alpha=.89$ , I=3), visuelle Ästhetik ( $\alpha=.89$ , I=3), soziale Identität: Status ( $\alpha=.83$ , I=3), soziale Identität: Bindung ( $\alpha=.86$ , I=3). „Nutzeremotion“ umfasst positive Emotionen ( $\alpha=.94$ , I=6) und negative Emotionen ( $\alpha=.92$ , I=6). „Konsequenzen“ umfasst Produktloyalität ( $\alpha=.86$ , I=3) und Nutzungsintention ( $\alpha=.76$ , I=3). In Klammern ist der Wert von Cronbachs Alpha ( $\alpha$ ) und die Anzahl der Items (I) angegeben. Das Antwortformat ist 7-stufig (vgl. Kapitel 3.3, Tabelle 4).

O’Brian und Toms (2010) erstellten einen Fragebogen zur Erfassung von „User Engagement“. Das Messinstrument basiert auf den fünf Faktoren: fokussierte Aufmerksamkeit ( $\alpha=.92$ , I=7), wahrgenommene Usability ( $\alpha=.91$ , I=8), Ästhetik ( $\alpha=.89$ , I=5), Einträglichkeit ( $\alpha=.87$ , I=5), Neuheit ( $\alpha=.59$ , I=3) und Einbindung ( $\alpha=.70$ , I=3). In Klammern ist der Wert von Cronbachs Alpha ( $\alpha$ ) und die Anzahl der Items (I) angegeben. Die Testpersonen bringen ihre Einschätzung auf einem 5-stufigem Antwortformat zum Ausdruck (vgl. Kapitel 3.3, Tabelle 4).

Hassenzahl, Burmester und Koller (2003) entwickelten den AttrakDiff zur Erfassung des wahrgenommenen Produktcharakters sowie der globalen Bewertung interaktiver Produkte. Mit Hilfe von Wortpaaren (Semantisches Differential) wird erhoben, wie Nutzer ein Produkt wahrnehmen und bewerten. Der AttrakDiff erfasst die Dimensionen Attraktivität, Pragmatische Qualität, Hedonische Qualität-Identität und Hedonische Qualität-Stimulation. Die Skala umfasst in jeder Dimension sieben gegensätzliche Adjektivpaare, zwischen denen die Testperson

ein Produkt einschätzt. Innerhalb der Wortpaare gibt es sieben Abstufungen. Die Autoren geben für Cronbachs Alpha folgende Werte an: Pragmatische Qualität ( $\alpha=.83-.85$ ), Hedonische Qualität-Stimulation ( $\alpha=.76-.90$ ) und Hedonische Qualität-Identität ( $\alpha=.73-.83$ ).

Ein umfangreiches, praxisorientiertes Handbuch zur Evaluation von User Experience stellen Diefenbach und Hassenzahl (2010) zur Verfügung.

Sauro (2015) erstellte in seiner Arbeit ab Seite 70 eine Übersicht über diverse Messinstrumente für Software und Website Qualität. Für weitere Messinstrumente sei auf diese Publikation verwiesen.

#### 2.2.5. Zentrale Variablen der Website-Rezeption

Aufgrund seiner qualitativen und quantitativen Studien benennt Thielsch (2008) in seinem Modell (Abbildung 2) zentrale Variablen, die bei der Website-Rezeption von Bedeutung sind. Dabei wird Farbe als wichtige ästhetisch wirkende Variable hervorgehoben.

Der Inhalt hat sich gemäss Thielsch (2008) zwar relativ unbeeinflusst von ästhetischer Gestaltung gezeigt, ist aber bezüglich der Analyse von Aspekten der Passung zwischen den beiden Konstrukten ein weiterer Forschungsansatz. Besonders die Wirkung von Inhaltsvariationen auf ästhetische Wahrnehmungen ist hierbei interessant. Überraschend sind die hohen Korrelationen zwischen den Bewertungen von Ästhetik auf der einen sowie Usability und Inhalt auf der anderen Seite, obwohl sich kein Einfluss der vorgenommenen ästhetischen Gestaltungsvariationen auf die letzteren beiden finden lässt. Entsprechend ist auch die allgemeine Betrachtung der Wirkungszusammenhänge zwischen den drei Konstrukten Ästhetik, Usability und Inhalt ein weiterer Ansatzpunkt für Forschungen, zusätzlich auch der Einfluss von weiteren technischen oder personenbezogenen Variablen auf diese.

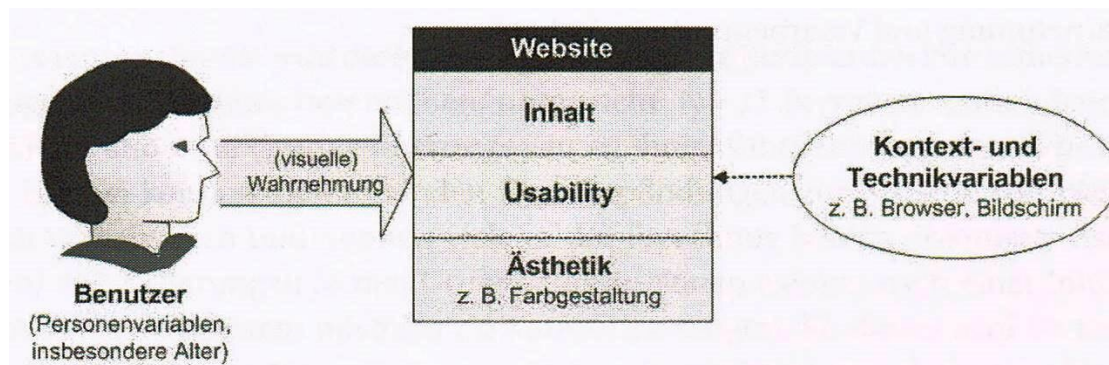


Abbildung 2: Zentrale Variablen bei der Website-Rezeption (Thielsch, 2008, S. 263)

### 2.3. Lernkontext

Als grundlegendes Modell im Bereich Multimedia-Lernen können die Basis-Prinzipien von Mayer (2014b) beim Design von multimedialen Lernumgebungen angesehen werden. Sie werden in der Tabelle 2 zusammenfassend dargestellt. In dieser Arbeit sind sie nur am Rande relevant, deshalb wird nicht näher darauf eingegangen. Dagegen werden Motivation und Emotionen im Zusammenhang mit Multimedia-Lernen näher erläutert. Im Zusammenhang mit Motivation und Multimedia-Lernen schreibt Mayer (2014, S. 171) in seinem Artikel: „Overall, motivational features can improve student learning by fostering generative processing as long as the learner is not continually overloaded with extraneous processing or overly distracted from essential processing.“



Tabelle 2: Mayers Multimedia-Prinzipien, eigene Darstellung nach Mayer (2014b)

<b>Mayers Prinzip</b>	<b>Beschreibung</b>
Multimediaprinzip	Menschen lernen besser mit Text und Bild als ausschliesslich mit Text.
Redundanzprinzip	Menschen lernen besser wenn die gleiche Information nicht mehr als einem Format gezeigt wird.
Modalitätsprinzip	Menschen lernen besser mit Grafiken und Erzählungen als mit Grafiken und gedrucktem Text.
Signalisierungsprinzip	Menschen lernen besser, wenn Hinweisreize hinzugefügt werden, die Schlüsselinformationen und deren Organisation hervorheben.
Räumliches Kontiguitätsprinzip	Menschen lernen besser, wenn zusammengehörende Bild- und Textinformationen räumlich nah zueinander präsentiert werden.
Zeitliches Kontiguitätsprinzip	Menschen lernen besser, wenn zusammengehörende Bild- und Textinformationen gleichzeitig präsentiert werden. Im Gegensatz zu sukzessiv aufeinanderfolgenden Präsentation.

Leutner (2014) diskutiert in seiner Arbeit drei Studien im Vergleich zu Morenos „cognitive-affective theory of learning with media“ (CATLM) (Moreno, 2006, zitiert nach Leutner, 2014). Dazu schreibt er: Im Einklang mit Morenos CATLM berücksichtigen alle Autoren Mediations-Hypothesen. Experimentell induzierte Unterschiede in Motivation/Emotion führen zu Unterschieden in der kognitiven Verarbeitung von Lernmaterial und führen zu Unterschieden in den Lernresultaten. Leider wurde die kritische Beziehung zwischen Motivation/Emotion und der kognitiven Verarbeitung nicht direkt untersucht, wobei dies an der mangelnden Verfügbarkeit von Messinstrumenten für die kognitive Verarbeitung liegen könnte. Weitere Forschungen sollten unabhängige Messinstrumente für die kognitive Verarbeitung entwickeln und anwenden, damit diese Mediations-Hypothese getestet werden kann.

### 2.3.1. Emotion

Nach Harrer (2012) besteht in der Emotionsforschung Konsens darüber, dass Emotionen Prozesse darstellen, die sich aus affektiven, kognitiven, expressiven, motivationalen und physiologischen Komponenten zusammensetzen (Scherer, 1993, zitiert nach Götz et al., 2004;vgl.

Pekrun & Hofmann, 1999; Hagenauer, 2011, zitiert nach Harrer, 2012). Anhand dieser Komponentendefinition der Emotionen wird die Relevanz von Emotionen für den Lern- und Leistungskontext sichtbar. Dabei ist vor allem die motivationale Komponente, die sich direkt auf die Handlungstendenzen bezieht, von besonderer Bedeutung für das Lernen und Erbringen von Leistung.

Unter dem Begriff Leistungsemotion (engl. achievement emotion) sind gemäss Harrer (2012) solche Emotionen zu verstehen, die im Zusammenhang mit Lern- und Leistungssituationen erscheinen und sich direkt auf die Aktivität (engl. activity-related achievement emotions) oder das Resultat einer Aktivität (engl. outcome-related achievement emotions) beziehen (Pekrun, 2000; Pekrun, 2006, zitiert nach Harrer, 2012). In diesem Zusammenhang handelt es sich um Emotionen, die zum Beispiel im Unterricht, während Prüfungen, beim Lernen zuhause oder anderen lern- und leistungsassoziierten Situationen oder Tätigkeiten erlebt werden (vgl. Götz et al., 2004; Götz, Zirngibl, Pekrun, & Hall, 2003, zitiert nach Harrer, 2012). Obwohl der Begriff Leistungsemotion in der deutschsprachigen Emotionsforschung fälschlicherweise vermuten lässt, dass es sich hierbei ausschließlich um Emotionen handelt, die sich auf Leistung respektive das Resultat einer leistungsassoziierten Tätigkeit beziehen, sind unter diesem Begriff auch Lernemotionen zu subsumieren. Lernemotionen beziehen sich im Gegensatz zu Leistungsemotionen nicht auf das Resultat, sondern direkt auf die Aktivität des Lernens wie zum Beispiel Freude darüber, beim Lernen gut voran zu kommen oder Frustration aufgrund des schwierigen Lernstoffes.

Die Kontroll-Wert Theorie ist gemäss Harrer (2012) ein integratives Modell, das verschiedene Konzepte der Motivationspsychologie und Emotionsforschung vereint (vgl. Pekrun, 2006, zitiert nach Harrer, 2012). Dazu zählen neben Erwartungs-Wert Ansätzen der Emotionen (z.B. Turner & Schallert, 2001, Pekrun 2006, zitiert nach Harrer, 2012) und Kontrollwahrnehmungstheorien (z.B. Patrick, Skinner, & Connell, 1993 Pekrun, 2006, zitiert nach Harrer, 2012) auch

Attributionstheorien von Lern- und Leistungsempfindungen (Weiner, 1985, zitiert nach Harrer).  
Abbildung 3 zeigt die Variablen der Kontrollwert-Theorie (Pekrun, 2006, zitiert nach Harrer, 2012).

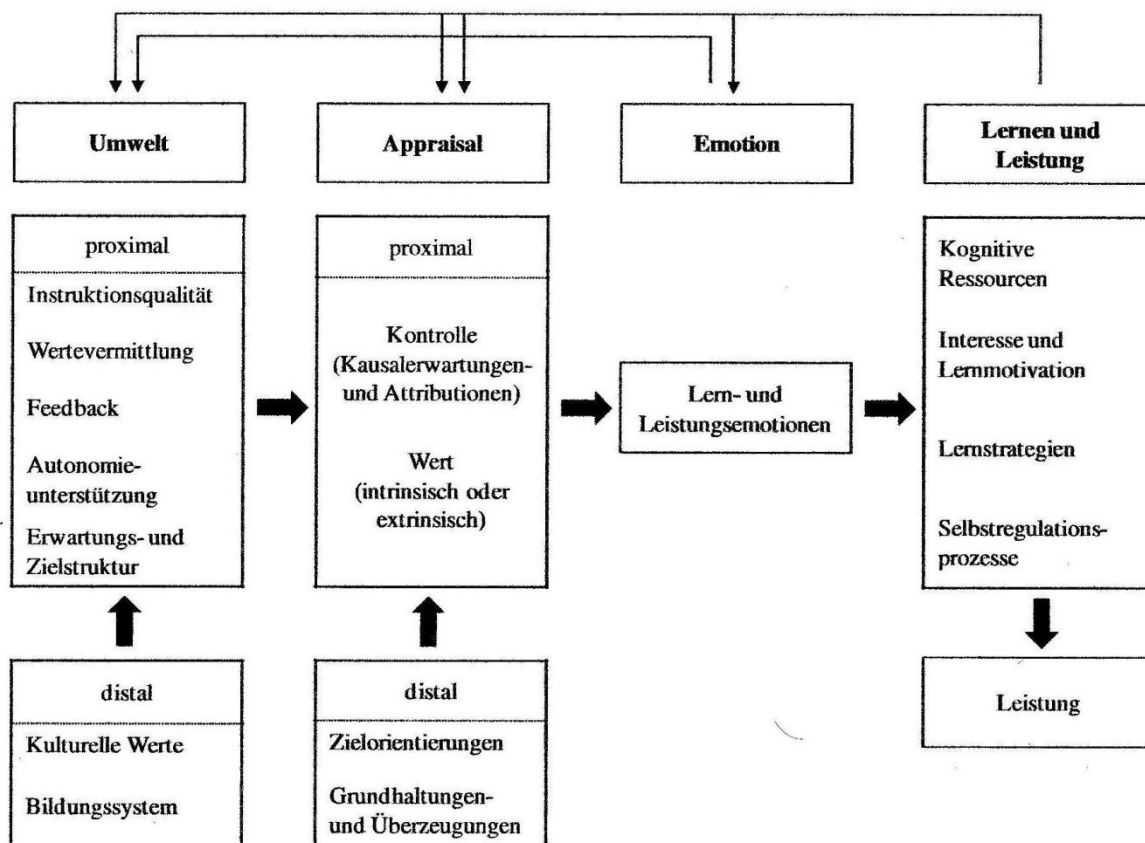


Abbildung 3: Kontroll-Wert Theorie, nach Pekrun (2006, adaptiert von Harrer 2012, S. 23)

### Messinstrumente

Pekrun, Goetz, Titz und Perry (2002) entwickelten auf der Basis der Kontroll-Wert-Theorie (vgl. Abbildung 3) einen Fragebogen zu Lern- und Leistungsempfindungen. Dazu sind umfangreiche Skalen in Bezug auf Klassen-, Lern- und Testbezogene Emotionen erhältlich (Pekrun, Goetz, & Perry P., 2005). Die Skala für „Lernbezogene Langeweile“ umfasst 11 Items. Cronbachs Alpha wird mit ( $\alpha=.92$ ) angegeben. Das Antwortformat ist 5-stufig (vgl. Kapitel 3.3, Tabelle 4).

Mit einer deutschen Version des PANAS (Positive and negative Affect Schedule) können positive und negative Affekte gemessen werden (Krohne, Egloff, Kohlmann & Tausch, 1996). Dabei schätzen die Versuchspersonen ihren emotionalen Zustand anhand von 20 Adjektiven ein, die Emotionen charakterisieren. Jeweils die Hälfte der Adjektive beschreibt positive, bzw. negative Emotionen. Das Antwortformat ist 5-stufig. Cronbachs Alpha beträgt für die Skala der positiven Emotionen  $\alpha=.85$ , für die Skala der negativen Emotionen  $\alpha=.85$  (vgl. Kapitel 3.3, Tabelle 4).

### 2.3.2. Lernmotivation

Gemäss Spinath (2011) kommt der Lernmotivation in der Bildungspraxis und -forschung eine wichtige Rolle zu. Lernmotivation ist einerseits eine wichtige Voraussetzung für Lernen und Leistung andererseits auch für Wohlbefinden und Selbstbestimmtheit. Ausserdem ist eine hohe Lernmotivation auch ein Bildungsziel in sich, da lebenslanges Lernen eine lebenslange Lernmotivation erfordert. Da Lernmotivation durch Merkmale der Lernumwelt manipulierbar ist, bildet sie einen zentralen Ansatzpunkt für pädagogisch-psychologische Interventionen. In Tabelle 3 sind zentrale Konstrukte und Theorien der Lernmotivation zusammengefasst (Spinath, 2011).

**Tabelle 3: Lernmotivation: Zentrale Konstrukte und Theorien, nach Spinath (2011, S. 50)**

<b>Konstrukt</b>	<b>Theorie</b>	<b>Vertreter/innen und grundlegende Publikationen</b>
Intrinsische/Extrinsische Motivation	Selbstbestimmungstheorie	Deci & Ryan (1985) Ryan & Deci (2000)
Interesse	Interesstheorie	U. Schiefele (1996) Krapp (1999)
Erwartungen und Werte	Erwartungs-Wert-Theorie	Atkinson (1957) Eccles & Wigfield (2002)
Zielorientierung	Zielorientierungstheorie	Nicholls (1984) Elliot (1999)
Leistungsmotiv	Leistungsmotivtheorie	Mc Clelland et al. (1953) Heckhausen (1972)

Im Folgenden wird das Konstrukt „intrinsische/extrinsische Motivation“ erläutert. Auf die nähere Beschreibung der anderen Theorien wird verzichtet, da sie im konstruierten Fragebogen keine Verwendung finden.

Um zu beschreiben, warum eine Person Lernaktivitäten unternimmt, wird nach Spinath (2011) oftmals zwischen intrinsischer und extrinsischer Motivation unterschieden. Unter intrinsischer Motivation wird verstanden, dass die Gründe für die Auseinandersetzung mit der Aufgabe in der Aufgabe selbst liegen und nicht in etwaigen Konsequenzen (Deci & Ryan, 1985; Ryan & Deci, 2000, zitiert nach Spinath, 2011). Jemand ist intrinsisch motiviert, wenn die Ausübung der Tätigkeit selbst als belohnend empfunden wird. Intrinsische Motivation gilt deshalb als besonders wünschenswerte Art der Lernmotivation, da sich die Lernenden als selbstbestimmt erleben können und Freude beim Lernen empfinden. Gemäss der Selbstbestimmungstheorie von Deci und Ryan (1985, zitiert nach Spinath, 2011) ist das Auftreten von intrinsischer Motivation an die Befriedigung der Grundbedürfnisse nach Kompetenzerleben und Selbstbestimmung geknüpft.

Im Gegensatz dazu wird von extrinsischer Motivation gesprochen, wenn die Gründe für die Beschäftigung mit einer Aufgabe eher in den Konsequenzen liegen (Deci & Ryan, 1985; Ryan & Deci, 2000, zitiert nach Spinath, 2011). Beispielsweise könnten sich Schülerinnen oder Schüler nur deshalb mit ihren Mathematikaufgaben befassen, weil sie gute Noten haben oder sich vor anderen nicht blamieren möchten. Extrinsische Motivation wird generell als wenig wünschenswerte Motivation angesehen, weil die Beschäftigung mit der Aufgabe nur Mittel zum Zweck ist und bei Wegfall der vorweg genommenen Konsequenzen auch keine Beschäftigung mit der Sache mehr erfolgen wird. In der Realität sind intrinsische und extrinsische Motivation natürlich nicht in ihrer Reinform anzutreffen. Anstelle sind von Ryan und Deci (2000, zitiert nach Spinath) verschiedene Zwischenformen beschrieben worden, die auf einem Kontinuum zwischen intrinsischer und extrinsischer Motivation liegen.

## Messinstrumente

Isen & Reeve (2005) konstruierten einen Fragebogen zu Intrinsischer Motivation mit acht Items. Cronbachs Alpha beträgt  $\alpha=.92$ . Die Testpersonen beurteilen die Aussagen auf einer 7-stufigen Skala (vgl. Kapitel 3.3, Tabelle 4).

### 2.3.3. Emotional Design

Im Folgenden werden drei Studien zu emotionalem Design zusammenfassend dargestellt. Die Autorinnen und Autoren beziehen sich in ihren Untersuchungen unter anderem auf die Kontroll-Wert-Theorie von Pekrun (2006) und auf die Arbeiten zum Multimedia-Lernen von Mayer (2009) sowie Moreno und Mayer (2007).

#### **Studie von Heidig, Müller und Reichelt (2015)**

**Titel:** Differenzierung zwischen relevanten Design-Merkmalen und ihren Effekt auf Emotionen und Lernen.

**Stichprobe:** Deutsche Studenten ( $N=334$ ,  $w=273$ , Alter:  $MW=22.3$ ,  $SD=2.1$ )

##### Resultate

Objektive Unterschiede in Ästhetik und Usability beeinflusst den emotionalen Zustand von Lernenden nicht.

Wahrgenommene Ästhetik und Usability beeinflusst den emotionalen Zustand von Lernenden positiv.

Der emotionale Zustand von Lernenden hat einen kleineren Einfluss auf Lernresultate aber einen grösseren Einfluss auf die intrinsische Motivation, einschliesslich die Motivation, mit dem Lernmaterial weiter zu arbeiten.

#### **Studie von Um, Plass, Hayward und Homer (2002)**

**Titel:** Emotionales Design beim Multimedia-Lernen

**Stichprobe:** Amerikanische Studenten ( $N=118$ ,  $w=79$ , Alter:  $MW=24.9$ ,  $SD=6.4$ )

##### Resultate

Die Verwendung von emotionalen Design-Prinzipien bei Lernmaterial kann zu positiven Emotionen führen und positive Emotionen im Multimedia-Lernen können kognitive Verar-

beitungsprozesse und lernen vereinfachen.

Positive Emotionen durch „mood induction“ erhöhen die berichtete mentale Anstrengung signifikant, wobei positives emotionales Design die wahrgenommene Schwierigkeit der Lernaufgabe reduziert.

Positive Emotionen erhöhen Motivation, Befriedigung und Wahrnehmungen bezüglich Materials.

Mediationsanalysen zeigen, dass der external herbeigeführte Effekt auf positive Emotionen von Motivation und mentaler Anstrengung mediiert wird. Aber es finden sich keine Mediatoren für Emotionen, die via emotionales Design herbeigeführt werden.

Es zeigt sich, dass positives emotionales Design hat einen direkteren Einfluss auf das Lernen als external herbeigeführte Emotionen.

### **Studie von Plass, Heidig, Hayward, Homer und Um (2014)**

Titel: Effekte von Form und Farbe auf Emotion und Lernen.

Stichprobe: Deutsche Studenten ( $N=121$ ,  $w=101$ , Alter:  $MW=22.0$ ,  $SD=2.1$ )

#### Resultate

Studie 1 zeigt, dass gut designtes Material zu positiven Emotionen führt und das Verständnis erleichtert, wobei die Transferleistung nicht vom emotionalen Design beeinflusst wird.

Studie 2 zeigt, dass runde, gesichtsähnliche Formen alleine und im Zusammenhang mit warmen Farben zu positiven Emotionen führen.

Warme Farben alleine dagegen beeinflussen die Emotionen der Lernenden nicht. Das Verständnis wird durch warme Farben und runde, gesichtsähnliche Formen erleichtert, zusammen aber auch unabhängig voneinander. Der Transfer wird durch runde, gesichtsähnliche Formen erleichtert, wenn neutrale Farben benützt werden.

## 2.4. Arbeitsmodell

Aus den vorangehend erläuterten Konzepten erfolgt die Bildung eines Arbeitsmodells zur Definition und Erfassung von „ästhetischem Empfinden“. Dabei sind die besonderen Rahmenbedingungen der Fragestellung wie etwa Lernkontext, Zielgruppe und Applikation berücksichtigt. Ziel dieser Untersuchung ist nicht etwa die empirische Überprüfung dieses Arbeitsmodells. Das Arbeitsmodell dient als Grundlage, um Variablen respektive Konstrukte zu identifizieren, die bei der ästhetischen Empfindung und deren Erfassung von Bedeutung sind. Ausserdem begründet es die Auswahl der relevanten Dimensionen, Skalen und Items der bereits vorhandenen Messinstrumente, die beim neu konzipierten Fragebogen zur Erfassung des ästhetischen Empfindens verwendet werden.

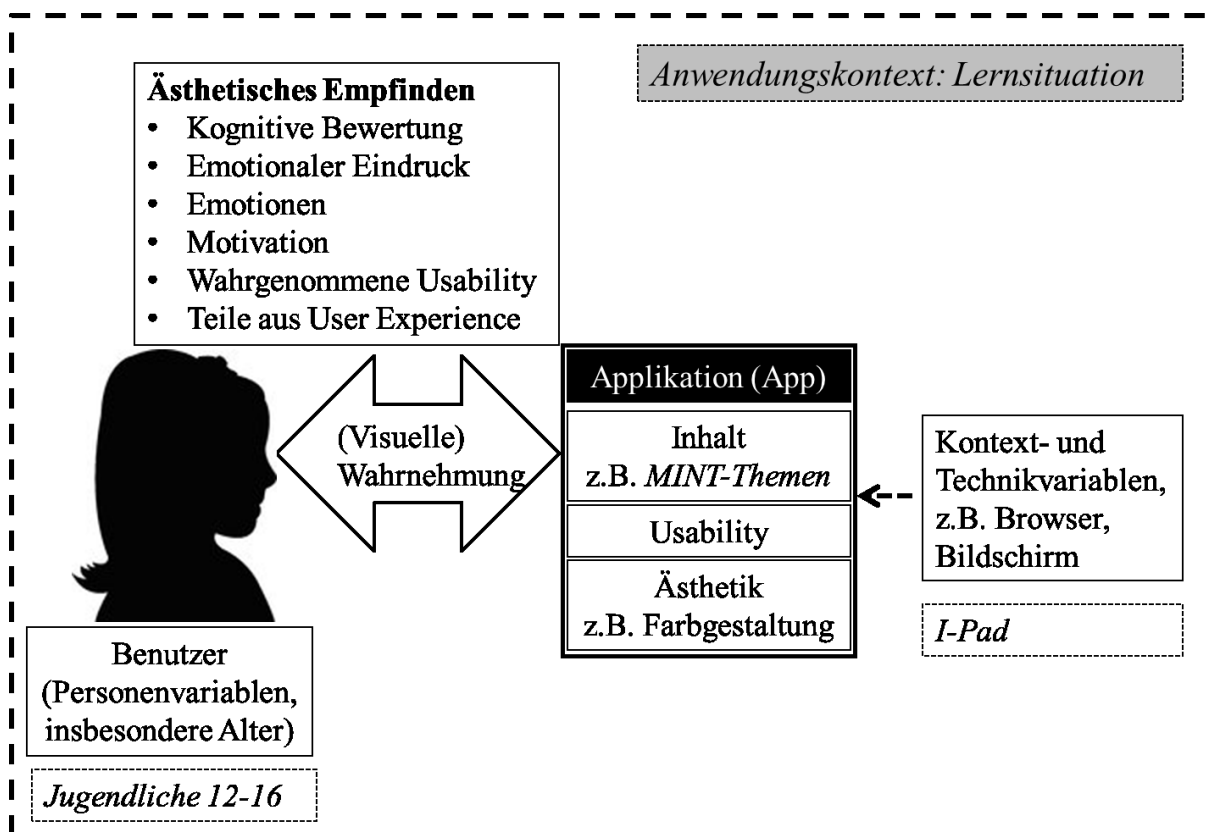


Abbildung 4: Arbeitsmodell ästhetisches Empfinden

Grundlage des Arbeitsmodells (Abbildung 4) bildet das Modell von Thielsch (2008), das zentrale Variablen bei der Website-Rezeption definiert. Das Arbeitsmodell liegt in einem definier-



ten Anwendungskontext, nämlich einer Lernsituation. Der Untersuchungsgegenstand ist im Gegensatz zu Thielsch (2008) keine Website, sondern eine Applikation, bei deren Gestaltung Variablen wie Inhalt, Usability und Ästhetik zu berücksichtigen sind. Vor allem die Konzeption des Inhalts kann durch weitere Variablen aufgelöst werden, da bei dessen Gestaltung Elemente des Multimedia-Lernens (Mayer, 2009) oder des Emotional Designs (Um et al., 2012) und anderen einfließen können. Weiter wird der Inhalt durch Themen aus den MINT-Fächern eingegrenzt.

Die Technikvariable wird durch die Anwendung eines I-Pads beschrieben. Bildschirmgröße und Steuerung können Auswirkungen auf die Usability haben, wie beispielsweise Loranger und Nielsen (2013) darstellen. Bei den Personenvariablen ist zu berücksichtigen, dass es sich um Jugendliche handelt, die in der Schweiz zur Schule gehen.

In Anlehnung an Leder et al. (2004) ist der Pfeil bei der (visuellen) Wahrnehmung doppelseitig gezeichnet. Dies symbolisiert, dass bei der Rezeption top-down und bottom-up-Prozesse stattfinden.

Das Konstrukt „ästhetische Empfinden“ stellt sich zwischen Subjekt (Benutzer) und Objekt (App) ein. Hauptbestandteile sind eine „kognitive Bewertung“ und ein „emotionaler Eindruck“ nach Thielsch (2008), die in Analogie zum „ästhetischen Urteil“ aufgrund kognitiver Verarbeitungsprozesse und zur „ästhetischen Emotion“ nach Leder et al. (2004) gesehen werden können.

Zusätzlich werden weitere Konstrukte wie Emotionen, Motivation, Wahrgenommene Usability und Teile aus User Experience integriert. „Emotionen“ und „Motivation“ werden als Variablen im Zusammenhang mit dem Lernkontext (vgl. Kapitel 2.3) verwendet. „Wahrgenommene Usability“ unterstreicht den Umstand, dass es sich beim „ästhetischen Empfinden“ um eine subjektive Einschätzung handelt. Die Variable „Teile aus User Experience“ beinhaltet eine

Auswahl aus Konstrukten, die aus dem Bereich „User Experience“ (vgl. Kapitel 2.2.4) stammen wie zum Beispiel „Focused Attention“ und „Involvement“ von O’Brian und Toms (2010). User Experience als Ganzes wäre ein viel zu breit gefasstes Konstrukt und ist teilweise unscharf abgegrenzt. Die Dimension Ästhetik und Emotionen werden zudem häufig als Teilbereiche in User Experience angesehen (vgl. Kapitel 2.2.4).

Zusammengefasst ist ästhetisches Empfinden demnach ein Konstrukt, das sich aufgrund von Wahrnehmungs- und Bewertungsprozessen zwischen den Gestaltungsmerkmalen der App (Inhalt, Usability und Ästhetik) und der Betrachterin oder dem Betrachter einstellt. Die Wahrnehmungs- und Bewertungsprozesse beinhalten kognitive, emotionale, motivationale und andere Komponenten.

### 3. Methoden

In diesem Kapitel werden die methodischen Grundlagen sowie die Durchführung der einzelnen Untersuchungsschritte wie Fragebogenkonstruktion, Testkonstruktion, Testdurchführung und Auswertung aufgezeigt. Zusätzlich erfolgt bereits in diesem Kapitel eine Diskussion der Gütekriterien Durchführungs- und Auswertungsobjektivität.

#### 3.1. Untersuchungs-Design

Zur Beantwortung der Fragestellung wurde ein quantitatives Untersuchungs-Design im Sinne einer klassischen Fragebogenstudie gewählt. Abbildung 5 zeigt die verschiedenen Phasen.



Abbildung 5: Untersuchungs-Design

Die einzelnen Phasen werden in den folgenden Abschnitten ausführlich beschrieben. Auf das Literarurstudium wird nicht explizit eingegangen. Es war jedoch ein integraler Bestandteil jeder Untersuchungs-Phase. Die quantitative Untersuchung mit einer Stichprobe von 160 Testpersonen ermöglichte eine erste Validierung des neu konstruierten Fragebogens. Aus den statistischen Daten konnten Erkenntnisse gewonnen werden, die zur einer Bestätigung oder Anpassung der ausgewählten Items und Dimensionen führten.

### 3.2. Methodische Grundlagen zur Fragebogenkonstruktion

Folgende Abschnitte geben eine Übersicht, welche Grundlagen und Kriterien bei der Konstruktion eines Fragebogens beachtet werden müssen.

#### 3.2.1. Itemformulierung

Nachfolgend sind die wichtigsten Regeln in Bezug auf die Formulierung von Items zusammengefasst. Punkt 1 stammt aus Bühner (2011), Punkt 4, 5 und 7 aus Jonkisz, Moosbrugger und Brandt (2012) und Punkt 2, 3, 6 und 8 aus Porst (2011).

1. Einfache und eindeutige Begriffe verwenden.
2. Pro Item nur eine Aussage formulieren.
3. Lange und komplexe Sätze vermeiden.
4. Items positiv formulieren – möglichst keine negative Formulierungen verwenden.
5. Keine Universaldrücke wie „immer“, „nie“ oder „alle“ verwenden.
6. Es soll ein eindeutiger zeitlicher Bezug hergestellt werden.
7. Abkürzungen und Fachbegriffe vermeiden.
8. Items sollen ohne den Kontext der anderen Items beantwortet werden können.

#### 3.2.2. Itempolung

Gemäss Bühner (2006) zeigt sich, dass besonders negativ formulierte Items einen Einfluss auf das Antwortverhalten haben bzw. die Testpersonen verwirren. Die negative Formulierung wird dazu verwendet, einer Zustimmungs- oder Ja-sage-Tendenz entgegenzuwirken, welche als Akquieszenz bezeichnet wird (Hossiep, Paschen & Mühlhaus, 2000, zitiert nach Bühner, 2006).

Die Resultate zur Itempolung variieren und sind somit uneinheitlich. Es ist davon auszugehen, dass ein solcher Effekt auftritt. Wenn man positiv und negativ gepolte Items verwendet, sollte nach Bühner (2006) das Verhältnis zumindest ausgeglichen sein.

### 3.2.3. Skalenniveau

Für die Auswahl eines korrekten statistischen Verfahrens ist nach Zöfel (2003) die Feststellung des Skalenniveaus (Messniveau) der verwendeten Variablen von entscheidender Bedeutung. Es werden das Nominal-, Ordinal-, Intervall- und Verhältnisniveau unterschieden. In SPSS 22 werden das Intervall- und Verhältnisniveau als Skala bezeichnet.

**Nominalniveau:** Die Ziffern der Kodierung weisen keinerlei empirische Bedeutung auf. Als Beispiel einer nominalskalierten Variablen kann der Familienstand genannt werden. Einen Spezialfall bilden die dichotom nominalskalierten Variablen, z. B. ja/nein, richtig/falsch. Sie besitzen eine Ordnungsrelation und bilden den Übergang zwischen Nominal- und Ordinalniveau.

**Ordinalniveau:** Bei dieser Variable kommt den verwendeten Kodezahlen eine empirische Relevanz hinsichtlich ihrer Ordnung zu. Sie bezieht sich aber nicht auf die Differenz zweier Kodezahlen. Typische Beispiele sind die Schulbildung oder die Altersklasseneinteilung.

**Intervallniveau:** Dieser Variablen wird eine empirische Relevanz zugesprochen, weil sie eine Rangordnung aufweist und den Differenzen eine Bedeutung zukommt. Das Alter oder das Körpergewicht sind häufig verwendete Variablen. Der Übergang von Ordinal- zu Intervallniveau ist fließend und eine Einordnung in eines der beiden Niveaus erscheint manchmal als schwierig.

**Verhältnisniveau:** Bei diesen Variablen kommt nicht nur der Differenz zweier Werte, sondern auch dem Verhältnis zweier Werte eine empirische Bedeutung zu. Es handelt sich dabei um intervallskalierte Variablen, die den Wert Null annehmen können, wie beispielsweise die Temperatur in Grad Celsius oder der Intelligenzquotient.

### 3.2.4. Testgütekriterien

Gemäss Moosbrugger und Kelava (2012) stellen die Testgütekriterien ein Instrument der Qualitätsbeurteilung psychologischer Tests dar. Im Weiteren werden die drei Gütekriterien Objektivität, Reliabilität und Validität vorgestellt.

### 3.2.5. Objektivität

Nach Bühner (2011) sollte ein psychometrischer Test bei der Durchführung, der Auswertung und der Interpretation objektiv sein. Das heisst die Testleistung einer Person variiert nicht, auch wenn verschiedene Testleiter den Test durchführen, auswerten oder interpretieren. Es werden Durchführungs-, Auswertungs- und Interpretationsobjektivität unterschieden.

#### **Durchführungsobjektivität**

Die Durchführungsobjektivität ist gegeben, wenn das Testergebnis unabhängig davon ist, welcher Testleiter den Test durchführt. Mittels Standardisierung kann sie erhöht werden. Andere Bedingungen sollen konstant und kontrolliert sein, damit sie nicht als Störvariablen wirken können. Im Testmanual wird dazu definiert, wie und unter welchen Bedingungen ein Fragebogen durchzuführen ist. Es sind genaue Anweisungen zu Zeitbegrenzungen oder Hilfestellungen zur Beantwortung der Fragen anzugeben. Ausserdem muss festgelegt werden, wie mit allfälligen Fragen der Testperson umzugehen ist. Normalerweise verweist man auf die Instruktionen im Fragebogen (Bühner, 2011; Moosbrugger & Kelava, 2012).

#### **Auswertungsobjektivität**

Wenn bei einem vorliegenden Testprotokoll das Testergebnis nicht vom Testauswerter abhängt, ist die Auswertungsobjektivität gegeben – was von Multiple-Choice Aufgaben (Mehrfachwahlaufgaben) im Allgemeinen problemlos erreicht wird. Bei offenen Antwortformaten müssten detaillierte Auswertungsregeln vorliegen, deren einheitliche Anwendung empirisch zu überprüfen ist (Moosbrugger & Kelava, 2012).

## **Interpretationsobjektivität**

Die Standardisierung eines Testes beinhaltet klare Regeln für die Interpretation, welche über die Durchführungs- und Auswertungsvorschriften hinausgehen. Wenn unterschiedliche Testanwender bei Testpersonen mit demselben Testwert zu den gleichen Schlussfolgerungen kommen, liegt Interpretationsobjektivität vor. Dazu können im Testmanual Angaben von Ergebnissen (Normtabellen) aus der Eichstichprobe gemacht werden, die den Vergleich der Testperson mit der relevanten Bezugsgruppe ermöglicht (Moosbrugger & Kelava, 2012).

### 3.2.6. Reliabilität

„Die Reliabilität gibt den Grad der Messgenauigkeit eines Messwertes an“ (Bühner, 2011, S.60). Es geht also um den Grad der Genauigkeit, mit dem ein Test ein spezifisches Merkmal misst, unabhängig davon, was er zu messen beansprucht. Laut Moosbrugger und Kelava (2012) wird der Wert der Reliabilität über den Reliabilitätskoeffizienten (Rel.) angegeben – er liegt zwischen Null und Eins ( $0 \leq \text{Rel.} \leq 1$ ). Wenn der Wert Eins beträgt, sind keine Messfehler vorhanden. Null bedeutet, dass das Ergebnis ausschliesslich durch Messfehler zustande gekommen ist.

In der klassischen Testtheorie werden vier Methoden beschrieben, die das Ausmass der Reliabilität bestimmen: Die Retest-, Paralleltest- und Testhalbierungs-Reliabilität und die innere Konsistenz. Für diese Masterarbeit ist die innere Konsistenz relevant, welche nachfolgend beschrieben wird.

### **Innere Konsistenz**

Interne bzw. innere Konsistenz stellt eine Verallgemeinerung der Testhalbierungsmethode dar. Jedes Item wird als eigenständiger Testteil angesehen (Moosbrugger & Kelava, 2012). Somit entstehen gedanklich so viele Teile, wie Items vorhanden sind (Rauchfleisch, 2008). Je stärker die einzelnen Testteile untereinander positiv korrelieren, desto höher ist die interne Konsistenz,

welche als Cronbachs Alpha angegeben wird. Nach Bühner (2011) steht die innere Konsistenz für die Messgenauigkeit des Tests zu einem bestimmten Messzeitpunkt.

### 3.2.7. Validität

„Die Validität gibt an, ob der Test das auch wirklich misst, was er zu messen beansprucht“ (Bühner, 2011, S.60). Für Moosbrugger und Kelava (2012) handelt es sich bei der Validität, bezogen auf die Testpraxis, um das wichtigste Gütekriterium überhaupt. Objektivität und Reliabilität ermöglichen zwar eine hohe Messgenauigkeit, liefern aber nicht hinreichende Bedingungen für hohe Validität. Somit kann ein Test, der eine niedrige Reliabilität aufweist, keine hohe Validität haben. Ist die Validität aber hoch, so erlauben die Ergebnisse des Tests die Generalisierung des in der Testsituation beobachteten Verhaltens auf das zu messende Verhalten ausserhalb der Testsituation. Um ein differenziertes Bild der Validität (Gültigkeit) zu erhalten, untersucht man sinnvollerweise folgende Aspekte: Inhalts-, Augenschein-, Konstrukt- und Kriteriumsvalidität. Für die Kriteriumsvalidität sei auf die Literatur von Bühner (2011) sowie Moosbrugger und Kelava (2012) hingewiesen. Die anderen Aspekte, welche für diese Masterarbeit relevant sind, werden nachfolgend beschrieben.

#### **Inhaltsvalidität**

„Unter Inhaltsvalidität versteht man, inwieweit ein Test oder ein Test-Item das zu messende Merkmal repräsentativ erfasst“ (Moosbrugger & Kelava, 2012, S.15).

Dabei wird von einem Repräsentationsschluss ausgegangen. Das bedeutet, dass die Testitems eine repräsentative Stichprobe aus dem Item-Universum darstellen, womit das Merkmal erfasst wird. Die Inhaltsvalidität wird in der Regel „aufgrund logischer und fachlicher Überlegungen“ und nicht numerisch anhand eines Kennwertes bestimmt Michel und Conrad (1982, zitiert nach Bühner, 2011). Die Beurteilung der Inhaltsvalidität durch Experten spielt dabei eine bedeutende Rolle.



### **Augenscheinvalidität**

„Augenscheinvalidität gibt an, inwieweit der Validitätsanspruch eines Tests, vom blossen Augenschein her einem Laien gerechtfertigt erscheint“ (Moosbrugger & Kelava, 2012, S.15).

Bühner (2011) weist darauf hin, dass die Begriffe Augenscheinvalidität und Inhaltsvalidität eng verbunden sind. Jedoch ist die Augenscheinvalidität kein ausreichendes Validitätskriterium für die Güte eines Tests, da es sich nicht um ein wissenschaftliches Konzept handelt.

### **Konstruktvalidität**

„Ein Test weist Konstruktvalidität auf, wenn der Rückschluss vom Verhalten der Testperson innerhalb der Testsituation auf zugrunde liegende psychologische Persönlichkeitsmerkmale („Konstrukte“, „latente Variablen“, „Traits“) wie Fähigkeiten, Dispositionen, Charakterzüge, Einstellungen wissenschaftlich fundiert ist. Die Enge dieser Beziehung wird aufgrund von testtheoretischen Annahmen und Modellen überprüft“ (Moosbrugger & Kelava, 2012, S.16).

Bühner (2011) sowie Moosbrugger und Kelava (2012) unterscheiden zwischen konvergenter Validität und diskriminanter/divergenter Validität (struktursuchendes Vorgehen) sowie faktorieller Validität (strukturprüfendes Vorgehen). Für diese Masterarbeit ist die faktorielle Validität von Bedeutung.

### **Faktorielle Validität**

Mittels Faktorenanalyse werden häufig Zusammenhänge zwischen verschiedenen Tests untersucht. Dadurch werden einerseits homogene konstruktnahe Inhaltsbereiche verknüpft und andererseits konstruktferne Bereiche getrennt. Mithilfe der Faktorenanalyse kann die Zugehörigkeit bestimmter Items zu bestimmten Konstruktbereichen aufgezeigt werden (Bühner, 2011). Im Kapitel 0 wird detaillierter auf die Faktorenanalyse eingegangen.

### **Beziehung zwischen Objektivität, Reliabilität und Validität**

Die Hauptgütekriterien stehen in einem bestimmten Abhängigkeitsverhältnis zueinander. Ein Test erreicht höchst wahrscheinlich keine optimale Reliabilität, wenn er nicht objektiv ist. Es können Fehler bei der Ermittlung und der Interpretation der Ergebnisse entstehen, welche das Testergebnis verzerren und die Reliabilität des Tests beeinträchtigen. Wenn die Reliabilität gering ist, kann die Vorhersage auf ein Kriterium (Validität) nicht hoch sein (Bühner, 2011).

#### 3.3. Fragebogenkonstruktion

Die Fragebogenkonstruktion erfolgte in mehreren Schritten. Im ersten Schritt wurde mittels Literaturrecherche verschiedene Instrumente zur Erfassung von Visueller Ästhetik, User Experience/Nutzungserleben, Intrinsischer Motivation und Emotionen gesucht. Die verwendeten Instrumente sind in Tabelle 4 dargestellt. Eine ergänzte Darstellung mit Beispielimens befindet sich im Anhang.

Tabelle 4: Verwendete Instrument und Skalen zur Fragebogenkonstruktion

Instrument	Quelle	Subskalen	Cronbachs Alpha	Anzahl Items	Antwortformat
Dimensions of perceived visual aesthetics of web sites	Lavie & Tractinsky (2004)	Expressive Aesthetics	.87	5	7-stufig strongly disagree – strongly agree
		Pleasurable Interaction	.86	3	– strongly agree
Standardized User Experience Percentile Rank Questionnaire (SUPR-Q)	Sauro (2015)	Usability	.88	2	5-stufig strongly disagree – strongly agree
		Loyalty	.64	2	11-stufig not at all likely – extremely likely
User Engagement Scale (UES)	O'Brian & Toms (2009)	Focused attention	.92	7	5-stufig strongly disagree – strongly agree
		Involvement	.71	3	– strongly agree
		Perceived Usability	.91	8	
Modularer Fragebogen zur Erfassung des Nutzungserleben (meCUE)	Minge & Riedel (2013)	Nützlichkeit	.83	3	7-stufig stimme völlig zu – lehne völlig ab
		Benutzbarkeit	.89	3	– lehne völlig ab
Intrinsic Motivation	Isen & Reeve (2005)	Intrinsic Motivation	.92	8	7-stufig strongly disagree – strongly agree
Positive and Negative Affect Schedule (PANAS)	Krohne et al. (1996)	Positive Emotion	.85	10	5 stufig gar nicht - äusserst
		Negative Emotion	.86	10	
Ars reception survey (ARS)	Hager et al. (2012)	Positive attraction	.83	5	5-stufig completely disagree – completely agree
		Cognitive Stimulation	.90	5	
Achievment emotions questionnaire (AEQ)	Pekrun (2005)	Learning related boredom	.92	11	5-stufig strongly disagree – strongly agree

Im zweiten Schritt erfolgte eine provisorische Auswahl von maximal 25 Items. Diese Obergrenze wurde im Hinblick auf die Testsituation ausgewählt. Einerseits sollte es aus organisatorischen Gründen möglich sein, zwei Apps innerhalb von rund 60 Minuten zu bewerten, andererseits sollte die Aufmerksamkeitsspanne der Testpersonen nicht überstrapaziert werden.

Der dritte Schritt diente dazu, die englisch formulierten Items auf Deutsch zu übersetzen und anschliessend von einer englischsprachigen Person, die das Original-Item nicht kennt, wieder ins Deutsch rück-übersetzen zu lassen. Ziel dieses Vorgehens war die Vermeidung von Übersetzungsfehlern. Im Anhang sind die Übersetzungen der Items dokumentiert.

Der vierte Schritt beinhaltete die konkrete Itemformulierung. Dabei wurden die im Kap. 3.2.1 beschriebenen methodischen Grundlagen berücksichtigt. Daneben orientierte sich die Formulierung an der Testsituation, dem Testobjekt (App) und der Zielgruppe (Jugendliche). Entsprechend wurden Wörter wie „Webseite“ oder „Gemälde“ durch das Wort „App“ ersetzt, komplexe Formulierungen angepasst oder einzelne Wörter auf einen Aussagesatz ergänzt.

Im fünften Schritt kamen zwei neue Items zur Beurteilung des Tones dazu, die in dieser Form in keinem bestehenden Instrument zu finden waren. Zusätzlich erfolgte die Reduktion des Fragebogens auf 22 Items.

Der letzte Schritt beinhaltete die Durchführung eines Pretests mit den gleichen Unterlagen und Rahmenbedingungen wie der Haupttest. Der Pretest sollte unter anderem zeigen, ob die formulierten Items von den Schülerinnen und Schülern verstanden wurden. Dazu erhielt die Klasse nach dem Test nochmals einen Fragebogen ausgehändigt. Darauf markierten sie individuell, bei welchen Items sie Verständlichkeitsprobleme hatten. Nach einer Pause, in welcher die markierten Items miteinander verglichen wurden, erfolgte eine Diskussion über entsprechende Formulierungen in der Klasse. Daraus resultierte die Konkretisierung von Item 22 wie folgt: Aus „Das Benützen der App regt meine Gedanken an“ wurde „Das Benützen der App regt meine Gedanken zum Thema Astronomie/Physik an“, wobei sich „Astronomie“ auf die erste und „Physik“ auf die zweite App bezieht. Die restlichen Items wurden als verständlich eingestuft.

Tabelle 5 zeigt die definitive Version der Itemformulierungen. Daraus ist ersichtlich, dass die Empfehlungen von Bühner (2011) bezüglich Polung nicht vollständig einhalten wurden. Ziel war, hauptsächlich positiv gepolte Items zu verwenden. Die Verwendung von einzelnen negativ gepolten Items liegt daran, dass eine hohe Ähnlichkeit zu den Originalitems angestrebt wurde.

Tabelle 5: Definitive Version Fragebogen

Subskalen	Originalitem	Anpassung auf Testobjekt und Zielgruppe, definitive Formulierung	Nr.
Expressive Aesthetics	Fascinating Design.	Das Design der App ist faszinierend.	1
Expressive Aesthetics	Creative Design.	Das Design der App ist originell.	2
Expressive Aesthetics	Use of special effects.	Die Spezialeffekte in der App sind gut gemacht.	3
Pleasurable Interaction	I feel joyful.	Das Benützen der App macht Spass	4
Usability	The website is easy to use.	Die App ist einfach zu gebrauchen.	5
Usability	It's easy to navigate within the website.	Es ist einfach in der Webseite zu navigieren.	6
Loyalty	How likely are you to recommend this website to a friend or colleague?	Ich werde diese App weiter empfehlen.	7
Focused attention	I blocked out things around me, when I was shopping on this website	Bei der Benützung der App vergesse ich meine Umgebung.	8
Involvement	I was really drawn into my shopping task.	Ich werde richtig in das Thema reingezogen.	9
Learning related boredom	The material bores me to death.	Die App langweilt mich.	10
Perceived Usability	This shopping experience is demanding.	Die Benützung der App ist anspruchsvoll.	11
Nützlichkeit	Mithilfe des Produkts kann ich meine Ziele erreichen.	Mit der App kann ich die Lernaufgabe lösen.	12
Benutzbarkeit	Die Bedienung des Produkts ist verständlich.	Die Bedienung der App ist verständlich.	13
Ton, positiv		Der Ton passt zum Design.	14
Ton, negativ		Der Ton stört mich.	15
Intrinsic Motivation	It makes me feel curious about it.	Die App macht mich neugierig auf das Thema (Astronomie/Physik).	16
Intrinsic Motivation	It makes me want to explore it further.	Ich möchte das Thema am liebsten mit der App weiter erkunden.	17
Intrinsic Motivation	It's interesting.	Das Thema interessiert mich.	18
Positive Emotion	Wie fühlen Sie sich im Moment? begeistert.	Die App begeistert mich.	19
Negative Emotion	Wie fühlen Sie sich im Moment? verärgert.	Die App verärgert mich.	20
Positive attraction	This painting thrills me.	Die Benützung der App ist aufregend.	21
Cognitive Stimulation	This painting is thought-provoking.	Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	22

*Anmerkungen.* Dimensionen und Items in Originalsprache.

## **Antwortskala**

Bei der verwendeten Antwortskala im Fragebogen handelt es sich um eine 5-stufige unipolare Ratingskala in Anlehnung an Rohrmann (1978, zitiert nach Bühner, 2011). Dabei sind allerdings nur die Pole der Skala mit den verbalen Ankern „stimme gar nicht zu“ und „stimme voll zu“ versehen. Die Zwischenschritte wurden mit den Symbolen „-“, „-/“ und „+“ gekennzeichnet. Die Richtung der Zustimmung verläuft von links nach rechts. Der Verzicht auf die verbalen Anker in den Zwischenschritten sollte einerseits die Gleichabständigkeit gewährleisten sowie die Beantwortung erleichtern. Bei verbalen Ankern besteht nach Bühner (2011) die Möglichkeit, dass die Wortanker subjektiv unterschiedlich interpretiert werden können.

Obwohl in den Originalskalen häufig eine 7-stufige Likert-Skala verwendet wird, erschien die Reduktion auf fünf Stufen sinnvoller. Einerseits reichte dieser Differenzierungsgrad und andererseits sollte die Beantwortung im Hinblick auf die Zielgruppe nicht zu komplex und mit zu hohem kognitiven Aufwand verbunden sein. Für die statistische Auswertung wird dieses Antwortformat als „intervallskaliert“ angesehen.

## **Weitere Variablen**

Neben dem Fragebogen wurden weitere Variablen wie Alter, Geschlecht und „App-Nutzung pro Woche“ erhoben. Die „App-Nutzung“ diente als Kontrollvariable bezüglich Nutzergewohnheiten und Expertise im Zusammenhang mit Lern-Apps. Eine Übersicht aller erhobenen Variablen befindet sich im Anhang. Daten zum Schulort, Niveau, Klasse entstammen aus Vorabklärungen und mussten nicht mittels Fragebogen erhoben werden.

### 3.4. Testkonstruktion

Für die erste Validierung des Fragebogens wurde ein Test konstruiert, in dem zwei unterschiedliche Apps von Oberstufenschülerinnen und –schülern (Sekundarstufe I) beurteilt wurden. Die Auswahl der Apps erfolgte anhand folgender Kriterien:

- verfügbar für I-Pad (App-Store)
- inhaltlicher Bezug zu MINT-Themen v.a. Naturwissenschaft und Technik
- einfache Bedienbarkeit (mit kurzer Instruktion bedienbar)
- an Alter (12-16) der Stichprobe angepasst
- unterscheidbar bezüglich ihrer Ästhetik
- keine bis geringe Kosten
- keine Internetverbindung für Anwendung notwendig

Mithilfe von Link-Listen zu empfehlenswerten Lern-Apps, die Institutionen im Bereich Medienpädagogik veröffentlichten, wurde die Recherche vereinfacht. Erwähnt sei das Merkblatt „Lern-Apps: Interaktive und multimediale Lernbausteine“ vom Schulpsychologischen Dienst des Kantons Basellandschaft (2015) oder die Webseite „App-Advisor“ von der Beratungsstelle imedias. (Fachhochschule Nordwestschweiz, Pädagogische Hochschule, Institut Weiterbildung und Beratung, Beratungsstelle Digitale Medien in Unterricht und Schule-imedias, 2016).

Die zwei ausgewählten Apps „Solar Walk“ und „Physics at school lite“ sind in Tabelle 6 gegenübergestellt. Die Apps wurden vom Usability-Labor der Fachhochschule Nordwestschweiz in Olten auf die verwendeten I-Pads überspielt. Die Unabhängigkeit von einer Internetverbindung war insofern wichtig, als dass nicht alle Schulhäuser über eine funktionierende W-LAN-Infrastruktur verfügen.

Tabelle 6: Vergleich der Test-Apps

Kriterien	Solar Walk	Physics at school lite
Thema	Astronomie	Physik
Entwickler	Vito Technology	Vladimir Vascak
Kosten	Fr. 2.40	gratis
Ton	ja, sphärische Klänge als Hintergrundgeräusch abschaltbar	ja, passend zu den Animationen, z.B. Geräusch einer Dampflokomotive abschaltbar
Navigation	über Display und Menü	über Display und Menü
Steuerung	Vergrössern, verkleinern und drehen direkt auf Screen, Buttons	Buttons, Schieberegler
Inhalt	Informationen über unser Sonnensystem mit Planeten, Asteroiden, Kometen	steuerbare Animationen über physikalische Phänomene und deren Formeln
Effekte	3-D	Animationen
Interaktions-Stil*	eher metaphor-basiert	eher menü-basiert

\* vgl. Sutcliffe & De Angeli (2005)

Solar Walk ist eine App, in der unserem Sonnensystem im Weltall in 3-D-Ansicht rekonstruiert ist. Der Nutzende befindet sich mitten im All, kann zu den verschiedenen Himmelskörpern fliegen und erhält diverse Fakten und Zahlen zu den Objekten.

Physics at school lite, im weiteren Physics Lite genannt, ist eine App, in der verschiedene physikalische Phänomene in einfachen Animationen dargestellt werden. Der Nutzende kann die Animation manipulieren und einzelne Parameter verändern. Bei einigen Animationen ist die physikalische Formel dargestellt, deren Werte sich analog zu den Parametern verändern. Die zwei Apps unterscheiden sich auch hinsichtlich der Menüführung, respektive des Interaktionsstils in Anlehnung an Sutcliffe und De Angeli (2005). In Solar Walk ist der Bildschirm eher „metaphor-based“ gestaltet, während Physics Lite eher traditionell „menu-based“ kreiert wurde. Im Anhang sind zu jeder App Screenshots dargestellt.



Damit sich die Schülerinnen und Schüler mit fokussierter Aufmerksamkeit (vgl. Kap. 3.6) mit den zwei Apps beschäftigen, wurden pro App zwei Lernfragen gestellt. Somit konnten die Testpersonen innerhalb der App navigieren, Informationen finden und ein physikalisches Phänomen (Geschwindigkeit der Dampflokomotive) erklären. Der Schwierigkeitsgrad der Aufgaben wurde so gewählt, dass es gut möglich war, die Fragen zu beantworten. Eine zu hohe Schwierigkeit hätte womöglich demotivierend wirken können. In Kapitel 4.1 sind die Ergebnisse der Lernfragen wiedergegeben. Um ein simultanes Arbeiten mit dem I-Pad zu gewährleisten, beantworteten die Testpersonen den Fragebogen auf Papier.

Die Erläuterungen in den Testunterlagen wurden mit Bezug auf die Zielgruppe in einer verständlichen Sprache verfasst. Weiter sind zur guten Lesbarkeit Textabsätze und genügend Leerraum eingebaut. Im Fragebogen ist jedes zweite Item grau unterlegt, um das Verrutschen oder Überspringen der Zeilen zu verringern. Die Testunterlagen finden sich im Anhang.

### 3.5. Stichprobe

Der Zugang zur Stichprobe erfolgte über das persönliche Netzwerk des Autors. Entsprechend wurden verschiedene Lehrpersonen, die auf der Oberstufe (Sekundarstufe I) unterrichten, angefragt, ob sie mit ihrer Klasse oder ihren Klassen an dieser Erhebung teilnehmen möchten. Bei der Zusammensetzung sollten alle drei Jahrgänge und die drei verschiedenen Leistungsniveaus vertreten sein. Im Kanton Aargau entspricht die Realschule dem untersten, die Sekundarschule dem mittleren und die Bezirksschule dem höchsten Niveau. Im Kanton Zürich kann die Sekundarschule je nach Typus (A, B, C) dem unteren oder mittleren Niveau zugeordnet werden.

Schüler und Schülerinnen im höchsten Niveau besuchen das Gymnasium. Die Zusammensetzung der Stichprobe bezüglich Niveau und Jahrgängen erfüllte somit die gewünschte Unterschiedlichkeit (vgl. Tabelle 7). Das Durchschnittsalter betrug 13.74 Jahre ( $SD = 1.04$ ). Die Stichprobengröße richtete sich nach den statistischen Auswertungsmethoden. Sie sollte min-

destens so gross sein, dass eine Faktorenanalyse gerechnet werden konnte. Mit 160 Versuchspersonen entsprach sie den Anforderungen.

**Tabelle 7: Stichprobe**

Schulort	Kanton	Stufe/Klasse	Anzahl	männlich	weiblich	Bemerkung
Meisterschwanden	AG	Sek 2b	18	11	7	Pretest
Meisterschwanden	AG	Sek 2a	16	9	7	
Lenzburg	AG	Sek 3	15	7	8	
Lenzburg	AG	Sek 1b	25	17	8	
Lenzburg	AG	Sek 1a	23	13	10	
Zürich, Stettbach	ZH	Sek 2	19	13	6	
Zürich, Stettbach	ZH	Sek 3	17	10	7	
Dottikon	AG	Bez2a	16	8	8	
Dottikon	AG	Bez2b	15	7	8	
Dottikon	AG	Sek 3a	14	7	7	
Total Versuchspersonen			160 <sup>a</sup>	91 <sup>a</sup>	69 <sup>a</sup>	

<sup>a</sup>Total ohne Pretest

### 3.6. Testdurchführung

Wie im Kapitel 3.5 erwähnt, erfolgte der Zugang zu den Lehrpersonen mittels persönlicher Kontakte. Zudem vermittelten einige Lehrpersonen weitere Testklassen innerhalb ihrer Schulhäuser. Nach der generellen Zustimmung konnten die Lehrpersonen den Test-Termin mit wenigen Einschränkungen frei bestimmen. Als Bedingung galt lediglich ein verfügbares Zeitfenster von rund 60 Minuten. Die Lehrpersonen erhielten den Auftrag, die Schülerinnen und Schüler im Voraus zu fragen, ob sie am Test teilnehmen möchten. Es wäre möglich gewesen, dass einzelne Personen nicht teilgenommen und diese Zeit anders genutzt hätten. Insofern war die Teilnahme freiwillig, was bezüglich ethischen Überlegungen gemäss Huber (2009) anzustreben ist. Nach Absprache mit den Lehrpersonen wurde auch auf eine Einverständniserklärung der Eltern verzichtet, da die Befragung anonym und freiwillig durchgeführt wurde. Ein solches Verfahren hätte zudem einen hohen administrativen Aufwand bedeutet.

Um die Durchführungsobjektivität zu gewährleisten, folgte die konkrete Testdurchführung einem genauen Skript (siehe Anhang). Alle Tests wurden vom Autor persönlich realisiert. Der Pretest lieferte wichtige Informationen zum Zeitbedarf beim Ausfüllen und zur Chronologie der einzelnen Schritte. Hier ist anzumerken, dass auf starres Zeitregime verzichtet wurde. Je nach Klassen, die sich bezüglich Grösse, Alter und Schulniveau unterschieden, brauchten die Einführung, die Lernfragen oder das Ausfüllen der Fragebogen etwas mehr Zeit. Ziel der Untersuchung war das „ästhetische Empfinden“ und nicht die Erfassung einer Leistung. Zeitlicher Stress oder Verständlichkeitsprobleme hätten womöglich als Störvariablen die Ergebnisse verzerren können. Alles in allem bewegte sich der Zeitbedarf zwischen 50 und 65 Minuten, wobei keine Pause zwischengeschaltet wurde. Ab und zu mussten einzelne Schülerinnen oder Schüler darauf hingewiesen werden, nicht zu trödeln, damit der Zeitplan eingehalten werden konnte. Schnelle Schülerinnen und Schüler durften nicht selbständig mit der zweiten App beginnen.

In der Testeinführung wurden einige Schlüsselwörter (z.B. Design, Spezialeffekte), die in den Items vorkommen, erklärt (vgl. Skript Testdurchführung). Zusätzlich hing gut sichtbar ein Flip-Chart mit deren Bedeutungen. Auf dieses Vorgehen wird im Kapitel 3.8 eingegangen.

Aus dem Skript zur Testdurchführung ist ersichtlich, dass sich die Testpersonen in zwei Modi mit der App beschäftigen. Mit einer zielorientierten Herangehensweise, nämlich der Beantwortung der Lernfragen und einer selbstgesteuerten, erkundungsorientierten Herangehensweise. Die Zielorientierung sollte eher Aspekte der kognitiven Bewertung und der Usability der App provozieren, die freie Erkundung eher Aspekte der affektiven Bewertung. Als zusätzliche Information und zum Festhalten besonderer Ereignisse wurde für jeden Test ein Beobachtungsprotokoll ausgefüllt.

### 3.6.1. Methodische Grundlagen zur Datenauswertung

In folgenden Abschnitten werden Methoden erläutert, die bei der Analyse von Fragebogendaten relevant sind.

### 3.6.2. Itemanalyse

Gemäss Moosbrugger und Kelava (2012) werden nach der Planungs- und Entwicklungsphase eines Fragebogens in einem nächsten Schritt die Items an einer möglichst repräsentativen Stichprobe einer deskriptivstatischen Evaluation unterzogen. Die Itemanalyse dient zur Überprüfung des Fragebogens und stellt laut Bortz und Döring (2009) einen zentralen Aspekt der Testkonstruktion und Bewertung dar. Die Itemanalyse enthält folgende deskriptivstatistische Untersuchungsschritte, welche in den nachfolgenden Unterkapiteln genauer beschrieben werden: Rohwertverteilung, Itemschwierigkeit, Trennschärfe, Homogenität und Dimensionalität.

### 3.6.3. Rohwertverteilung

Durch die Häufigkeitsverteilung der Testwerte wird ein erster Überblick über das Antwortverhalten der untersuchten Testpersonen ermittelt. Die Werte der einzelnen Items lassen sich dabei grafisch als Histogramme darstellen. Darin ist z. B. abzulesen, wie stark die Testresultate streuen. Häufig entsteht ein Interesse, ob die Rohwertverteilung einer Normalverteilung entspricht, weil weitere statistische Verfahren normalverteilte Werte voraussetzen (Bortz & Döring, 2009). Die Testung auf Normalverteilung kann gemäss Bühner (2011) mit dem Kolmogorov-Smirnov-Anpassungstest (KSA) vorgenommen werden. Mittels Nullhypothese werden die Items auf Normalverteilung überprüft. Lehnt der KSA die Nullhypothese ab ( $p < 0.05$ ), sind die Daten nicht normalverteilt. Moosbrugger und Kelava (2012) legen dar, dass unter anderem die Schiefe entscheidet, ob die Testwertverteilung von der Normalverteilung oder einer anderen symmetrischen Verteilung abweicht. Wenn die Verteilung nicht symmetrisch ist, wird sie als schief respektive rechtssteil oder linkssteil bezeichnet.

#### 3.6.4. Itemschwierigkeit

Items besitzen nach Bortz und Döring (2009) unterschiedliche Zustimmungsraten, welche als Itemschwierigkeiten quantifizierbar sind. Items, die nur von wenigen Probanden bejaht werden, nennt man schwierige Items. Dagegen werden leichte Items von fast allen Testpersonen bejaht. Durch die Itemschwierigkeiten wird die Verteilung der Testwerte wesentlich beeinflusst. Bei einem Fragebogen ist darauf zu achten, dass die Items eine möglichst breite Schwierigkeitsstreuung aufweisen, um in allen Merkmalsbereichen gut differenzieren zu können. Items, die differenzieren sind informativ und machen Personenunterschiede sichtbar. Bei mehrstufigen Items (mehr als zwei Antwortmöglichkeiten) verwendet man gemäss Fisseni (2004) den Mittelwert als äquivalent zur Itemschwierigkeit. Er weist jedoch darauf hin, dass Mittelwerte ohne parallele Beachtung des Streuungsmasses wenig aussagekräftig sind.

#### 3.6.5. Trennschärfe

Inhaltlich drückt die Trennschärfe aus, wie gut ein Item die gesamte Sub-Dimension widerspiegelt, beziehungsweise wie hoch die Korrelation eines Items mit der entsprechenden Sub-Dimension ist. Nicht berücksichtigt wird die Korrelation mit anderen Sub-Dimensionen (Bühner, 2011). Laut Weise (1975, zitiert nach Fisseni, 2004) werden Trennschärfen unter 0.3 als niedrig, zwischen 0.3 bis 0.5 als mittel und über 0.5 als hoch betrachtet.

Bühner (2011) weist jedoch darauf hin, dass die Entfernung des Items auch inhaltlich und nicht nur durch die geringe Trennschärfe begründet werden soll. Laut Moosbrugger und Kelava (2012) begünstigt eine hohe Itemvarianz (Quadrat der Standardabweichung) eine erwünschte hohe Trennschärfe bei dichotomen sowie bei intervallskalierten Items. Dabei wird unter der Itemvarianz die Differenzierungsfähigkeit eines Items bezüglich der untersuchten Probandenstichprobe verstanden. Die Trennschärfe entspricht dabei den Gütekriterien Reliabilität und interne Validität (Rost, 2004).

### 3.6.6. Homogenität

Die Items von einem eindimensionalen Instrument stellen nach Bortz und Döring (2009) Operationalisierungen identischer Konstrukte dar. Aus diesem Grund sollten die Items untereinander korrelieren. Der Wert dieser wechselseitigen Korrelationen nennt man Homogenität. Somit erfassen die Items eines Tests bei hoher Homogenität ähnliche Informationen.

Die Mittelwerte der Inter-Item-Korrelationen (Itemhomogenitäten) gehen laut Bortz und Döring (2009) in den Alphakoeffizienten von Cronbach ein. Der Alphakoeffizient, gelegentlich Homogenitätsindex genannt, gibt den Grad der Homogenität für den Gesamttest an (Fisseni, 2004). Brosius (2011) beschreibt Cronbachs Alpha als ein zentrales Ergebnis der Reliabilitätsanalyse. Der Alphakoeffizient ist sowohl ein Mass für die Reliabilität einer aus mehreren Items zusammengesetzten Sub-Dimension als auch der inneren Konsistenz. Das heisst, der Wert gibt Aufschluss darüber, wie konsistent sich die Items zu Sub-Dimensionen zusammenfügen. Damit eine Sub-Dimension als zuverlässig angesehen werden kann, sollte der Cronbachs Alphakoeffizient einen Wert von mindestens 0.80 aufweisen, wobei dieser Wert nur eine grobe Orientierung darstellt.

### 3.6.7. Dimensionalität

Gemäss Bortz und Döring (2009, S. 221) gibt die Dimensionalität eines Tests an, „ob er nur ein Merkmal bzw. Konstrukt erfasst (eindimensionaler Test), oder ob mit den Testitems mehrere Konstrukte bzw. Teilkonstrukte operationalisiert werden (mehrdimensionaler Test).“ Die Dimensionalitätsüberprüfung wird üblicherweise mit explorativen oder konfirmatorischen Faktorenanalysen durchgeführt. Bühner (2011) weist darauf hin, dass die Faktorenanalyse Zusammenhänge zwischen Items auf latente Variablen (nicht direkt beobachtbare Grössen) aufzeigt. Die Korrelation zwischen Items desselben Konstrukts sollte höher sein, als die Korrelation zwischen Items unterschiedlicher Konstrukte. Die Faktorenanalyse ordnet demnach Items nach ihren korrelativen Ähnlichkeiten. Die Zuordnung entsteht aufgrund von Ähnlichkeiten der I-

temantworten. Das Ziel der Faktorenanalyse ist nach Brosius (2011) die Anzahl von Variablen auf möglichst wenige Faktoren (Komponenten), die hinter den beobachtbaren Items stehen, zu reduzieren. Dieser Vorgang ermöglicht letztendlich die Interpretierbarkeit der einzelnen Faktoren. Dabei spricht Weise (1975, zitiert nach Bortz & Döring, 2009) von einer mittelmässigen Validität, wenn der Koeffizient der Faktorladungen zwischen 0.4 und 0.6 ist und von hoher Validität bei einem Wert über 0.6.

Gemäss Brosius (2011) kommt die exploratorische Faktorenanalyse zur Anwendung, wenn keine spezifischen Hypothesen über die Anzahl oder Variablenstruktur von Faktoren in einem Datensatz vorhanden sind. Ziel dieser Faktorenanalyse ist eine generelle Datenreduktion oder die Untergliederung eines komplexeren Merkmalbereichs. Die Faktorenanalyse lässt sich nach Bühner (2011, S. 349) in vier Schritte einteilen: Dateneignung, Wahl der Methode, Extraktionskriterien für Faktoren und Wahl der Rotationsmethode. Diese vier Schritte werden nachfolgend aufgezeigt.

#### 3.6.8. Dateneignung

Vor der eigentlichen Faktorenanalyse werden die Daten auf ihre Dateneignung hin geprüft (Bühner, 2011). Als erstes werden gemäss Brosius (2011) Korrelationsmatrizen für alle in die Faktorenanalyse miteinbezogenen Variablen berechnet und dargestellt. Daraus kann abgeleitet werden, welche Variablen nur sehr gering mit den übrigen Variablen korrelieren und daher möglicherweise nicht berücksichtigt werden sollten. Sollte es der Fall sein, dass alle aufgezeigten Korrelationskoeffizienten nur sehr geringe relative Werte aufweisen (unter 0.3), würde es wenig Sinn machen, die Faktorenanalyse fortzusetzen. Dann kann angenommen werden, dass keine gemeinsamen Faktoren existieren.

In einem nächsten Schritt wird der Bartlett-Test durchgeführt. Dieses Verfahren überprüft laut Brosius (2011) die Nullhypothese, dass sämtliche Korrelationen zwischen den Variablen in der

Grundgesamtheit den Wert Null haben. Damit die Faktorenanalyse durchgeführt werden kann, muss die Nullhypothese abgelehnt werden.

Anschliessend soll der KMO-Test (Kaiser-Mayer-Olkin-Mass) Aufschluss darüber geben, ob sich die Dimensionen für eine Faktorenanalyse eignen. Bühner (2011) setzt eine hohe Korrelation zwischen den einzelnen Items für eine Dateneignung voraus. Werte, die eine Korrelation kleiner 0.5 aufweisen sind inkompatibel mit der Durchführung einer Faktorenanalyse (siehe Tabelle 8).

**Tabelle 8: Bewertung KMO-Koeffizient**

<b>Bewertung KMO-Koeffizient</b>	<b>Wert</b>
inkompatibel mit der Durchführung	> .50
schlecht	.50 – .59
mässig	.60 – .69
mittel	.70 – .79
gut	.80 – .89
sehr gut	≥ .90

Mithilfe des MSA-Koeffizienten (Measure of Sample Adequacy) kann nach Bühner (2011) eine Aussage getroffen werden, inwiefern sich das einzelne Item für die Faktorenanalyse eignet. Hier werden Korrelationen bzw. Partialkorrelationen eines Items mit den restlichen Items betrachtet. Zur Interpretation der Zahlenwerte gelten die gleichen Masse wie beim KMO-Koeffizient (Tabelle 8). Niedrige MSA-Koeffizienten können aber auch andere Gründe haben: Extreme Itemschwierigkeit respektive eine geringe Itemvarianz (Skala enthält ähnlich formulierte Items) oder eine geringe Reliabilität. Der MSA-Koeffizient kann aus der Anti-Image-Matrix abgelesen werden.



### 3.6.9. Methoden für exploratorische Faktorenanalyse

Es stehen verschiedene Verfahren zur Verfügung, um die den beobachteten Variablen zugrunde liegenden Faktoren zu bestimmen (Brosius, 2011). Die am häufigsten angewandten werden in folgenden Abschnitten erklärt.

Bei der Hauptkomponentenanalyse ist eine lineare Zerlegung der Itemvarianzen (Quadrat der Standardabweichung) in wenige Komponenten das Ziel. Sie soll lediglich beschreiben, welche Items deskriptiv zu einer Komponente zusammengefasst werden können (Bühner, 2011).

Die Hauptachsenanalyse wird angewendet, um Faktoren zu identifizieren, mit denen sich die Beziehungsstruktur beobachteter Variablen erklären lässt (Brosius, 2011)

Die dritte Methode ist die Maximum-Likelihood-Analyse. Dies ist ein Signifikanztest zur Ermittlung der Faktorenzahl. Diese Methode wird verwendet, wenn die Korrelationen der Items durch weniger Faktoren erklärt werden sollen (Bühner, 2011).

### 3.6.10. Extraktionskriterien für Faktoren

Anschliessend gilt es in einem nächsten Schritt die Faktorenzahl zu bestimmen. Bühner (2011) gibt als Methoden die explizite Theorie zur Faktorenextraktion und Scree-Test nach Cattell an. Bei der ersten Variante wird die Anzahl der zu extrahierenden Faktoren vorgegeben. Sie wird verwendet, wenn die Zuordnung der Items zu den Skalen bereits bekannt ist. Beim Scree-Test nach Cattell wird nach einem bedeutsamen Eigenwertabfall gesucht. Bei dieser Extraktionsmethode werden die Eigenwerte zuerst in einem Scree-Plott abgetragen und anschliessend ein Knick im Scree-Plot von links nach rechts gesucht. Die Anzahl Faktoren ergibt sich durch den Knick. Je weiter rechts sich der Knick befindet, um desto mehr Faktoren handelt es sich. Ein weiteres Extraktionskriterium für die Anzahl Faktoren stellt das Kaiser-Guttman-Kriterium dar, wonach nur Faktoren mit Eigenwerten  $> 1$  aufgenommen werden. Der Eigenwert gibt an, wie gut ein Faktor in der Lage ist, Unterschiede in der Beantwortung aller Items zu erklären.

Wenn der Eigenwert grösser als eins ist, kann er mehr Unterschiede aufklären als ein einziges Item und wird deshalb als bedeutsam angesehen (Bühner, 2011).

#### 3.6.11. Wahl der Rotationsmethode

In den bisherigen Schritten wurden die Faktoren lediglich als rein rechentechnische Ergebnisse betrachtet. Um die Faktoren besser interpretieren zu können, werden die Faktorenladungen in einem Koordinatensystem dargestellt (Brosius, 2011). Dabei unterscheidet Bühner (2011) zwei Rotationsmethoden. Einerseits die Orthogonale Rotation und andererseits die Oblique Rotation. Die Orthogonale Rotation wird bei unkorrelierten Faktoren gewählt. Diese Form von Rotation kann durch verschiedene Methoden durchgeführt werden. Es wird jedoch die Verwendung der Varimax-Methode empfohlen.

Demgegenüber wird die Oblique Rotation bei korrelierten Faktoren verwendet. Falls unklar ist, ob die Faktoren korrelieren, ist die Oblique Rotation vorzuziehen. Als empfohlene Methode wird hier die Promax-Methode angegeben.

#### 3.6.12. Weitere Werte der Faktorenanalyse

Als Kommunalität ( $h^2$ ) wird jener Teil der Itemvarianz bezeichnet, welcher durch die Faktoren erklärt werden kann. Bei der Hauptkomponentenanalyse wird davon ausgegangen, dass die gesamte Varianz der Items durch die Faktoren erklärt werden kann. Der Wert entspricht 100%, also Eins. Im Unterschied dazu wird bei der Hauptachsenanalyse schon zu Beginn angenommen, dass nicht die gesamte Itemvarianz erklärt werden kann. Die Kommunalitäten wird deshalb geschätzt und ist immer kleiner als eins (Leonhart, 2010).

#### 3.6.13. Zahlenwerte für Items und Skalen

Für eine bessere Übersicht sind die Zahlenwerte für die Analyse von Items und Skalen aus den vorangehenden Abschnitten in der Tabelle 9 zusammengefasst. In der Fachliteratur werden je nach Autor die Zahlenwerte auch anders interpretiert.

Tabelle 9: Kennwerte für Items und Skalen

Kennwerte	Niedrig	Mittel	Hoch
Trennschärfe	< .30	.30 – .50	> .50
Reliabilität	< .80	.80 – .90	> .90
Validität (Koeffizient Faktorladung)	< .40	.40 – .60	> .60

### 3.7. Vorgehen bei der statistischen Auswertung

In einem ersten Schritt wurden die Resultate aus den Papierfragebogen in eine Exceldatei übertragen und anschliessend in SPSS 22 importiert.

Die Kodierung der Variablen ist in Tabelle 10 ersichtlich. Fehlende Werte wurden mit Kode 99 versehen. Jeder Fragebogen erhielt eine ID-Nummer, sodass bei möglichen Daten-Auffälligkeiten auf das Original zurückgegriffen werden konnte. Die statistische Auswertung mit SPSS 22 erfolgte mit den Verfahren: Test auf Normalverteilung, Deskriptive Statistik, Dateneignung, Faktorenanalyse, Reliabilitätsanalyse, Mittelwertvergleiche, Korrelationen (vgl. Kapitel 0).

Bei der Faktorenanalyse wurden in einem iterativen Vorgehen verschiedene Verfahren mit unterschiedlichen Item-Auswahlen gerechnet (Hauptkomponentenanalyse, Hauptachsenanalyse, Varimax- und Promax-Rotation). In den Ergebnissen sind nur noch die Resultate der Hauptachsenanalyse mit Promax-Rotation wieder gegeben. Alle Berechnungen wurden für jede Test-App einzeln durchgeführt. Aus der Auswertung mussten die Fragebogen von zwei männlichen Testpersonen ausgeschlossen werden, da sie aufgrund sprachlicher Kompetenzen nur lückenhaft ausgefüllt waren. Entsprechend reduzierte sich die Stichprobe auf  $N=158$ .

Tabelle 10: Variablenübersicht

Variablen	Skala	Stufen	Antwort-Kodierung
Test-App	Nominal	2-stufig	1=SolarWalk, 2=PhysicsLite
Schulort	Nominal	4-stufig	1=Meisterschwanden, 2=Lenzburg, 3=Stettbach, 4=Dottikon
Niveau	Nominal	3-stufig	1=Realschule (tief), 2=Sekundarschule (mittel), 3=Bezirksschule (hoch)
Klasse	Nominal	3-stufig	1=1. Klasse, 2=2. Klasse, 3=3. Klasse
AE.ExprAe: Das Design der App ist faszinierend.	Intervall	5-stufig	1=stimme gar nicht zu, 2=-, 3=-/+ , 4=+, 5=stimme voll zu
AE.ExprAe: Das Design der App ist originell.	Intervall	wie oben	wie oben
AE.ExprAe: Die Spezialeffekte in der App sind gut gemacht.	Intervall	wie oben	wie oben
AE.Plea.Int: Das Benützen der App macht Spass.	Intervall	wie oben	wie oben
SPQ.Usa: Die App ist einfach zu gebrauchen.	Intervall	wie oben	wie oben
SPQ.Usa: Es ist einfach in der App zu navigieren.	Intervall	wie oben	wie oben
SPQ.Loy: Ich werde diese App weiter empfehlen.	Intervall	wie oben	wie oben
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	Intervall	wie oben	wie oben
UES.Inv: Ich werde richtig in das Thema reingezogen.	Intervall	wie oben	wie oben
LBor: Die App langweilt mich. <sup>a</sup>	Intervall	wie oben	wie oben
UES.Perc.Us: Die Benützung der App ist anspruchsvoll. <sup>a</sup>	Intervall	wie oben	wie oben
meCu: Mit der App kann ich die Lernaufgabe lösen.	Intervall	wie oben	wie oben
meCu: Die Bedienung der App ist verständlich.	Intervall	wie oben	wie oben
Ton: Der Ton passt zum Design.	Intervall	wie oben	wie oben
Ton: Der Ton stört mich. <sup>a</sup>	Intervall	wie oben	wie oben
In.Mot: Die App macht mich neugierig auf das Thema. (Astronomie/Physik).	Intervall	wie oben	wie oben
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	Intervall	wie oben	wie oben
In.Mot: Das Thema interessiert mich.	Intervall	wie oben	wie oben
Pan.Emot: Die App begeistert mich.	Intervall	wie oben	wie oben
Pan.Emot: Die App verärgert mich. <sup>a</sup>	Intervall	wie oben	wie oben
Ars.Patt: Die Benützung der App ist aufregend.	Intervall	wie oben	wie oben
Ars.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	Intervall	wie oben	wie oben
Alter	Verhältnis	Zahl	Anzahl Jahre
Geschlecht	Nominal	2-stufig	1=männlich, 2=weiblich
Leistung Lernaufgaben	Ordinal	3-stufig	0=keine Aufgabe richtig, 1=1 Aufgabe richtig, 2=2 Aufgaben richtig
App-Nutzung: Wie oft arbeitest du mit einer Lern-App innerhalb einer Woche?	Ordinal	3-stufig	1=nie, 2=einmal in der Woche, 3=mehrmals in der Woche

<sup>a</sup>Diese Items wurden invers rekodiert.

### 3.8. Bezugnahme zu Gütekriterien

#### **Durchführungsobjektivität**

Wie in Kapitel 3.6 beschrieben, wurde die Untersuchung im Feld bzw. in den Klassenzimmern durchgeführt. Insofern unterschieden sich die Orte und Zeitpunkte (morgens und mittags), was die Möglichkeit zur Kontrolle von entsprechenden Störvariablen ausschloss (vgl. Kapitel 3.2.5). Es wurde darauf geachtet, dass innerhalb der Klassenzimmer ruhige Bedingungen vorherrschten, so dass sich die Testpersonen konzentrieren konnten. Die Lehrpersonen waren jeweils anwesend und wiesen die Schülerinnen und Schüler bei Bedarf an, sich mit den gestellten Aufgaben zu befassen. Um die Verständlichkeit der Items zu gewährleisten, wurden einige Schlüsselwörter im Voraus erklärt (siehe Kapitel 3.6). Dieses Vorgehen ist im Zusammenhang mit Fragebogenstudien etwas unorthodox (vgl. Bühner, 2011). Im Hinblick auf die Stichprobe war es von entscheidender Bedeutung, dass die Testpersonen die Items auch wirklich verstanden. Bei Schülerinnen und Schülern in der Volksschule sind die sprachlichen Kompetenzen in Deutsch teilweise sehr unterschiedlich, unter anderem deshalb, weil je nach Schulort ein erheblicher Anteil der Jugendlichen einen Migrationshintergrund aufweist. Mit Bezug auf Kapitel 3.2.5 wurde der Umgang mit Fragen wie folgt geregelt: Grundsätzlich wurde auf die Testunterlagen oder auf den Flip-chart mit den Schlüsselwörtern hingewiesen. Bei schwerwiegenden sprachlichen Problemen erfolgte eine Hilfestellung. Bei Problemen mit der Navigation in der Software oder bei technischen Problemen gab es eine Unterstützung. Insgesamt musste wenig Hilfestellung geleistet werden.

Auf den Vertausch der Bearbeitungsreihenfolge der zwei Apps wurde verzichtet, um den Ablauf der Untersuchung in der gesamten Stichprobe gleich zu halten.

### **Interpretationsobjektivität**

Die Interpretationsobjektivität (vgl. Kapitel 3.2.5) der Untersuchung ist gegeben, da einerseits alle Auswertungsschritte nur vom Autor vorgenommen wurden und andererseits keine offenen Antwortformate zu bearbeiten waren.

## 4. Ergebnisse

In diesem Kapitel sind die Ergebnisse der statistischen Überprüfung beschrieben, die zur Anpassung des Fragebogens in Kapitel 4.8 führen.

### 4.1. Deskriptive Statistik

Die deskriptive Statistik gibt Aufschluss über die Rohwertverteilung und die Itemschwierigkeit. Der Kolmogorov-Smirnov-Anpassungstest (KSA) zeigt bei allen Items des Fragebogens keine Normalverteilung (siehe Anhang). Die Werte der deskriptiven Statistik für die Fragebogenitems sind aus Tabelle 11 und Tabelle 12 zu entnehmen. Um die Verteilung der Häufigkeiten bei den verschiedenen Items besser analysieren zu können, wurden Histogramme ausgegeben (siehe Anhang). Die Häufigkeiten der weiteren Variablen wie Geschlecht, Niveau, Lernaufgabe, App-Nutzung pro Woche sind in den Tabellen 13 bis 17 ersichtlich. Zur besseren Übersichtlichkeit sind Balkendiagramme direkt in die Tabellen integriert.

Tabelle 11: Deskriptive Statistik Solar Walk

	<i>N</i>	<i>Min.</i>	<i>Max.</i>	<i>MW</i>	<i>SD</i>	<i>V</i>	<i>Schiefe</i>	<i>SE</i>
AE.Expr.Ae: Das Design der App ist faszinierend.	158	3	5	4.42	0.698	0.487	-0.784	0.193
AE.Expr.Ae: Das Design der App ist originell.	158	1	5	4.17	0.808	0.652	-0.910	0.193
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	158	2	5	4.36	0.733	0.538	-0.880	0.193
AE.Plea.Int: Das Benützen der App macht Spass.	158	1	5	4.25	0.901	0.811	-1.092	0.193
SPQ.Usa: Die App ist einfach zu gebrauchen.	158	2	5	4.35	0.732	0.536	-0.962	0.193
SPQ.Usa: Es ist einfach in der App zu navigieren.	158	1	5	4.13	0.875	0.766	-0.841	0.193
SPQ.Loy: Ich werde diese App weiter empfehlen.	157	1	5	4.00	0.954	0.910	-0.987	0.194
UES.FA: Bei der Benützung der App ver- gesse ich meine Umgebung.	157	1	5	3.25	1.250	1.563	-0.216	0.194
UES.Inv: Ich werde richtig in das Thema reingezogen.	158	1	5	3.96	0.950	0.903	-0.510	0.193
R-LBor: Die App langweilt mich.*	158	2	5	4.34	0.894	0.800	-1.272	0.193
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.*	156	1	5	3.58	1.164	1.354	-0.428	0.194
meCu.Nütz: Mit der App kann ich die Lern- aufgabe lösen.	156	1	5	4.49	0.741	0.548	-1.668	0.194
meCu.Benütz: Die Bedienung der App ist verständlich.	158	1	5	4.36	0.759	0.576	-1.244	0.193
Ton: Der Ton passt zum Design.	158	1	5	4.16	0.990	0.979	-1.296	0.193
R-Ton: Der Ton stört mich.*	157	1	5	4.05	1.208	1.459	-1.160	0.194
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	157	1	5	3.80	1.036	1.074	-0.457	0.194
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	158	1	5	3.98	0.974	0.949	-0.841	0.193
In.Mot: Das Thema interessiert mich.	157	1	5	3.57	1.145	1.310	-0.481	0.194
Pan.Emot: Die App begeistert mich.	157	1	5	4.04	0.857	0.735	-0.580	0.194
R-Pan.Emot: Die App verärgert mich.*	156	2	5	4.65	0.689	0.475	-2.043	0.194
ARS.PAtt: Die Benützung der App ist auf- regend.	158	1	5	3.80	1.031	1.063	-0.621	0.193
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Phyik) an.	158	1	5	3.66	1.133	1.284	-0.546	0.193
Gültige Werte (Listenweise)	146							

Anmerkungen. *MW*=Mittelwert, *V*=Varianz, *SD*=Standardabweichung, *SE*=Standardfehler  
 \*invers kodierte Items



Tabelle 12: Deskriptive Statistik Physics Lite

	<i>N</i>	<i>Min.</i>	<i>Max.</i>	<i>MW</i>	<i>SD</i>	<i>V</i>	<i>Schiefe</i>	<i>SE</i>
AE.Expr.Ae: Das Design der App ist faszinierend.	158	1	5	3.04	0.986	0.973	-0.049	0.193
AE.Expr.Ae: Das Design der App ist originell.	158	1	5	3.19	1.054	1.110	-0.123	0.193
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	158	1	5	3.15	1.188	1.412	-0.101	0.193
AE.Plea.Int: Das Benützen der App macht Spass.	157	1	5	3.25	1.130	1.278	-0.153	0.194
SPQ.Usa: Die App ist einfach zu gebrauchen.	157	1	5	3.50	1.130	1.277	-0.316	0.194
SPQ.Usa: Es ist einfach in der App zu navigieren.	156	1	5	3.53	1.098	1.206	-0.333	0.194
SPQ.Loy: Ich werde diese App weiter empfehlen.	157	1	5	2.85	1.148	1.318	0.137	0.194
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	157	1	5	2.01	1.121	1.256	1.109	0.194
UES.Inv: Ich werde richtig in das Thema reingezogen.	157	1	5	2.42	1.127	1.271	0.418	0.194
R-LBor: Die App langweilt mich.*	157	1	5	3.11	1.166	1.359	-0.005	0.194
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.*	157	1	5	3.48	1.141	1.303	-0.222	0.194
meCu.Nütz: Mit der App kann ich die Lernaufgabe lösen.	155	1	5	3.55	1.135	1.288	-0.273	0.195
meCu.Benütz: Die Bedienung der App ist verständlich.	157	1	5	3.55	1.106	1.224	-0.266	0.194
Ton: Der Ton passt zum Design.	155	1	5	3.52	1.235	1.524	-0.436	0.195
R-Ton: Der Ton stört mich.*	155	1	5	3.79	1.190	1.415	-0.657	0.195
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	157	1	5	2.66	1.072	1.148	0.266	0.194
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	157	1	5	2.76	1.140	1.300	0.147	0.194
In.Mot: Das Thema interessiert mich.	156	1	5	3.12	1.180	1.393	-0.036	0.194
Pan.Emot: Die App begeistert mich.	156	1	5	2.73	1.074	1.153	0.271	0.194
R-Pan.Emot: Die App verärgert mich.*	155	1	5	3.71	1.168	1.363	-0.555	0.195
ARS.PAtt: Die Benützung der App ist aufregend.	156	1	5	2.71	1.072	1.149	0.248	0.194
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Phyik) an.	156	1	5	2.64	1.113	1.238	0.208	0.194
Gültige Werte (Listenweise)	147							

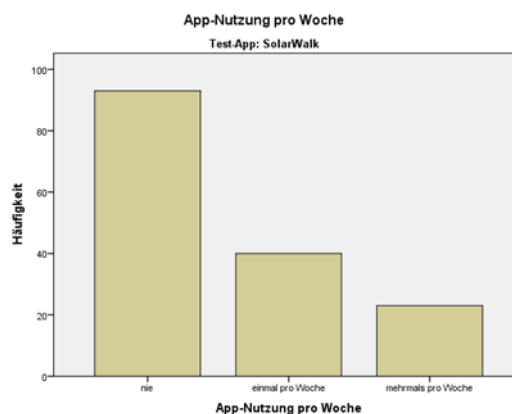
Anmerkungen. *MW*=Mittelwert, *V*=Varianz, *SD*=Standardabweichung, *SE*=Standardfehler  
 \*invers kodierte Items

**Tabelle 13: Deskriptive Statistik Geschlecht, Niveau, Lernaufgabe, App-Nutzung, Alter**

Werte	Geschlecht	Niveau	Lernaufgabe	App-Nutzung pro Woche	Alter
<i>N</i> Gültig	158	158	158	156	157
<i>N</i> Fehlend	0	0	0	2	1
Mittelwert <i>MW</i>	1.44	1.98	1.82	1.55	13.74
Standardabweichung <i>SD</i>	0.498	0.643	0.435	0.739	1.039
Varianz <i>V</i>	0.248	0.414	0.189	0.546	1.079
Minimum	1	1	0	1	12
Maximum	2	3	2	3	17

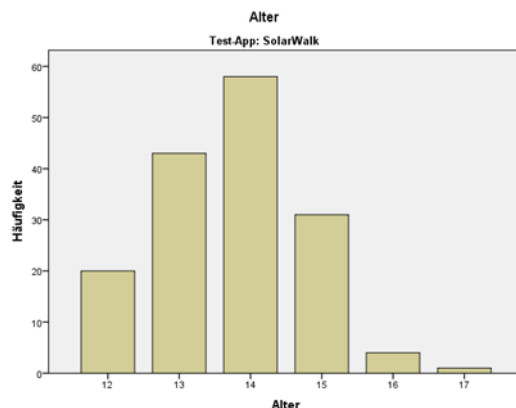
**Tabelle 14: Häufigkeiten App-Nutzung**

App-Nutzung	Häufigkeit	Prozent	Gültige Prozente
nie	92	58.9	59.6
einmal pro Woche	41	25.3	25.6
mehrmals pro Woche	23	14.6	14.7
Gesamt	156	98.7	100.0
Fehlend 99	2	1.3	
Gesamt	158	100.0	



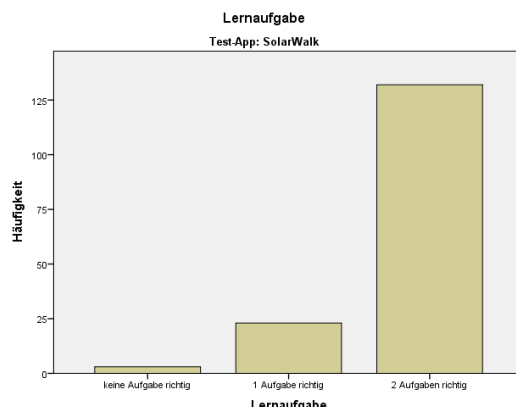
**Tabelle 15: Häufigkeiten Alter**

Alter	Häufigkeit	Prozent	Gültige Prozepte
12	20	12.7	12.7
13	43	27.2	27.4
14	58	36.7	36.9
15	31	19.6	19.7
16	4	2.5	2.5
17	1	0.6	0.6
Gesamt	157	99.4	100.0
Fehlend 99	1	0.6	
Gesamt	158	100.0	



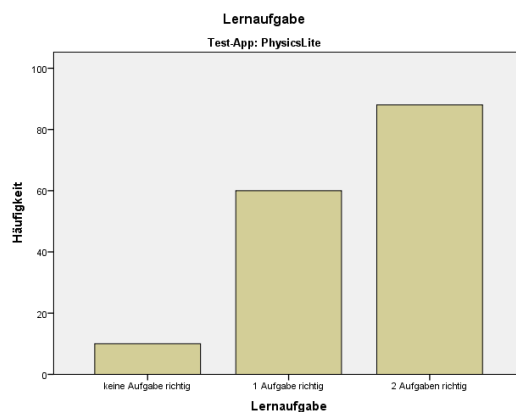
**Tabelle 16: Häufigkeiten Lernaufgaben Solar Walk**

Leistung	Häufigkeit	Prozent	Gültige Prozepte
keine Aufgabe richtig	3	1.9	1.9
1 Aufgabe richtig	23	14.6	14.6
2 Aufgaben richtig	132	83.5	83.5
Gesamt	158	100.0	100.0



**Tabelle 17: Häufigkeiten Lernaufgaben Physics Lite**

Leistung	Häufigkeit	Prozent	Gültige Prozepte
keine Aufgabe richtig	10	6.3	6.3
1 Aufgabe richtig	60	38.0	38.0
2 Aufgaben richtig	88	55.7	55.7
Gesamt	158	100.0	100.0



## 4.2. Faktorenanalyse

Die Dimensionalitätsüberprüfung des Fragebogens wurde mit einer exploratorischen Faktorenanalyse durchgeführt. Wie in Kapitel 3.6.8 beschrieben, erfolgte zuerst eine Prüfung der Werte auf Dateneignung. In einem ersten Schritt wurden alle Items des Fragebogens berücksichtigt. In einem zweiten Schritt erfolgte der Ausschluss des Items Nr. 18. „Das Thema interessiert mich.“ Die Begründung dazu ist im Kapitel 4.6.1 erläutert. Im Folgenden sind in Tabelle 18 die Werte dieser Berechnungsgrundlage zusammengefasst. Um diese Daten von den anderen Versionen abzugrenzen, sind die Tabellen mit dem Zusatz V1 versehen.

**Tabelle 18: KMO- und Bartlett-Test V1**

KMO- und Bartlett-Test		Solar Walk	Physics Lite
Maß der Stichprobeneignung nach Kaiser-Meyer-Olkin.		0.802	0.894
Bartlett-Test auf Sphärizität	Ungefähres Chi-Quadrat	1216.332	1567.198
	<i>df</i>	210	210
	Signifikanz nach Bartlett	0.000	0.000

*Anmerkungen.* Hauptachsenanalyse mit Promax-Rotation, *df*=Freiheitsgrade

Die MSA-Koeffizienten aus der Anti-Image-Matrix bewegen sich bei Solar Walk zwischen .56 und .90. Das Item „Die Benützung der App ist anspruchsvoll.“ ist mit einem Wert von .40 auffällig niedrig. Bei Physics Lite variieren die Werte zwischen .61 und .93. Tabelle 19 zeigt die Kommunalitäten vor und nach der Extraktion.

Tabelle 19: Kommunalitäten Solar Walk, Physics Lite V1

Items	<i>Solar Walk</i>		<i>Physics Lite</i>	
	$h^2$ Anfänglich	$h^2$ Extraktion	$h^2$ Anfänglich	$h^2$ Extraktion
AE.Expr.Ae: Das Design der App ist faszinierend.	.517	.522	.573	.617
AE.Expr.Ae: Das Design der App ist originell.	.544	.662	.484	.471
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	.240	.165	.618	.626
AE.Plea.Int: Das Benützen der App macht Spass.	.569	.546	.520	.468
SPQ.Usa: Die App ist einfach zu gebrauchen.	.506	.464	.619	.703
SPQ.Usa: Es ist einfach in der App zu navigieren.	.409	.364	.607	.629
SPQ.Loy: Ich werde diese App weiter empfehlen.	.474	.502	.677	.661
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	.499	.410	.410	.336
UES.Inv: Ich werde richtig in das Thema reingezogen.	.544	.529	.606	.652
R-LBor: Die App langweilt mich.	.526	.457	.453	.443
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	.209	.090	.305	.257
meCu.Nütz: Mit der App kann ich die Lernaufgabe lösen.	.198	.090	.317	.273
meCu.Benütz: Die Bedienung der App ist verständlich.	.458	.492	.559	.664
Ton: Der Ton passt zum Design.	.551	.772	.377	.319
R-Ton: Der Ton stört mich.	.487	.432	.288	.236
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	.562	.579	.575	.601
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	.538	.544	.691	.707
Pan.Emot: Die App begeistert mich.	.458	.432	.634	.608
R-Pan.Emot: Die App verärgert mich.	.513	.423	.451	.318
ARS.PAtt: Die Benützung der App ist aufregend.	.526	.439	.534	.485
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	.474	.471	.594	.597

*Anmerkungen.* Extraktionsmethode: Hauptachsen-Faktorenanalyse mit Promax-Rotation,  $h^2$  = Kommunalität.

In weiteren Schritten wurden verschiedene Items sukzessive ausgeschlossen und die Berechnungen wiederholt. Die definitive Lösung beinhaltet den Ausschluss folgender Items:

- 7. Ich werde diese App weiter empfehlen.
- 12. Mit der App kann ich die Lernaufgabe lösen.
- 14. Der Ton passt zum Design.
- 15. Der Ton stört mich.
- 18. Das Thema interessiert mich.

Im Folgenden sind nur noch die Resultate der definitiven Version beschrieben. Tabelle 20 zeigt die Werte der Überprüfung auf Dateneignung.

**Tabelle 20: KMO- und Bartlett-Test**

<b>KMO- und Bartlett-Test</b>	<b>Solar Walk</b>	<b>Physics Lite</b>
Maß der Stichprobeneignung nach Kaiser-Meyer-Olkin.	0.821	0.897
Bartlett-Test auf Sphärizität	953.433	1263.473
	136	120
	0.000	0.000

*Anmerkungen.* Hauptachsenanalyse mit Promax-Rotation

Die Anti-Image-Matrizen mit den MSA-Koeffizienten sowie die Korrelationsmatrizen sind im Anhang ersichtlich. In Tabelle 21 sind die Kommunalitäten vor und nach der Extraktion angegeben.

Tabelle 21: Kommunalitäten Solar Walk, Physics Lite

Items	<i>Solar Walk</i>		<i>Physics Lite</i>	
	$h^2$ Anfänglich	$h^2$ Extraktion	$h^2$ Anfänglich	$h^2$ Extraktion
AE.Expr.Ae: Das Design der App ist faszinierend.	.469	.666	.568	.723
AE.Expr.Ae: Das Design der App ist originell.	.485	.598	.474	.537
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	.221	.170	.543	.571
AE.Plea.Int: Das Benützen der App macht Spass.	.562	.617	.517	.481
SPQ.Usa: Die App ist einfach zu gebrauchen.	.484	.593	.581	.687
SPQ.Usa: Es ist einfach in der App zu navigieren.	.355	.416	.591	.661
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	.487	.490	.372	.329
UES.Inv: Ich werde richtig in das Thema reingezogen.	.538	.595	.595	.629
R-LBor: Die App langweilt mich.	.476	.532	.413	.501
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	.190	.092	.274	.256
meCu.Benütz: Die Bedienung der App ist verständlich.	.416	.468	.507	.608
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	.559	.533	.564	.578
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	.527	.542	.658	.675
Pan.Emot: Die App begeistert mich.	.442	.545	.388	.405
R-Pan.Emot: Die App verärgert mich.	.431	.407	.616	.636
ARS.PAtt: Die Benützung der App ist aufregend.	.500	.431	.520	.504
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	.445	.483	.560	.571

*Anmerkungen.* Extraktionsmethode: Hauptachsen-Faktorenanalyse mit Promax-Rotation,  $h^2$ =Kommunalität

Die Kommunalität von Item „Die Benützung der App ist anspruchsvoll.“ ist bei beiden Apps auffällig niedrig mit  $h^2=0.11$  bis  $0.27$ .

Aus Tabelle 22 und Tabelle 23 lässt sich die erklärte Gesamtvarianz bei der Extraktion von vier Faktoren entnehmen. Die dazu gehörigen Scree-Plots der Eigenwerte sind in Abbildung 6 und Abbildung 7 ersichtlich.

**Tabelle 22: Erklärte Gesamtvarianz Solar Walk**

Faktor	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulier-te %	Gesamt
1	5.441	32.008	32.008	4.964	29.201	29.201	4.028
2	2.088	12.280	44.288	1.568	9.226	38.427	3.675
3	1.477	8.689	52.977	1.038	6.106	44.533	2.978
4	1.092	6.425	59.402	0.606	3.567	48.100	2.290
5	1.020	6.002	65.404				
6	0.861	5.066	70.470				
7	0.746	4.386	74.857				
8	0.681	4.008	78.865				
9	0.605	3.556	82.421				
10	0.507	2.981	85.402				
11	0.467	2.748	88.150				
12	0.428	2.519	90.670				
13	0.384	2.262	92.931				
14	0.352	2.072	95.004				
15	0.326	1.921	96.924				
16	0.295	1.736	98.660				
17	0.228	1.340	100.000				

*Anmerkungen.* Extraktionsmethode: Hauptachsen-Faktorenanalyse mit Promax-Rotation.

<sup>a</sup> Wenn Faktoren korreliert sind, können die Summen der quadrierten Ladungen nicht addiert werden, um eine Gesamtvarianz zu erhalten.

Aus der rotierten Summe der quadrierten Ladungen kann gemäss Bühner (2011) mit der Formel „Eigenwert/Anzahl Items x 100“ die aufgeklärte Varianz ausgerechnet werden. Dies ergibt für Faktor 1=24%, Faktor 2=21%, Faktor 3=17%, Faktor 4=13%.



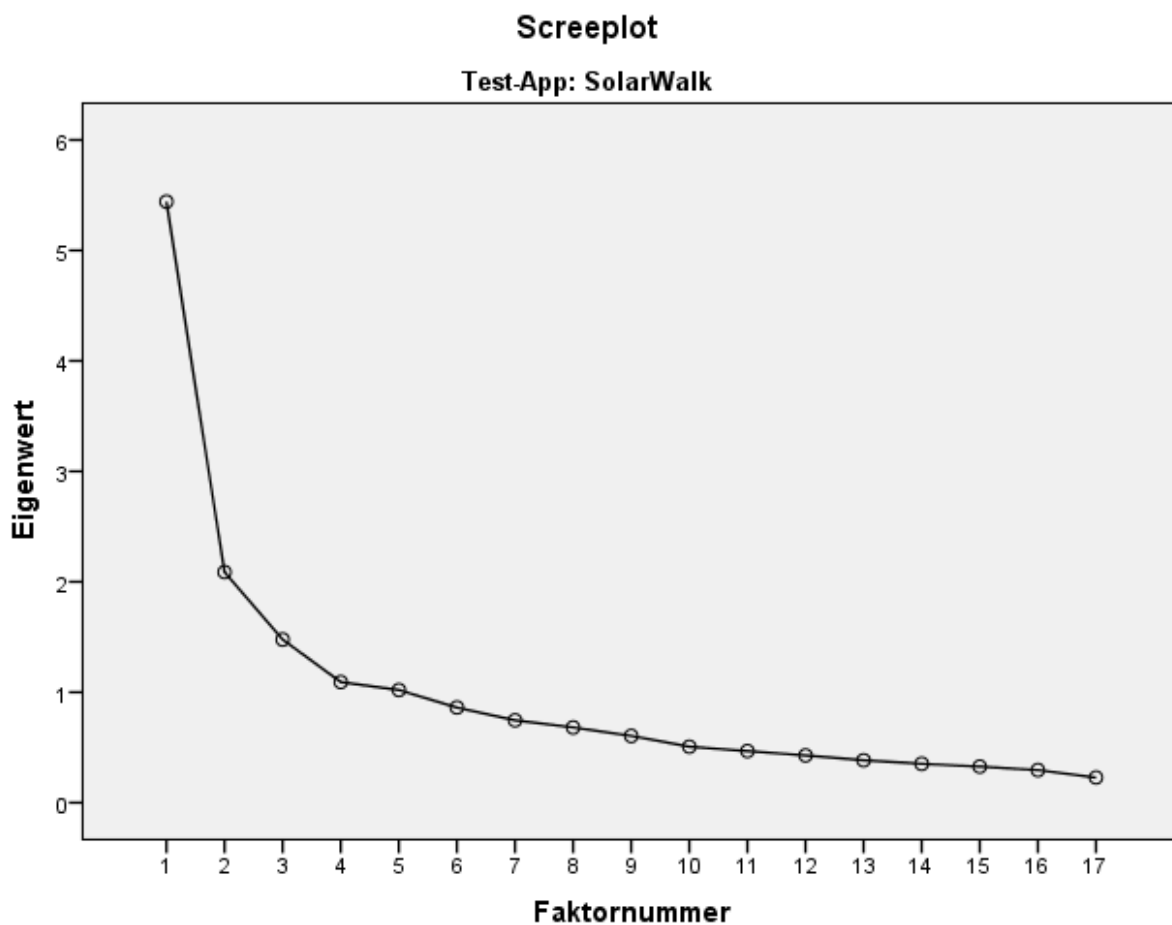


Abbildung 6: Scree-Plot Solar Walk

Tabelle 23: Erklärte Gesamtvarianz, Physics Lite

Faktor	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	Gesamt
1	6.985	41.090	41.090	6.564	38.610	38.610	5.775
2	1.988	11.696	52.786	1.562	9.187	47.797	3.381
3	1.165	6.855	59.641	0.700	4.115	51.912	5.059
4	0.991	5.829	65.469	0.528	3.103	55.015	4.511
5	0.858	5.048	70.517				
6	0.746	4.389	74.906				
7	0.627	3.690	78.596				
8	0.550	3.238	81.834				
9	0.482	2.832	84.667				
10	0.427	2.513	87.180				
11	0.403	2.373	89.552				
12	0.370	2.176	91.729				
13	0.329	1.936	93.665				
14	0.318	1.872	95.537				
15	0.290	1.703	97.240				
16	0.259	1.525	98.765				
17	0.210	1.235	100.000				

*Anmerkungen.* Extraktionsmethode: Hauptachsen-Faktorenanalyse mit Promax-Rotation.

<sup>a</sup> Wenn Faktoren korreliert sind, können die Summen der quadrierten Ladungen nicht addiert werden, um eine Gesamtvarianz zu erhalten.

Die aufgeklärte Varianz aus den rotierten Summe der quadrierten Ladungen ergibt für Faktor 1=34%, Faktor 2=20%, Faktor 3=30%, Faktor 4=26%.

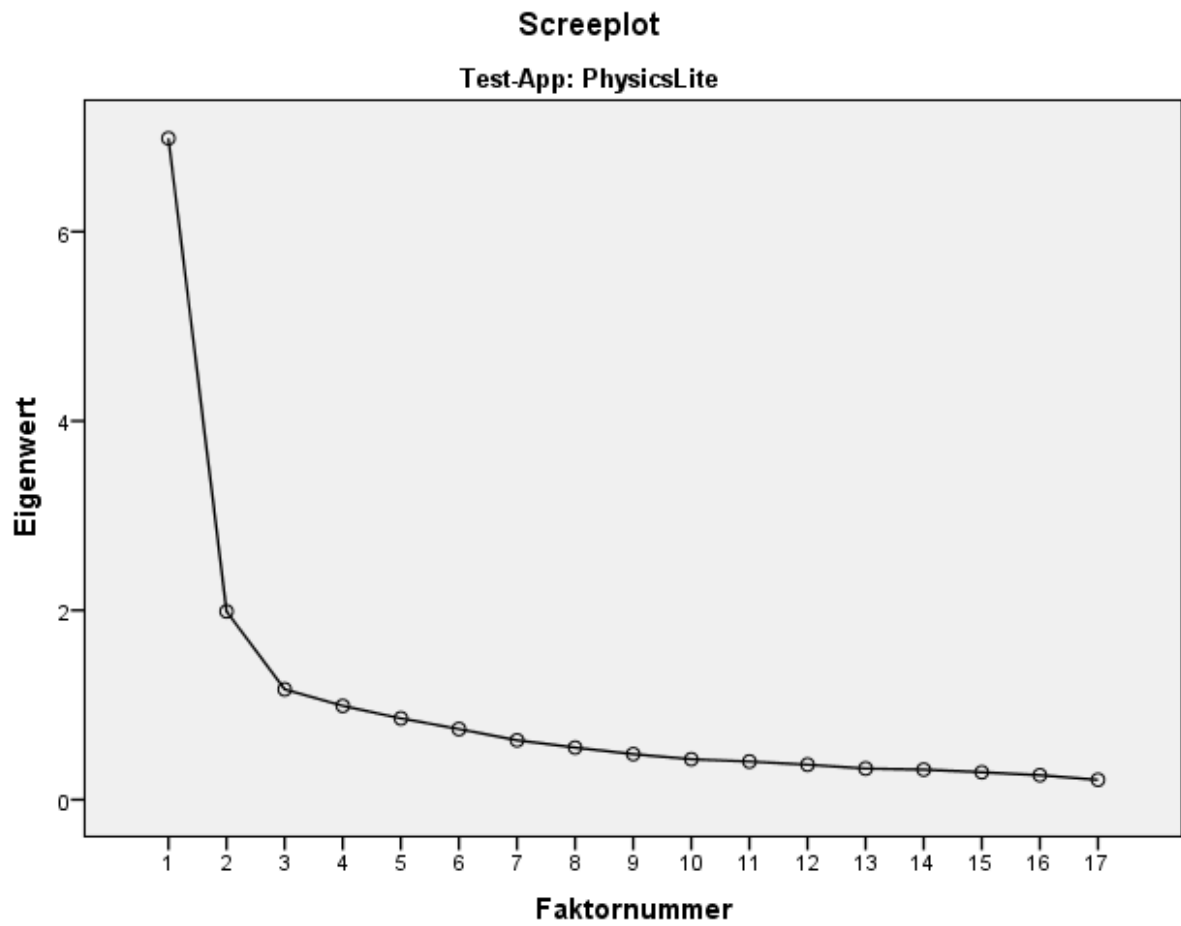


Abbildung 7: Scree-Plot Physics Lite

Die Mustermatrizen (Tabelle 24, Tabelle 25) zeigen die Koeffizienten der Faktorladungen für die einzelnen Items. Zur besseren Verständlichkeit sind die extrahierten Faktoren bereits als Dimensionen bezeichnet. Dieser Schritt wird im Kapitel 4.6.2 erläutert.

**Tabelle 24: Faktorladungen Solar Walk**

Items	Kognitive Stimulation	Emotion	Expressive Ästhetik	Usability
UES.Inv: Ich werde richtig in das Thema reingezogen.	.781			
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	.753			
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	.677	-.281	.261	
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	.621	.239		
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	.501	.365		
ARS.PAtt: Die Benützung der App ist aufregend.	.427	.221		
R-Pan.Emot: Die App verärgert mich.		.749		
R-LBor: Die App langweilt mich.		.681		
AE.Plea.Int: Das Benützen der App macht Spass.		.595	.224	
AE.Expr.Ae: Das Design der App ist faszinierend.			.875	
AE.Expr.Ae: Das Design der App ist originell.			.737	
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.			.346	
Pan.Emot: Die App begeistert mich.		.240	.313	
SPQ.Usa: Die App ist einfach zu gebrauchen.				.779
meCu.Benütz: Die Bedienung der App ist verständlich.				.657
SPQ.Usa: Es ist einfach in der App zu navigieren.				.626
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.				.250

*Anmerkungen.* Extraktionsmethode: Hauptachsenfaktorenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung. Die Rotation ist in Iterationen konvergiert. Werte unter .2 sind nicht angezeigt.

Tabelle 25: Faktorladungen Physics Lite

Items	Kognitive Stimulation	Usability	Emotion	Expressive Ästhetik
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	.919			
UES.Inv: Ich werde richtig in das Thema reingezogen.	.791			
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	.703			
ARS.PAtt: Die Benützung der App ist aufregend.	.594			
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	.591		-.284	.239
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	.591		.319	
SPQ.Usa: Die App ist einfach zu gebrauchen.		.847		
SPQ.Usa: Es ist einfach in der App zu navigieren.		.764		
meCu.Benütz: Die Bedienung der App ist verständlich.		.728		
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.		.551		
R-LBor: Die App langweilt mich.			.760	
R-Pan.Emot: Die App verärgert mich.			.731	
Pan.Emot: Die App begeistert mich.	.354		.457	
AE.Plea.Int: Das Benützen der App macht Spass.	.312		.337	
AE.Expr.Ae: Das Design der App ist faszinierend.				.820
AE.Expr.Ae: Das Design der App ist originell.				.725
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.			.347	.465

*Anmerkungen.* Extraktionsmethode: Hauptachsenfaktorenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung. Die Rotation ist in 5 Iterationen konvergiert. Werte unter .2 sind nicht angezeigt.

### 4.3. Reliabilitätsanalyse

Um die Homogenität der einzelnen Dimensionen zu erfassen wurde eine Reliabilitätsanalyse durchgeführt. Tabelle 26 zeigt die Reliabilitätsstatistik der Items.

Tabelle 26: Reliabilitätsstatistik

Dimensionen und Items	Solar Walk			Physics Lite		
	Cronbachs Alpha	Trennschärfe	Cronbachs Alpha, wenn Item weggelassen	Cronbachs Alpha	Trennschärfe	Cronbachs Alpha, wenn Item weggelassen
<b>Gesamtskala*</b>	.862			.909		
<b>Expressive Ästhetik</b>	.680			.806		
AE.Expr.Ae: Das Design der App ist faszinierend.		.574	.489		.690	.704
AE.Expr.Ae: Das Design der App ist originell.		.543	.519		.660	.727
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.		.380	.724		.624	.776
<b>Emotion</b>	.787			.772		
AE.Plea.Int: Das Benützen der App macht Spass.		.704	.674		.563	.722
LBor: Die App langweilt mich.		.603	.731		.553	.728
Pan.Emot: Die App begeistert mich.		.501	.782		.688	.658
R-Pan.Emot: Die App verärgert mich.		.598	.742		.500	.755
<b>Usability</b>	.610			.811		
SPQ.Usa: Die App ist einfach zu gebrauchen.		.540	.455		.716	.720
SPQ.Usa: Es ist einfach in der App zu navigieren.		.444	.499		.689	.734
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.		.207	.735		.456	.843
meCu.Benütz: Die Bedienung der App ist verständlich.		.494	.479		.670	.743
<b>Kognitive Stimulation</b>	.831			.864		
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.		.499	.831		.477	.873
UES.Inv: Ich werde richtig in das Thema reingezogen.		.688	.789		.729	.828
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).		.632	.798		.675	.838
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.		.623	.801		.740	.826
ARS.PAtt: Die Benützung der App ist aufregend.		.582	.808		.661	.841
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.		.637	.796		.676	.838

*Anmerkungen.* Reliabilität Gesamtskala\* ohne Item „Das Benützen der App ist verständlich“.

## Faktor-Korrelationen

Für die Analyse der Faktor-Korrelationen wurde mit den Summenwerten der Faktoren (Dimensionen) gerechnet, da mit der regressionsanalytischen Schätzung der Faktorwerten nach Fromm (2004) die Unabhängigkeit der einzelnen Faktoren erzwungen wird. Da die Daten nicht normalverteilt sind, entspricht der Korrelationskoeffizient dem Spearman-Rho. Tabelle 27 und Tabelle 28 zeigen die Faktorenkorrelationen der beiden Apps.

**Tabelle 27: Faktorenkorrelationen Solar Walk**

		<b>Kognitive Stimulation</b>	<b>Expressive Ästhetik</b>	<b>Emotion</b>	<b>Usability</b>
<b>Kognitive Stimulation</b>	Korrelationskoeffizient	–			
	Signifikanz (2-seitig)	–			
	<i>N</i>	–			
<b>Expressive Ästhetik</b>	Korrelationskoeffizient	.411**			
	Signifikanz (2-seitig)	.000			
	<i>N</i>	158			
<b>Emotion</b>	Korrelationskoeffizient	.622**	.433**		
	Signifikanz (2-seitig)	.000	.000		
	<i>N</i>	158	158		
<b>Usability</b>	Korrelationskoeffizient	.092	.214**	.269**	–
	Signifikanz (2-seitig)	.252	.007	.001	–
	<i>N</i>	158	158	158	–

*Anmerkungen.* Korrelationskoeffizient=Spearman-Rho, *N*=Stichprobe  
 \*\*. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Tabelle 28: Faktorenkorrelationen Physics Lite

		Kognitive Stimulation	Expressive Ästhetik	Emotion	Usability
<b>Kognitive Stimulation</b>	Korrelationskoeffizient	–			
	Signifikanz (2-seitig)	–			
	<i>N</i>	–			
<b>Expressive Ästhetik</b>	Korrelationskoeffizient	.562**			
	Signifikanz (2-seitig)	.000			
	<i>N</i>	157			
<b>Emotion</b>	Korrelationskoeffizient	.637**	.632**		
	Signifikanz (2-seitig)	.000	.000		
	<i>N</i>	157	157		
<b>Usability</b>	Korrelationskoeffizient	.315**	.314**	.348**	–
	Signifikanz (2-seitig)	.000	.000	.000	–
	<i>N</i>	157	157	157	–

*Anmerkungen.* Korrelationskoeffizient=Spearman-Rho, *N*=Stichprobe  
 \*\*. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

#### 4.4. Mittelwertvergleiche

Der Wilcoxon-Test für abhängige und nicht normalverteilte Stichproben zeigt beim Mittelwertvergleich zwischen den zwei verschiedenen Test-Apps bei 21 Items signifikante Unterschiede. Nur bei Item 11 „Die Benützung der App ist anspruchsvoll.“ ist keine Signifikanz vorhanden. Die Testergebnisse finden sich im Anhang.

#### Vergleich der Apps

Abbildung 8 zeigt die Mittelwerte der einzelnen Dimensionen sowie den zwei Items „Ich werde diese App weiter empfehlen“ und „Mit der App kann ich die Lernaufgabe lösen“. Die App Solar Walk erzielt dabei auf allen Variablen höhere Werte als Physics Lite.



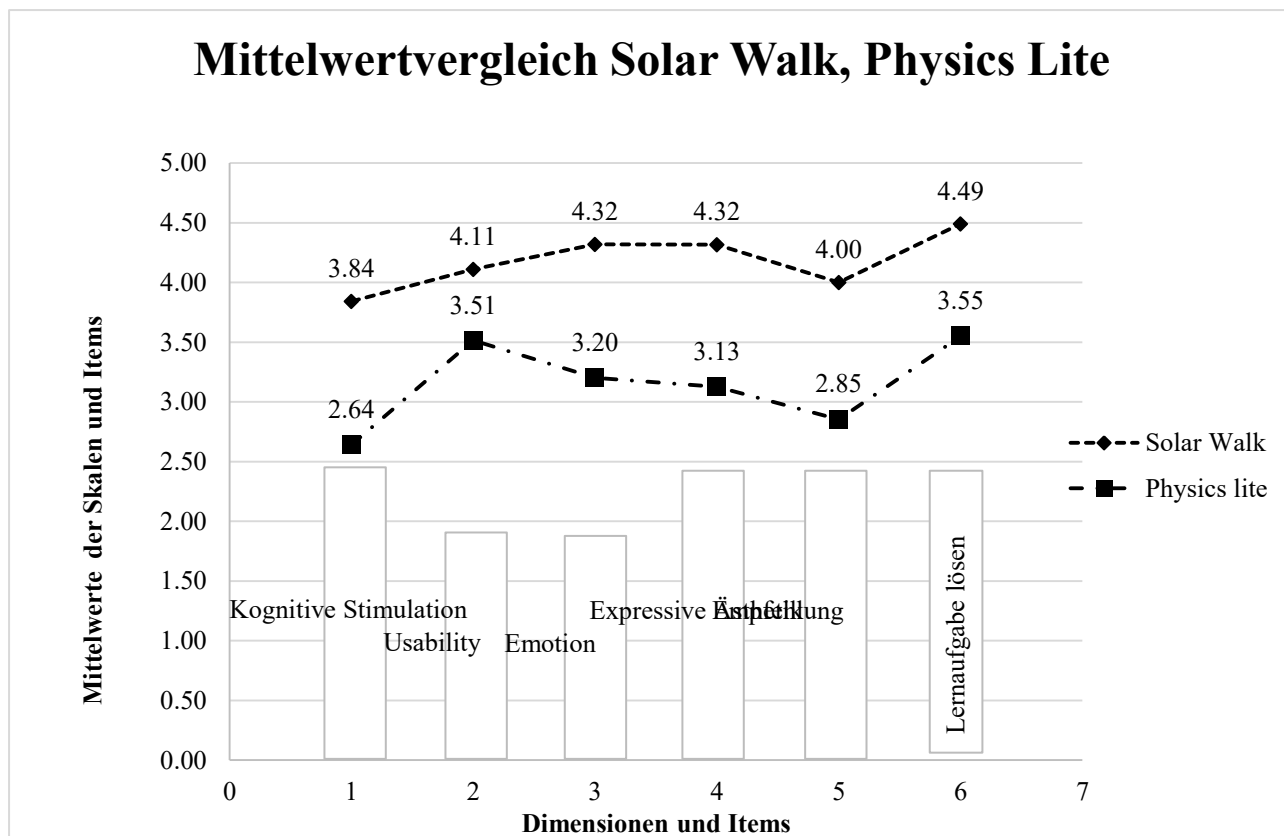


Abbildung 8: Mittelwertvergleich Solar Walk, Physics Lite

### Vergleich der Geschlechter

Mit einem T-Test für unabhängige Stichproben wurde getestet, ob sich das Antwortverhalten unter den Geschlechtern signifikant voneinander unterscheidet. Obwohl es sich um nicht normalverteilte Daten handelt, ist der T-Test für eine Stichprobe < 60 unproblematisch (Universität Zürich, 2016).

Die Tabellen befinden sich im Anhang. Bei Solar Walk unterscheiden sich „Das Design der App ist faszinierend“, „Das Design der App ist originell“ und „Ich werde die App weiter empfehlen“. Bei allen drei Items ist der Mittelwert der weiblichen Testpersonen höher. Bei Physics Lite unterscheiden sich „Das Design der App ist faszinierend“ und „Die App macht mich neugierig auf das Thema (Physik)“. Auch hier liegt der Mittelwert bei den weiblichen Testperso-

nen höher, wobei die Effekte jedoch gering sind. Die Effektgrösse ist mit dem Korrelationskoeffizienten  $r$  berechnet (Formel siehe Anhang).

Bei den Variablen „App-Nutzung pro Woche“ und der Leistung in den Lernaufgaben gibt es keine signifikanten Unterschiede zwischen den Geschlechtern (siehe Anhang).

#### 4.5. Korrelationen

Tabelle 29 zeigt die Korrelationen zwischen den vier Variablen: Lernaufgabe (richtige Antworten in der Lernaufgabe), Schulniveau (Niveau) und den Items „Mit der App kann ich die Lernaufgabe lösen“ sowie „Das Thema interessiert mich“. Bei der App Solar Walk ergeben sich keine signifikanten Korrelationen, bei der App Physics Lite geringe jedoch signifikante Korrelationen. Da es sich um verschiedene Antwortskalen handelt, sind die Variablen z-standardisiert.

**Tabelle 29: Korrelationen**

<b>App</b>			<b>z-Faktorwert: Niveau</b>	<b>z-Faktorwert: Lernaufgabe</b>	<b>z-Faktorwert: Mit der App kann ich die Lernaufgabe lösen.</b>	<b>z-Faktorwert: Das Thema interessiert mich.</b>
SolarWalk	<b>z-Faktorwert: Niveau</b>	Korrelationskoeffizient	-			
		Sig. (2-seitig)	-			
		<i>N</i>	-			
	<b>z-Faktorwert: Lernaufgabe</b>	Korrelationskoeffizient	.155			
		Sig. (2-seitig)	0.052			
		<i>N</i>	158			
	<b>z-Faktorwert: Mit der App kann ich die Lernaufgabe lösen.</b>	Korrelationskoeffizient	0.131	.100		
		Sig. (2-seitig)	0.104	0.216		
		<i>N</i>	156	156		
	<b>z-Faktorwert: Das Thema interessiert mich.</b>	Korrelationskoeffizient	0.039	0.033	.130	-
		Sig. (2-seitig)	0.626	0.680	0.106	-
		<i>N</i>	157	157	155	-
Physics Lite	<b>z-Faktorwert: Niveau</b>	Korrelationskoeffizient	-			
		Sig. (2-seitig)	-			
		<i>N</i>	-			
	<b>z-Faktorwert: Lernaufgabe</b>	Korrelationskoeffizient	.254**			
		Sig. (2-seitig)	0.001			
		<i>N</i>	158			
	<b>z-Faktorwert: Mit der App kann ich die Lernaufgabe lösen.</b>	Korrelationskoeffizient	0.146	.197*		
		Sig. (2-seitig)	0.070	0.014		
		<i>N</i>	155	155		
	<b>z-Faktorwert: Das Thema interessiert mich.</b>	Korrelationskoeffizient	.305**	0.123	.306**	-
		Sig. (2-seitig)	0.000	0.128	0.000	-
		<i>N</i>	156	156	154	-

*Anmerkungen.* Korrelationskoeffizient=Spearman-Rho, *N*=Stichprobe

\*\* . Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

\* . Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

#### 4.6. Erläuterungen zu den Ergebnissen

In den vorherigen Abschnitten sind die statistischen Werte der verschiedenen Items und weiteren Variablen der Untersuchung wiedergegeben. Nachfolgend sind einige Vorgehensschritte vertiefter erläutert.

##### 4.6.1. Itemauswahl für die Faktorenanalyse

Mit dem konstruierten Fragebogen soll das ästhetische Empfinden erfasst werden. Die Faktorenanalyse überprüft unter anderem, welche latenten Variablen dem Antwortverhalten zu Grunde liegen. Bei der Konzeption wurden auch Items verwendet, die voraussichtlich nicht mit ästhetischer Empfindung in Zusammenhang stehen würden, aber möglicherweise als Kontrollvariablen oder für zusätzliche Informationen gute Dienste leisten könnten. Die statistische Überprüfung konnte diese Annahme teilweise bestätigen. Im folgenden Abschnitt wird der Ausschluss der nicht verwendeten Items begründet.

##### **Item Nr. 7. Ich werde diese App weiter empfehlen.**

Dieses Item ist eine Kontrollvariable im Sinne einer Gesamtbeurteilung der App. Das Item liess sich in der vierfaktoriellen Variante der Faktorenanalyse nicht eindeutig einem Faktor zuordnen.

##### **Item Nr. 12. Mit der App kann ich die Lernaufgabe lösen.**

Dieses Item dient als Kontrollvariable im Sinne einer Selbstbeurteilung der Testperson zu den Resultaten der Lernaufgaben.

##### **Item Nr. 14. Der Ton passt zum Design.**

Das Item liess sich in der vierfaktoriellen Variante der Faktorenanalyse nicht eindeutig einem Faktor zuordnen.

##### **Item Nr.15. Der Ton stört mich.**

Das Item liess sich in der vierfaktoriellen Variante der Faktorenanalyse nicht eindeutig einem Faktor zuordnen.

**Item Nr. 18. Das Thema interessiert mich.**

Dieses Item dient als Kontrollvariable zur Erfragung des Interesses für das bearbeitete Thema.

**4.6.2. Benennung der Faktoren als Dimensionen**

Die Faktorenanalyse gibt nur aufgrund der statistischen Überprüfung der Items verschiedene Faktoren aus, die als latente Variablen das Antwortverhalten erklären können. Deshalb ist auch die Reihenfolge der vier Faktoren in den Mustermatrizen unterschiedlich. Die definitive Zuordnung der Items und die Benennung der Dimensionen berücksichtigen neben den statistischen Werten auch theoretische und inhaltliche Überlegungen.

Die Mustermatrizen bei beiden Apps zeigen, dass einzelne Items auf zwei Faktoren laden. In diesem Fall konnten die Matrizen miteinander verglichen werden. Für die Zuordnung gaben schliesslich die höhere Faktorladung sowie inhaltliche Überlegungen den Ausschlag. Die Dimensionen „Usability“, und „Expressive Ästhetik“ entsprechen mehrheitlich den Originalskalen, deshalb wurden auch deren Bezeichnungen übernommen.

Der Dimension „Emotion“ wurden die Items aus „Learning related boredom“ und „Pleasurable Interaction“ zugeordnet.

Die Dimension „Kognitive Stimulation“ ist in dieser Konstellation neu. Sie setzt sich zusammen aus „Cognitive Stimulation“, „Involvement“, „Intrinsic motivation“ und „Positive Attraction“. Die Bezeichnung „Kognitive Stimulation“ wurde deshalb gewählt, weil sie alle Teilbereiche als Oberbegriff gut repräsentiert. Diese Dimension ist aber von der Originalversion „Cognitive Stimulation“ aus Hager et al. (2012) deutlich zu unterscheiden.

**4.7. Zusammenfassung der Ergebnisse**

Die deskriptive Statistik zeigt beim Vergleich der verschiedenen Test-Apps, dass die App Solar Walk einige Items mit schiefen Verteilungen generiert. Bei Item „Das Design der App ist faszinierend“ wurden die Antwortkategorien 1 („stimme gar nicht zu“) und 2 („-“) nie ange-

kreuzt, zusätzlich ist das Item linkschief. Die Items „Das Design der App ist originell“, „Das Benützen der App macht Spass“, „Die App langweilt mich“, „Mit der App kann ich die Lernaufgabe lösen“, „Die Bedienung der App ist verständlich“, „Der Ton passt zum Design“, „Der Ton stört mich“ und „Die App verärgert mich“ sind alle linkschief verteilt. Ausserdem fehlt Antwortkategorie 1 („stimme gar nicht zu“) bei „Die App ist einfach zu gebrauchen“, „Die Spezialeffekte in der App sind gut gemacht“, „Die App langweilt mich“ und „Die App verärgert mich“.

Bei der App Physics Lite ist das Item „Bei der Benützung der App vergesse ich meine Umgebung“ rechtsschief verteilt.

Die Mittelwerte sind bei Solar Walk deutlich höher als bei Physics Lite. Alle Items sind nicht normalverteilt, was bei weiteren statistischen Verfahren berücksichtigt wurde.

Aufgrund der Mittelwerte, Verteilungsformen und der Auswahl der Antwortkategorien weist kein Item eine Auffälligkeit auf, die eine Entfernung aus dem Fragebogen rechtfertigen würde.

Aus den Ergebnissen der weiteren Variablen ist ersichtlich, dass rund 60 Prozent der Testpersonen nie eine Lern-App innerhalb einer Woche benützen und 26 Prozent einmal. Die Lernaufgaben wurden bei der App Solar Walk zu 84 Prozent richtig gelöst. Bei Physics Lite sind es 56 Prozent.

Bei der Faktorenanalyse ergibt die Überprüfung auf Dateneignung beim Kaiser-Meyer-Olkin-Mass gute bis sehr gute Werte und der Bartlett-Test ist höchst signifikant. Die MSA-Koeffizienten sind mit einer Faktorenanalyse kompatibel.

Das Kaiser-Guttman-Kriterium (Eigenwerte  $>1$ ) weist bei Solar Walk auf fünf, bei Physics Lite auf vier Faktoren hin. Der Knick beim Scree-Plot spricht bei Solar Walk eher für vier Faktoren, bei Physics Lite für drei. Die definitive Lösung beinhaltet nach dem Ausschluss von fünf Items vier Faktoren, die bei Solar Walk 59 Prozent der Gesamtvarianz der anfänglichen Ei-

genwerten aufklären und bei Physics Lite 60 Prozent. Die Koeffizienten der Faktorladungen sind bei den Apps unterschiedlich gross und laden nicht immer deutlich auf nur einen Faktor. Trotzdem können vier Faktoren extrahiert werden, die das ästhetische Empfinden erfassen. Sie entsprechen den Dimensionen „Kognitive Stimulation“, „Expressive Ästhetik“, „Emotion“ und „Usability“. Insofern bestätigt die Statistik, dass es sich um einen mehrdimensionalen Fragebogen handelt. Die Dimensionen korrelieren bei Solar Walk mit Ausnahmen von „Usability“ zu „Kognitiver Stimulation“ hoch signifikant zwischen .21 und .62. Bei Physics Lite korrelieren alle Dimensionen hoch signifikant zwischen .21 und .63.

Die Reliabilitätsanalyse dieser Dimensionen zeigen akzeptable Werte für Cronbachs Alpha (.68 bis .86) und mittlere bis hohe Trennschärfen (>.30). Nur bei Item „Die Benützung der App ist anspruchsvoll“ ist die Trennschärfe von .21 bei Solar Walk niedrig. Die Entfernung dieses Items aus der Skala erhöht den Wert für Cronbachs Alpha bei Solar Walk auf .74 und bei Physics Lite auf .84. Bei der Berechnung der Gesamtskala wurde dieses Item bereits entfernt. Cronbachs Alpha beträgt für die gesamte Skala „ästhetisches Empfinden“ bei Solar Walk .82 und bei Physics Lite .91.

Der Mittelwertvergleich zwischen den vier Dimensionen, sowie den Items „Ich werde diese App weiter empfehlen“ und „Mit der App kann ich die Lernaufgabe lösen“ zeigt deutliche Unterschiede. Das App Solar Walk wird auf allen verglichenen Variablen höher bewertet als Physics Lite.

Der Geschlechtervergleich zeigt bei insgesamt fünf Items signifikante, jedoch geringe Unterschiede mit kleinen Effekten an. Hinsichtlich der Resultate in den Lernaufgaben und der „App Nutzung pro Woche“ unterscheiden sich die Geschlechter nicht signifikant.

Die Korrelations-Berechnung der Variablen Lernaufgabe (richtige Antworten in der Lernaufgabe), Schulniveau (Niveau) und den Items „Mit der App kann ich die Lernaufgabe lösen“

sowie „Das Thema interessiert mich“ ergeben bei der App Physics Lite geringe jedoch signifikante Korrelationen.

#### 4.8. Anpassung des Fragebogen nach der statistischen Überprüfung

Aufgrund der statistischen Werte ist der Ausschluss des Items „Die Benützung der App ist anspruchsvoll“ angezeigt. Die Kommunalität ( $h^2=.11 - .26$ ), der MSA-Koeffizient (.40), die Trennschärfe (.21 - .46) und bei Solar Walk die Faktorladung (.27) sind auffällig niedrig. Beide selbst konstruierten Items zu Ton lassen sich nicht sinnvoll einer der vier Dimensionen zuordnen und werden ebenfalls ausgeschlossen. Die Items „Ich werde diese App weiter empfehlen“, „Mit der App kann ich die Lernaufgabe lösen“, „Das Thema interessiert mich“ bleiben bestehen. Sie können zwar nicht in eine Dimension integriert werden, generieren aber als Kontrollvariablen weitere Daten zur Beurteilung einer App. Tabelle 30 zeigt den angepassten Fragebogen nach der statistischen Überprüfung. Die ausgeschlossenen Items sind in der linken Spalte durchgestrichen. Eine weitere Version des definitiven Fragebogens mit anderer Reihenfolge der Items ist im Anhang eingefügt.



Tabelle 30: Fragebogen nach der statistischen Überprüfung

vor der statistischen Überprüfung	nach der statistischen Überprüfung
	<b>Ästhetisches Empfinden (<math>\alpha=.86-.91</math>)</b>
	<b>Dimension Expressive Ästhetik (<math>\alpha=.68-.86</math>)</b>
1. AE.Expr.Ae: Das Design der App ist faszinierend.	Das Design der App ist faszinierend.
2. AE.Expr.Ae: Das Design der App ist originell.	Das Design der App ist originell.
3. AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	Die Spezialeffekte in der App sind gut gemacht.
4. AE.Plea.Int: Das Benützen der App macht Spass.	<b>Dimension Usability (<math>\alpha=.74-.84</math>)</b>
5. SPQ.Usa: Die App ist einfach zu gebrauchen.	Die App ist einfach zu gebrauchen.
6. SPQ.Usa: Es ist einfach in der App zu navigieren.	Es ist einfach in der App zu navigieren.
7. SPQ.Loy: Ich werde diese App weiter empfehlen.	Die Bedienung der App ist verständlich.
8. UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	<b>Dimension Emotion (<math>\alpha=.77-.79</math>)</b>
9. UES.Inv: Ich werde richtig in das Thema reingezogen.	Das Benützen der App macht Spass.
10. LBo: Die App langweilt mich.	Die App langweilt mich.
<del>11. UES.Pere.Us: Die Benützung der App ist anspruchsvoll.</del>	Die App begeistert mich.
12. meCu.Nütz: Mit der App kann ich die Lernaufgabe lösen.	Die App verärgert mich.
13. meCu.Benütz: Die Bedienung der App ist verständlich.	<b>Dimension Kognitive Stimulation (<math>\alpha=.83-.86</math>)</b>
<del>14. Ton: Der Ton passt zum Design.</del>	Bei der Benützung der App vergesse ich meine Umgebung.
<del>15. Ton: Der Ton stört mich.</del>	Ich werde richtig in das Thema reingezogen.
16. In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	Die App macht mich neugierig auf das Thema (xxxxx).
17. In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	Ich möchte das Thema am liebsten mit der App weiter erkunden.
18. In.Mot: Das Thema interessiert mich.	Das Benützen der App regt meine Gedanken zum Thema (xxxxx) an.
19. Pan.Emot: Die App begeistert mich.	Die Benützung der App ist aufregend.
20. Pan.Emot: Die App verärgert mich.	<b>Kontrollvariablen</b>
21. ARS.PAtt: Die Benützung der App ist aufregend.	Ich werde diese App weiter empfehlen.
22. ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	Mit der App kann ich die Lernaufgabe lösen.
	Das Thema interessiert mich.

Anmerkungen. (xxxxx) = Thema, das in der App behandelt wird. Die durchgestrichenen Items werden in der angepassten Fassung nicht mehr verwendet.

## 5. Diskussion

In diesem Kapitel werden Theorie, Methoden und Ergebnisse diskutiert und abgeleitet aus den Erkenntnissen auf Implikationen für die Praxis hingewiesen.

### 5.1. Theorie

Zur Beantwortung der Fragestellung „Wie kann das ästhetische Empfinden von Jugendlichen bei der Verwendung digitaler Lernmedien (Applikation) erfasst werden?“ wird als Grundlage für die Konstruktion des Fragebogens ein Arbeitsmodell zu ästhetischem Empfinden entwickelt. Das Arbeitsmodell integriert die für diese Untersuchung relevanten Konstrukten aus verschiedenen Theorien, Modellen und empirischen Untersuchungen. Hauptreferenzen sind das Modell von Leder et al. (2004), die Arbeiten von Thielsch (2008) sowie Modelle und Konstrukte aus dem Lernkontext (vgl. Kapitel 2.3).

Nach der Untersuchung könnte man das Arbeitsmodell mit folgenden Worten zusammenfassen: Ästhetisches Empfinden ist ein Konstrukt, das sich aufgrund von Wahrnehmungs- und Bewertungsprozessen zwischen den Gestaltungsmerkmalen der App (Inhalt, Usability, Ästhetik) und der Betrachterin oder dem Betrachter einstellt. Ästhetisches Empfinden setzt sich aus den Dimensionen „Kognitive Stimulation“, „Expressive Ästhetik“, „Emotion“ und „Usability“ zusammen. Es ist hier aber zu berücksichtigen, dass nicht das Arbeitsmodell empirisch überprüft wurde, sondern nur der Fragebogen. Das Arbeitsmodell könnte aber als Grundlage für weitere Forschungen in Bereich Interface-Ästhetik dienen.

Beim Arbeitsmodell und dem Fragebogen werden allerdings Konstrukte wie soziale Faktoren (sozialer Austausch, soziale Meinungsbildung, Gruppendynamiken) Kollaboration (Kollaboratives Lernen) und weitere Personenvariablen (ästhetische Sensitivität, Persönlichkeitsstruktur) nicht berücksichtigt, was Gegenstände zukünftiger Forschungen sein könnte, um das Arbeitsmodell weiter zu verfeinern.

Das Arbeitsmodell beschreibt eine Definition zu ästhetischem Empfinden und gibt Hinweise darauf, dass die Gestaltungsmerkmale der App wie Inhalt, Usability und Ästhetik die ästhetische Empfindung beeinflussen. Insbesondere die Gestaltung des Inhalts ist bei Lernmedien von grosser Bedeutung (vgl. Kapitel 2.3) und erweitert ein Verständnis von Ästhetik, das sich nicht nur auf klassische Designmerkmale wie Farbigekeit, Form, Komplexität, Schriftgrössen, Schriftarten, Usability etc. abstützt.

Wird die Applikation als vermittelndes oder unterstützendes Medium in einer Ausstellung verwendet, so stellt sich die Frage, ob nicht noch weitere ästhetische Bewertungsprozesse im gesamten Kontext stattfinden (vgl. Kontextvariablen bei Leder et al., 2004). Ästhetik und ästhetische Wahrnehmung können schon bei der Art und Weise, wie die Applikation auf reale Objekte, auf die Umgebung und auf die User Bezug nimmt, eine Rolle spielen und eröffnen weitere Fragen, wie zum Beispiel: Bezieht sich die Applikation auf einen Gegenstand oder die ganze Ausstellung? Wie bezieht sich die App auf den Raum (Aussenraum, Innenraum)? Dient die App als mobile Datenbank, als Kollaborationsplattform, als Kontrollinstrument oder als „virtueller Führer“? Wie wird vorhandenes und neues Wissen in der Applikation rekonstruiert und dargestellt? Reduziert die Applikation die Realität oder bereichert sie sie im Sinne von „Augmented Reality“. Solche Fragen legen ein noch erweitertes Konzept ästhetischen Empfindens und entsprechend von Interface-Ästhetik nahe und sie können somit Gegenstand weiterer Forschungen sein.

## 5.2. Methoden

Für die Fragestellung „Wie kann das ästhetische Empfinden von Jugendlichen bei der Verwendung digitaler Lernmedien (Applikation) erfasst werden?“ wurden hauptsächlich zwei Methoden angewendet. Mittels Literaturstudium erfolgten die Formulierung eines Arbeitsmodells sowie die Auswahl an bereits vorhandenen Messinstrumenten, die zur Konstruktion des eigentlichen Fragebogens führten. Mit einer quantitativen Untersuchung wurde der Fragebogen sta-

tistisch überprüft. Neben den üblichen Verfahren wie Deskriptive Statistik, Korrelationen, Reliabilitäten, Mittelwertvergleiche wurde eine Explorative Faktorenanalyse (Hauptachsenanalyse, Promax) zur Dimensionsüberprüfung durchgeführt.

Die Beurteilung von zwei verschiedenen Apps pro Testperson hat sich als hilfreich erwiesen. Somit standen immer zwei Datensätze zur Verfügung, die miteinander verglichen und bei der Interpretation berücksichtigt werden konnten. Für eine erste Validierung des Fragebogens waren diese Methoden ausreichend und führten zu einer Anpassung des Fragebogens, indem Items mit ungenügenden statistischen Werten ausgeschlossen wurden. Die Stichprobengröße von ( $N=158$ ) hat sich als genügend gross erwiesen. Allerdings kam es zu schiefen Itemverteilungen, die möglicherweise bei einer grösseren Stichprobe nicht oder vermindert aufgetreten wären. Für eine weitere Überprüfung könnte die Anwendung einer konfirmatorischen Faktorenanalyse mit AMOS zusätzliche Informationen liefern, inwieweit sich die Dimensionen bestätigen lassen. Weiter würden sich zur Überprüfung von konvergenter sowie diskriminanter Validität Korrelationsberechnungen mit Tests gleicher oder anderer Gültigkeitsbereiche anbieten.

Ferner wurde der Fragebogen nicht in einem ausserschulischen Lernort (Museum, Ausstellung) und mit einer für diese Situation programmierten Applikation getestet. Für eine erste Überprüfung des Fragebogens ist die angewandte Untersuchung im schulischen Kontext ausreichend in Anbetracht dessen, dass weitere Validierungen folgen werden.

Zur Überprüfung des Arbeitsmodells, das zwar nicht Gegenstand dieser Untersuchung war, könnten Pfadanalysen die Wirkungszusammenhänge der einzelnen Konstrukte überprüfen. Bei der Identifikation von Gestaltungsmerkmalen für das Interface-Design, das besonders die Bedürfnisse der Anwender (Jugendliche) berücksichtigt, bieten sich als Erweiterung der experimentellen Untersuchungen qualitative Forschungsmethoden an, wie sie in der Arbeit von Lorange und Nielsen (2013) angewendet wurden.

### 5.3. Ergebnisse

Aus der Fragestellung „Wie kann das ästhetische Empfinden von Jugendlichen bei der Verwendung digitaler Lernmedien (Applikation) erfasst werden?“ resultiert der neu konstruierte und nach der ersten Überprüfung angepasste Fragebogen zur Erfassung des ästhetischen Empfindens (vgl. Kapitel 4.8). Das Messinstrument umfasst die vier Dimensionen „Kognitive Stimulation“ (6 Items), „Expressive Ästhetik“ (3 Items), „Emotion“ (4 Items) und „Usability“ (3 Items). Drei weitere Items bilden eine Skala mit Kontrollvariablen, die nicht der ästhetischen Empfindung zugeordnet werden. Das Item „Ich werde diese App weiter empfehlen“ kann im Sinne einer Gesamtbeurteilung der App als zusätzliche Datenquelle dienen. Die Items „Mit der App kann ich die Lernaufgabe lösen“ und „Das Thema interessiert mich“ sind in einer Testsituation sinnvoll, in der als Output des Tests eine messbare Lernleistung resultiert. Mit beiden Items kann die objektive Lernleistung mit Variablen der Selbsteinschätzung validiert werden. Aus den 22 ursprünglichen Items wurden drei Items entfernt. Beide Apps zu Ton „Der Ton passt zum Design“ und „Der Ton stört mich“ und „Die Benützung der App ist anspruchsvoll“. Bei letzterem ist die Eliminierung aufgrund niedriger statistischer Werte zu begründen. Zusätzlich ist es redundant mit dem Item „Die Bedienung der App ist verständlich“. Weiter ist das Item „Die Spezialeffekte in der App sind gut gemacht“ etwas kritisch, weil es das Vorhandensein von Spezialeffekten in einer App voraussetzt sowie die Identifikation von Spezialeffekten, falls vorhanden, erfordert. Die Problematik zeigt sich in Solar Walk, wo nicht alle Antwortkategorien benützt wurden und die Verteilung sehr linkschief ist, was wiederum Auswirkungen auf die Faktorenanalyse hat.

Die Items zu Ton laden einerseits nicht eindeutig auf einer der extrahierten Faktoren, andererseits bildeten sie auch bei den vorausgehenden Berechnungsmethoden keine deutliche, eigene Dimension. Ein Grund dafür ist möglicherweise die inkonsistente Formulierung wobei das eine Item den Ton mit dem Design verknüpft und das andere den Ton als störendes Element etiket-

tiert. Für eine genauere Analyse zu Ton wäre die Konstruktion einer eigenen Dimension mit mindestens vier Items sinnvoller. Bei der Gestaltung einer App könnte der Ton dazu dienen, positive Emotionen auszulösen oder kognitive Schemata in Bezug auf das Thema zu aktivieren, die die Integration von neuem Wissen erleichtern könnten (vgl. Kapitel 2.1). Die Histogramme zu den Ton-Items zeigen, dass die sphärischen Klänge bei Solar Walk durchwegs positiv beurteilt werden, auch bei Physics Lite überwiegen die positiven Antworten, jedoch nicht so deutlich wie bei Solar Walk. Zu berücksichtigen ist auch, dass bei Solar Walk der Ton stetig als Hintergrundgeräusch läuft und bei Physics Lite der Ton nur bei einigen Animationen als zusätzliches Element der Animation eingesetzt wird, wo es Sinn macht. Bei beiden Apps kann der Ton auch abgeschaltet werden.

Etwas problematisch sind die schiefen Verteilungen einzelner Items, vor allem bei Solar Walk. Gemäss Bühner (2011) erzeugen unterschiedlich verteilte Items schwach monotone Zusammenhänge und führen dazu, dass die Items nicht maximal miteinander korrelieren können. Weiter sollten pro Faktor mindestens vier Items vorliegen, was bei der Dimension „Expressive Ästhetik“ nicht der Fall ist. Schiefe Verteilungen können nach Bühner (2011) unter anderem auch zu niedriger Kommunalität ( $h^2 < 0.60$ ) führen, was auch als Mass für die Reliabilität eines Items angesehen werden kann. Insgesamt unterschreiten einige Items diesen Wert. Niedrige Kommunalitäten und wenige Items pro Faktor können nach Bühner (2011) durch die Grösse der Stichprobe kompensiert werden. Für diesen Fall gibt er eine Stichprobengrösse von mindestens 300 Versuchspersonen an, was in dieser Untersuchung nicht erreicht wurde. Für  $h^2 \approx 0.5$  empfiehlt Bühner (2011) eine Stichprobengrösse von 100-200 Probanden.

Da die Kriterien zur Durchführbarkeit einer Faktorenanalyse insgesamt alle erfüllt sind, werden nur aufgrund der niedrigen Kommunalitäten keine Items ausgeschlossen, jedoch die niedrigen Werte bei der Interpretation berücksichtigt.

Bei der Bildung der Dimensionen werden die Faktorladungen der Items aus den Mustermatrizen interpretiert. Wie schon erwähnt, laden nicht alle Items ausschliesslich auf einen Faktor und die Faktorladung ist nicht überall höher als 0.4 (vgl. Kap. 3.6.13). Insbesondere das Item „Bei der Benützung der App vergesse ich meine Umgebung“ lädt hauptsächlich auf „Kognitive Stimulation“ und nebenbei auf „Emotion“ und „Expressive Ästhetik“. Dies liegt möglicherweise an der etwas unspezifischen Formulierung, die tatsächlich einen Zusammenhang mit diesen Dimensionen plausibel macht. Zudem ist das Item für die Zielgruppe eventuell zu abstrakt und entsprechend etwas weit vom konkreten Gegenstand entfernt. Eine Möglichkeit wäre die Eliminierung dieses Items, was auch durch die niedrigsten Trennschärfen (.48, .50) in der Reliabilitätsstatistik begründet werden könnte. Trotzdem wurde es im Fragebogen belassen, damit weitere Untersuchungen zeigen können, inwieweit sich dieses Item bewährt.

Die Dimension „Usability“ lässt sich bei beiden Apps am deutlichsten bestätigen. Die negativen Emotionen mit „Die App langweilt mich“ und „Die App verärgert mich“ laden deutlich auf „Emotion“ wobei die positiven Emotionen mit „Das Benützen der App macht Spass“ und „Die App begeistert mich“ bei Solar Walk auch auf „Expressive Ästhetik“ und bei Physics Lite auch auf „Kognitive Stimulation“ laden. Dies zeigt, dass die negativen Emotionen als Konstrukt schärfer abgrenzbar sind. Für eine Weiterentwicklung des Fragebogens in der Dimension „Emotion“ wäre deshalb eine Anreicherung von passenden Items aus dem Achievement Emotion questionnaire von Pekrun Goetz und Perry (2005) oder anderen Skalen denkbar.

Sehr Interessant ist die Dimension „Kognitive Stimulation“, die sich deutlich aus den verschiedenen Konstrukten: Involvement, Cognitive Stimulation, Positive Attraction, Focused Attention und intrinsischer Motivation zusammensetzt. Anscheinend lassen sich hier kognitive Prozesse, die beim Lernen stattfinden, in einer Dimension zusammenfassen.

Um das Testgütekriterium Reliabilität (vgl. Kapitel 3.2.6) zu beurteilen, kann die interne Konsistenz herangezogen werden. Die Reliabilitätsstatistik über die gesamte Skala generiert ohne

das Item „Die Benützung der App ist anspruchsvoll“ bei Solar Walk einen mittleren Wert (vgl. Kapitel 3.6.13) von  $\alpha=.86$  und bei Physics Lite einen hohen Wert von  $\alpha=.91$ . Die Reliabilitätsstatistik der Dimensionen zeigt für das App Physics Lite durchwegs höhere Cronbachs Alpha Werte ( $\alpha=.77$  bis  $\alpha=.84$ ) als Solar Walk ( $\alpha=.61$  bis  $\alpha=.80$ ). Bei Eliminierung des Items „Die Benützung der App ist anspruchsvoll“ erhöht sich der Cronbachs Alpha Wert in „Usability“ auf  $\alpha=.74$  bei Solar Walk respektive  $\alpha=.84$  bei Physics Lite. Insgesamt können die Dimensions-Reliabilitäten gemäss Kapitel 3.6.13 als niedrig bis mittel eingestuft werden, was unter anderem mit der Anzahl der Items zusammenhängen könnte. Für eine Weiterentwicklung des Fragebogens wäre die Erhöhung bei den Dimensionen „Usability“, „Expressive Ästhetik“ und „Emotion“ auf fünf Items zu prüfen. Die Dimension „Kognitive Stimulation“ ist in dieser Zusammenstellung neu und verfügt über gute Kennzahlen.

Die geringen, teilweise nicht signifikanten Korrelationen zwischen „Usability“ und den anderen Faktoren zeigen, dass Usability nicht als unabhängiges Konstrukt extrahiert werden kann, wie beispielsweise von Lindgaard und Dudek (2003, zitiert nach Thielsch, 2008) angenommen. Die mittleren Korrelationen unter den anderen Faktoren bestätigen einen Zusammenhang dieser Konstrukte. Daraus lässt sich schliessen, dass Emotionen und kognitive Prozesse eng miteinander verknüpft sind, wie es auch durch die Theorie bestätigt wird (vgl. Kapitel 2).

Die tiefe, respektive mittlere und signifikante Korrelation von „Expressiver Ästhetik“ mit „Emotion“ und „Kognitiver Stimulation“ weist darauf hin, dass Elemente, die als faszinierend und originell wahrgenommen werden, entsprechend positive Emotionen auslösen und als stimulierend empfunden werden. Insgesamt können aber diese Korrelationen keine gerichteten Ursache-Wirkungszusammenhänge zwischen den Konstrukten aufzeigen, deshalb müssten sie in weiteren Studien und mit anderen Methoden genauer untersucht werden.



Der Mittelwertvergleich zwischen den Apps zeigt, dass Solar Walk auf allen Variablen höher eingeschätzt wird. Möglicherweise liegt das am Interaktionsstil des Menüs (vgl. Kapitel 3.4), bei dem sich der Nutzer mitten im Weltall befindet und zwischen den Planeten herumfliegen kann. Eventuell spielen die sehr realistische Darstellung und der einfache Abruf von Informationen, der eher spielerische Ansatz und die Originalität im Vergleich zu anderen Apps eine Rolle. Hierzu lassen sich nur Vermutungen anstellen. Allerdings scheinen die Korrelationen zu bestätigen, dass wahrgenommene hohe expressive Ästhetik auch hohe Werte bei Emotionen und kognitiver Stimulation nach sich ziehen, die sich wiederum in der Empfehlung der App niederschlagen.

Der Unterschied bezüglich der Lösung der Lernaufgabe erklärt sich hingegen durch die Schwierigkeit der Lernaufgabe. Sie war bei Solar Walk einfach konstruiert, sodass es 84 Prozent richtige Lösungen gab. Bei den schwierigeren Aufgaben in Physics Lite spielt das Bildungsniveau eine Rolle, was durch die signifikanten Korrelationen bestätigt wird. Ausserdem korreliert das Bildungsniveau auch mit dem Interesse am Thema Physik, was beim Thema Astronomie in Solar Walk nicht der Fall ist.

Beim Vergleich der Geschlechter zeigen sich nur sehr geringe Unterschiede mit kleinen Effekten, was die Theorie bestätigt.

Im Gegensatz zur häufigen Vorstellung, dass Digital Natives ständig elektronische Geräte benützen, zeigen die Resultate der „App Nutzung pro Woche“ ein komplett anderes Bild. 60 Prozent der Befragten benützen nie eine Lern-App, 25 Prozent einmal und nur 15 Prozent mehrmals in der Woche. Bei der Frage wurden alle Geräte (Tablet, Handy, Computer) eingeschlossen. Dies zeigt, dass sowohl in Schulen als auch privat der Gebrauch von Lern-Apps noch wenig verbreitet ist, womöglich aufgrund fehlender, brauchbarer Produkte und unterstreicht das Interesse an der Entwicklung von solchen Applikationen.

In Bezug auf die Validität als Testgütekriterium eines psychologischen Messinstruments (vgl. Kapitel 3.2.7) lassen sich folgende Aussagen formulieren: Inhalts- und Augenscheinvalidität können als gegeben betrachtet werden, da ein Grossteil der Items aus bereits validierten, von Fachleuten entwickelten Messinstrumenten übernommen wurden. Die Konstruktvalidität wurde nur mit der faktoriellen Validität untersucht, indem eine Faktorenanalyse durchgeführt wurde. Die Analyse zeigt die Zusammenfassung homogener, konstruktnaher Items respektive die vier vorher beschriebenen Dimensionen. Zu konvergenter sowie diskriminanter Validität können keine Aussagen gemacht werden, da keine Korrelationen mit Tests gleicher oder anderer Gültigkeitsbereiche vorliegen. Hierzu braucht es weitere Studien.

#### 5.4. Implikationen für die Praxis und Ausblick

Der Fragebogen zur Erfassung von ästhetischem Empfinden ist mehrdimensional und umfasst die vier Skalen „Kognitive Stimulation“, „Expressive Ästhetik“, „Emotion“ und „Usability“. Die Dimension „Kognitive Stimulation“ ist durch die Integration verschiedener relevanter Konstrukte im Kontext des Lernens neu. Sie könnte somit bei der Beurteilung von Ästhetik in Zusammenhang mit Lernmedien einen wichtigen Beitrag leisten. Denkbar ist auch eine Anwendung in Kombination mit anderen herkömmlichen Skalen zu Usability, User Experience und Ästhetik.

Die Korrelationen unter den Dimensionen „Kognitive Stimulation“, „Expressive Ästhetik“, „Emotion“ und „Usability“ zeigen, dass sie sich bei der ästhetischen Empfindung wechselseitig beeinflussen. Entsprechend sollte dies bei der Gestaltung von Interfaces berücksichtigt werden. Das Ausschliessen oder das Isolieren eines oder mehrerer Konstrukte ist deshalb wohl nicht angebracht.

Entscheidend ist die Frage, welche Gestaltungselemente einer Applikation die entsprechenden Konstrukte bei den ästhetischen Wahrnehmungs- und Beurteilungsprozessen beeinflussen. Wie

bereits erwähnt, scheint ein originelles und faszinierendes Design mit passenden Spezialeffekten positive Auswirkungen auf „Kognitive Stimulation“ und „Emotion“ zu haben.

Die Gestaltung des Inhalts, insbesondere bei Lernthemen scheint ein wichtiges Element zu sein. Der Inhalt sollte so aufbereitet und gestaltet sein, dass er Neugierde, Inspiration und Aufmerksamkeit erzeugt, Langeweile vermeidet und den User dazu motiviert sich mit dem präsentierten Thema auseinanderzusetzen. Hier könnten interaktive Features gemäss Loranger und Nielsen (2013) einen Beitrag leisten.

Obwohl die Verwendung von Sprache und Text nicht explizit Teil der Untersuchung war, sind sie im Zusammenhang mit Lernen (vgl. Mayer, 2014b) wichtige Gestaltungsmerkmale. Beide getesteten Apps kommen mit sehr wenig Text aus, möglicherweise hatte dies einen Einfluss auf die eher hohen Bewertungen. In diesem Zusammenhang ist auch besonders zu berücksichtigen, dass Jugendliche, die in der Schweiz in die Volksschule gehen, sehr unterschiedliche Kompetenzen in der deutschen Sprache haben. Sei es aufgrund ihres Migrationshintergrundes, ihres Bildungsniveaus oder anderer Ursachen (vgl. Loranger & Nielsen, 2013).

Emotionen können durch verschiedene Gestaltungsmerkmale ausgelöst werden (vgl. Kapitel 2.3.3) und begleiten den ganzen Wahrnehmungs- und Bewertungsprozess. Die erfolgreiche Integration von Wissen bei der kognitiven Verarbeitung (vgl. Leder, 2004) legt nahe, dass der Inhalt der App so aufbereitet wird, dass Anknüpfungspunkte aus der Lebens- und Wissenswelt der Jugendlichen bestehen und somit ihrer Expertise in diesem Bereich gerecht wird. Das gleiche gilt wahrscheinlich für die Formensprache, die Komplexität, den Stil, die Menüführung etc. Weiter führt eine verständliche, einfache und funktionierende Usability zu positiven Emotionen, respektive vermeidet negative Emotionen. Hier sind die Besonderheiten der Applikation, wie sie Loranger und Nielsen (2013) beschreiben, zu berücksichtigen.

Weiter kann der soziale Austausch, der in dieser Arbeit nicht berücksichtigt wurde, positive Auswirkungen auf die „Kognitive Stimulation“ und „Emotionen“ haben (vgl. Loranger und Nielsen, 2013).

Alles in allem bietet die Applikation durch enorme Gestaltungsmöglichkeiten grosses Potential, um bei der Vermittlung von MINT-Themen das interessierte Lernen zu fördern. Die genauen Wirkungszusammenhänge in diesem Anwendungskontext sind jedoch noch wenig bekannt und bedürfen noch weiterer Forschungen.

Werden Videos (vgl. Derry, Sherin & Sherin, 2014) und Games (vgl. Tobias, Fletcher, Bediou, Wind & Chen, 2014) Teil der Applikation, so gewinnt der Ton als kritische Variable an Bedeutung. Die Anwendung von Augmented Reality (vgl. Dunleavy, 2013) erweitert das Feld der Gestaltungsmöglichkeiten weiter und könnte vor allem in der lernförderlichen Vermittlung von realen Objekten und abstrakten Wissensinhalten eine grosse Rolle spielen.

Die genannten Aspekte zeigen, dass Interface-Ästhetik in Verbindung mit Multimedia-Lern-Design ein wichtiges Forschungsfeld für die Zukunft bleiben wird.

## Abbildungsverzeichnis

Abbildung 1: Modell der ästhetischen Erfahrung nach Leder et al. (2004, S. 492).....	5
Abbildung 2: Zentrale Variablen bei der Website-Rezeption (Thielsch, 2008, S. 263) .....	25
Abbildung 3: Kontroll-Wert Theorie, nach Pekrun (2006, adaptiert von Harrer 2012, S. 23) ..	28
Abbildung 4: Arbeitsmodell ästhetisches Empfinden .....	33
Abbildung 5: Untersuchungs-Design .....	36
Abbildung 6: Scree-Plot Solar Walk .....	74
Abbildung 7: Scree-Plot Physics Lite .....	76
Abbildung 8: Mittelwertvergleich Solar Walk, Physics Lite .....	82

## Tabellenverzeichnis

Tabelle 1: Web-Design für Altersgruppe, nach Loranger & Nielsen (2013, S. 5).....	20
Tabelle 2: Mayers Multimedia-Prinzipien, eigene Darstellung nach Mayer (2014b).....	26
Tabelle 3: Lernmotivation: Zentrale Konstrukte und Theorien, nach Spinath (2011, S. 50).....	29
Tabelle 4: Verwendete Instrument und Skalen zur Fragebogenkonstruktion .....	44
Tabelle 5: Definitive Version Fragebogen .....	46
Tabelle 6: Vergleich der Test-Apps .....	49
Tabelle 7: Stichprobe.....	51
Tabelle 8: Bewertung KMO-Koeffizient .....	57
Tabelle 9: Kennwerte für Items und Skalen .....	60
Tabelle 10: Variablenübersicht .....	61
Tabelle 11: Deskriptive Statistik Solar Walk .....	65
Tabelle 12: Deskriptive Statistik Physics Lite .....	66
Tabelle 13: Deskriptive Statistik Geschlecht, Niveau, Lernaufgabe, App-Nutzung, Alter .....	67
Tabelle 14: Häufigkeiten App-Nutzung .....	67
Tabelle 15: Häufigkeiten Alter.....	68
Tabelle 16: Häufigkeiten Lernaufgaben Solar Walk.....	68
Tabelle 17: Häufigkeiten Lernaufgaben Physics Lite .....	68
Tabelle 18: KMO- und Bartlett-Test V1 .....	69
Tabelle 19: Kommunalitäten Solar Walk, Physics Lite V1 .....	70
Tabelle 20: KMO- und Bartlett-Test .....	71
Tabelle 21: Kommunalitäten Solar Walk, Physics Lite .....	72
Tabelle 22: Erklärte Gesamtvarianz Solar Walk.....	73
Tabelle 23: Erklärte Gesamtvarianz, Physics Lite .....	75
Tabelle 24: Faktorladungen Solar Walk.....	77
Tabelle 25: Faktorladungen Physics Lite .....	78
Tabelle 26: Reliabilitätsstatistik .....	79
Tabelle 27: Faktorenkorrelationen Solar Walk .....	80
Tabelle 28: Faktorenkorrelationen Physics Lite.....	81
Tabelle 29: Korrelationen.....	84
Tabelle 30: Fragebogen nach der statistischen Überprüfung .....	90

## 6. Literaturverzeichnis

- Bargas-Avila, J. A. & Hornbaek, K., (2011). *Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience*. In CHI'11 May 7-12, 2011. Vancouver, BC, Canada (pp. 2689-2698). New York: ACM.
- Belke, B. & Leder, H. (2006). Annahmen eines Modells der ästhetischen Erfahrung aus kognitionspsychologischer Perspektive. In: *Sonderforschungsbereich 626 (Hrsg.): Ästhetische Erfahrung: Gegenstände, Konzepte, Geschichtlichkeit*. Freie Universität Berlin. Verfügbar unter: [http://www.sfb626.de/veroeffentlichungen/online/aesth\\_erfahrung/aufsaeetze/belke\\_leder.pdf](http://www.sfb626.de/veroeffentlichungen/online/aesth_erfahrung/aufsaeetze/belke_leder.pdf) [03.01.2017].
- Bonnardel, N., Piolat, A. & Le Bigot, L. (2011). The impact of color on website appeal and users' cognitive processes. *Displays*, 32, 69-80.
- Bortz, J. & Döring, N. (2009). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Aufl.). Berlin: Springer.
- Brosius, F. (2011). *SPSS 19* (1. Aufl.). Heidelberg: mitp.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion* (2. Aufl.). München: Pearson.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3. Aufl.). München: Pearson.
- Derry, S. J., Sherin, M. G. & Sherin, B. L. (2014). Introduction to Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (p. 785-812). New York: Cambridge University Press.
- Deutsches Institut für Normung e.V. EN ISO 9241-115 (DIN EN ISO 9241-115). (2008). *Ergonomie der Mensch-System-Interaktion – Teil 151: Leitlinien zur Gestaltung von Benutzungsschnittstellen für das World Wide Web*. Berlin: Beuth.
- Deutsches Institut für Normung e.V. EN ISO 9241-210 (DIN EN ISO 9241-110). (2008). *Ergonomie der Mensch-System-Interaktion – Teil 110: Grundsätze der Dialoggestaltung*. Berlin: Beuth.
- Deutsches Institut für Normung e.V. EN ISO 9241-210 (DIN EN ISO 9241-210). (2010). *Ergonomie der Mensch-System-Interaktion – Teil 210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme*. Berlin: Beuth.
- Diefenbach, S. & Hassenzahl, M. (2010). *Handbuch zur Fun-ni Toolbox. User Experience Evaluation auf drei Ebenen*. Verfügbar unter: [http://fun-ni.org/wp-content/uploads/Diefenbach+Hassenzahl\\_2010\\_HandbuchFun-niToolbox.pdf](http://fun-ni.org/wp-content/uploads/Diefenbach+Hassenzahl_2010_HandbuchFun-niToolbox.pdf) [12.01.2017].
- Dunleavy, M. (2013). Design principles for augmented reality learning. *TechTrends*, 58 (1), 28-34.
- Fachhochschule Nordwestschweiz, Brugg, Forschung und Entwicklung (2017). Strategische Initiative «EduNaT». Verfügbar unter: <http://www.fhnw.ch/forschung-und-entwicklung/strategische-initiativen/edunat> [03.01.2017].
- Fachhochschule Nordwestschweiz, Brugg, Pädagogische Hochschule, Institut Weiterbildung und Beratung, Beratungsstelle Digitale Medien in Unterricht und Schule-imedias (2016). App-Advisor. Verfügbar unter: <https://www.imedias.ch/appadvisor/> [13.12.16].
- Fisseni, H. J. (2004). *Lehrbuch der psychologischen Diagnostik* (3. Aufl.). Göttingen: Hogrefe.
- Fromm, S. (2004). Faktorenanalyse. In N. Baur & S. Fromm (Hrsg.), *Datenanalyse mit SPSS für Fortgeschrittene-Ein Arbeitsbuch*. Wiesbaden: Verlag für Sozialwissenschaften.
- Hager, M., Hagemann, D., Danner, D. & Schankin, A. (2012). Assessing aesthetic appreciation of visual artworks - The construction of the art reception survey (ARS). *Psychology of Aesthetics, Creativity, and the Arts*, 6 (4), 320-333.

- Harrer, K. (2012). Empirische Überprüfung des 3 x 2 Modells der Zielorientierungen und dessen Zusammenhang mit Lernfreude und Langeweile. Unveröffentlichte Diplomarbeit, Universität Wien. Verfügbar unter: [http://othes.univie.ac.at/22363/1/2012-09-10\\_0606443.pdf](http://othes.univie.ac.at/22363/1/2012-09-10_0606443.pdf) [02.01.2017].
- Harrison, L., Reinecke, K. & Chang, R. (2015). *Infographic Aesthetics: Designing for the first impression*. CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. (p 1187-1190). New York: ACM.
- Hassenzahl, M. & Tractinsky, N. (2006). User Experience – a research agenda (editorial). *Behaviour and Information Technology* 25 (2), 91-97.
- Hassenzahl, M. (2001). The Effect of Perceived Hedonic Quality on Product Appealingness. *International Journal of Human-Computer Interaction*, 13 (4), 481-499.
- Hassenzahl, M., Burmester, M. & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In J. Ziegler & G. Szwillus (Hrsg.), *Mensch & Computer 2003. Interaktion in Bewegung* (S. 187-196). Stuttgart: B. G. Teubner.
- Heidig, S., Müller, J. & Reichelt, M. (2015). Emotional Design in multimedia learning: Differentiation on relevant design features and their effects on emotions and learning. *Computers in Human Behaviour* 44, 81-95.
- Huber, O. (2009). *Das psychologische Experiment: Eine Einführung*. Bern: Hans Huber.
- Isen, A. M. & Reeve J. (2005). The influence of positive affect on intrinsic and extrinsic motivation: Facilitating enjoyment of play, responsible work behaviour and self-control. *Motivation and Emotion*, 29 (4), 297-325.
- Jonkisz, E., Moosbrugger, H. & Brandt, H. (2012). Planung und Entwicklung von Tests und Fragebogen. In K. Moosbrugger (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 27–74). Berlin: Springer.
- Krohne, H. W. & Egloff, B., Kohlmann, C-W. & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der “Positive and Negative Affect Schedule” (PANAS). *Diagnostica*, 42 (2), 139-156.
- Lavie, T. & Tractinsky N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies* 60, 269-298.
- Leder, H., Belke, B., Oeberst & A., Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgements. *British Journal of Psychology*, 95, 489-508.
- Leonhart, R. (2010). *Datenanalyse mit SPSS*. Göttingen: Hogrefe.
- Leutner, D. (2014). Motivation and Emotion as mediators in multimedia learning. *Learning and Instruction* 29, 174-175.
- Lindgaard, G., Fernandes, G., Dudek, C. & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression!. *Behaviour & Information Technology*, 25 (2), 115-126.
- Loranger, H. & Nielsen, J. (2013). *Teenage Usability: Designing Teen-Targeted Websites*. Verfügbar unter: <https://www.nngroup.com/articles/usability-of-websites-for-teenagers/> [12.01.2017].
- Mayer, R. E. (2009). *Multimedia learning* (2. Aufl.). Cambridge: Cambridge University Press.
- Mayer, R. E. (2014). Incorporating motivation into multimedia learning. *Learning and Instruction* 29, 171-173.
- Mayer, R. E. (2014b). Introduction to Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (p. 01-24). New York: Cambridge University Press.

- Minge, M., Riedel, L. & Thüring, M. (2013). *Modulare Evaluation von Technik. Entwicklung und Validierung des meCUE Fragebogens zur Messung der User Experience*. Verfügbar unter: [http://mecue.de/Homepage%20Content/02%20Links%20%26%20Literatur/minge\\_riedel\\_thuring\\_2013b.pdf](http://mecue.de/Homepage%20Content/02%20Links%20%26%20Literatur/minge_riedel_thuring_2013b.pdf) [12.01.2017].
- Moosbrugger, H. & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion* (2. Aufl.). Berlin: Springer.
- Moreno, R. & Mayer, R. E., (2007). Interactive multimodal learning environments. *Educational Psychology Review*, 19, 309-326.
- O'Brian, H. L. & Toms E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the american society for information science and technology*, 61 (1), 50-69.
- Pekrun R, Goetz, T. & Perry P. (2005) *Achievement Emotions Questionnaire (AEQ): User's Manual*. Verfügbar unter: <https://de.scribd.com/doc/217451779/2005-AEQ-Manual> [03.01.2017].
- Pekrun, R. & Perry, R. P. (2014). Control Value Theory of Achievement Emotions. In: R. Pekrun and L. Linnebrink-Garcia (Eds.), *International Handbook of Emotions in Education* (p. 120-141). New York: Routledge.
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries and implications for educational research and practice. *Educational Psychology Review*, 18, 315-341.
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of quantitative and qualitative research. *Educational Psychologist*, 37, 91-106.
- Plass, J. L., Heidig, S., Hayard, E. O., Homer, B. D., & Um, E. (2014). Emotional Design in multimedia Learning: Effects of shape and color on affect and learning. *Learning and Instruction* 29, 128-140.
- Porst, R. (2011). *Fragebogen - Ein Arbeitsbuch* (3. Aufl.). Wiesbaden: VS.
- Rauchfleisch, U. (2008). *Testpsychologie – Eine Einführung in die Psychodiagnostik* (5. Aufl.). Göttingen: UTB.
- Rost, J. (2004). *Lehrbuch Testtheorie, Testkonstruktion* (2. Aufl.). Bern: Huber.
- Sauro, J. (2015). SUPR-Q: A Comprehensive Measure of the Quality of the Website User Experience. *Journal of Usability Studies*, 10 (2), 68-86.
- Schrepp, M. & Müller K. E. (2015). Übersichtlichkeit als Mediator zwischen Ästhetik und Usability? In S. Diefenbach, N. Henze & M. Pielot (Hrsg.), *Mensch und Computer 2015 Tagungsband* (S. 73-82). Stuttgart: Oldenbourg Wissenschaftsverlag.
- Schulpsychologischer Dienst Kanton Baselland (2015). *Lern-Apps: Interaktive und multimediale Lernbausteine*. Verfügbar unter: [https://www.baselland.ch/fileadmin/baselland/files/docs/ekd/spd/lern-app\\_flyer.pdf](https://www.baselland.ch/fileadmin/baselland/files/docs/ekd/spd/lern-app_flyer.pdf) [13.12.16].
- Seckler, M. Opwis, K. & Tuch, A. N. (2015). Linking objective design factors with subjective aesthetics: An experimental study on how structure and color of websites affect the facets of users' visual aesthetic perception. *Computers in Human Behaviour*, 49, 375-389.
- Spinath, B. (2011). Lernmotivation. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung* (S. 45-56). Wiesbaden: Verlag für Sozialwissenschaften.
- Sutcliffe, A. & De Angeli, A. (2005). Assessing interaction Styles in web user interfaces. In M. F. Costabile & F. Paterno (Eds.). *Human-Computer-Interaction-Interact 2005* (S. 405-417). Berlin: Springer.
- Thielsch, M. T. & Moshagen, M. (2010). Erfassung visueller Ästhetik mit dem VisAWI. In H. Brau, A., Lehmann, K. Petrovic & M. C. Schroeder (Hrsg.), *Usability Professionals 2011* (S. 260-265). Stuttgart: German UPA.
- Thielsch, M. T. (2008). *Ästhetik von Websites*. Münster: MV-Verlag.



- Thüring, M. & Mahlke, S. (2007). Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology*, 42 (4), 253–264.
- Tobias, S., Fletcher, J. D., Bediou, B., Wind, A. P. & Chen, F. (2014). Multimedia Learning with Computer Games. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (p. 762-784). New York: Cambridge University Press.
- Tuch, A. N., Bargas-Avila, J. A., Opwis, K. & Wilhelm F. H. (2009). Visual complexity of websites: Effects on users' experience, physiology, performance and memory. *International Journal of Human-Computer Studies*, 67, 703-715.
- Tuch, A. N., Presslaber, E. E., Stöcklin, M., Opwis, K. & Bargas-Avila, J. A. (2012). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgements. *International Journal of Human-Computer Studies*, 70, 794-811.
- Tuch, A. N., Roth, S. P., Hornbaek, K., M., Opwis, K. & Bargas-Avila, J. A. (2012). Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI. *Computers in Human Behaviour*, 28, 1596-1607.
- Um, E. R., Plass, L., Hayward, E. O. & Homer, B. D. (2012). Emotional Design in Multimedia Learning. *Journal of Educational Psychology*, 104 (2), 485-498.
- Universität Zürich, Zürich, Methodenberatung (2016). t-Test für unabhängige Stichproben. Verfügbar unter: <http://www.methodenberatung.uzh.ch/de/datenanalyse/unterschiede/zentral/ttestunabh.html> [03.01.2017].
- Wang, H. F. (2014). Picture perfect: Girls' and boys' preferences towards visual complexity in children's websites. *Computers in Human Behavior*, 31, 551–557.
- Zahn, D. & Agotai, D. (2015). *Forschungsantrag Einfluss der Interface-Ästhetik digitaler Medien auf das Lernen von MINT-Themen in Ausstellungen*. Unveröffentlichtes Dokument, APS, Fachhochschule Nordwestschweiz in Olten.

## 7. Anhang

### Inhaltsverzeichnis

	Seite
Kapitel 3: Methoden	
Übersicht über verwendete Fragebogen und Skalen	01
Übersetzung und Rückübersetzungen der Items	02
Printscreens Solar Walk	03
Printscreens Physics Lite	04
Skript Testdurchführung	05
Testunterlagen	07
Kapitel 4: Ergebnisse	
KSA-Test Solar Walk	14
KSA-Test Physics Lite	15
Histogramme	16
Mittelwertvergleich Geschlecht Solar Walk	22
Mittelwertvergleich Geschlecht Physics Lite	23
Mittelwertvergleich Geschlecht weitere Variablen	24
Korrelationen Solar Walk	25
Anti-Image-Matrix Solar Walk	27
Korrelationen Physics Lite	28
Anti-Image-Matrix Physics Lite	30
Wilcoxon-Test für abhängige Stichproben	31
Fragebogen zur Erfassung des ästhetischen Empfindens	32

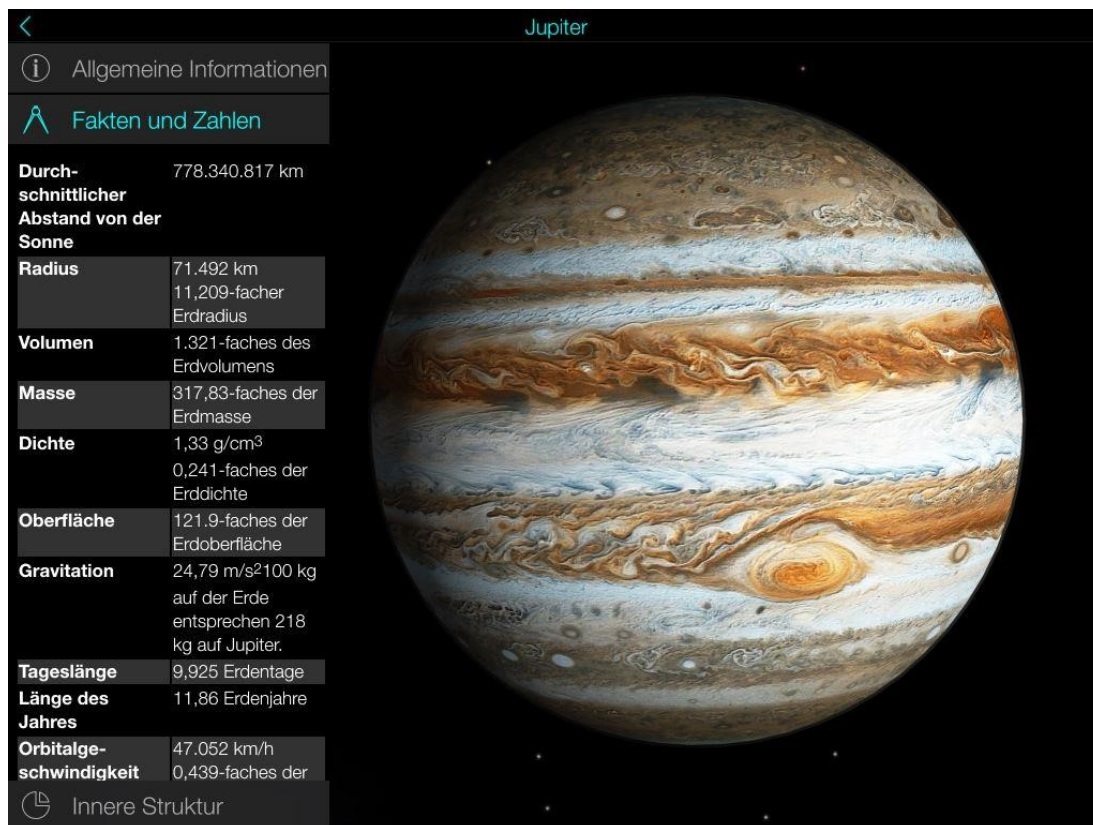
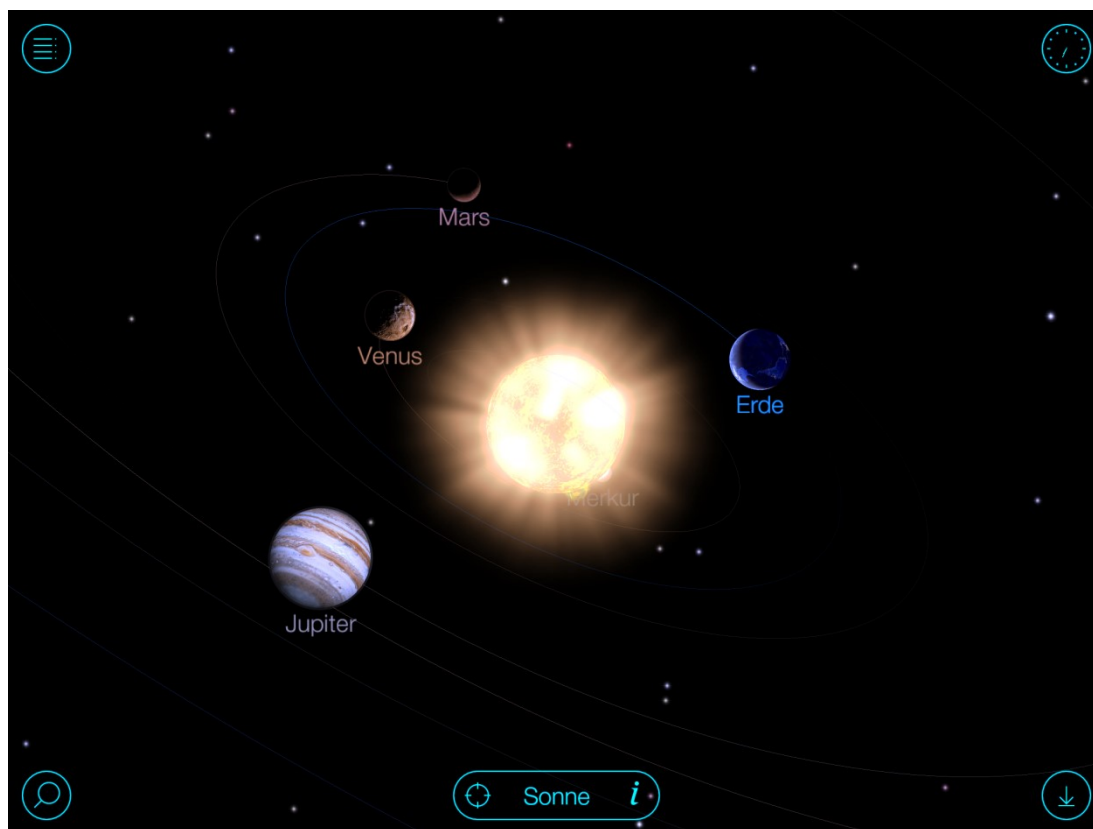
## Übersicht über verwendete Fragebogen und Skalen

Instrument	Quelle	Subskalen	Cronbachs Alpha	Antwortformat	Anzahl der Items	Beispielitems
Dimensions of perceived visual aesthetics of web sites	Lavie & Tractinsky (2004)	Expressive Aesthetics	.87	7-stufig	5	<i>Fascinating Design.</i> <i>I feel joyful.</i>
		Pleasurable Interaction	.86	strongly disagree – strongly agree	3	
Standardized User Experience Percentile Rank Questionnaire (SUPR-Q)	Sauro (2015)	Usability	.88	5-stufig strongly disagree – strongly agree	2	<i>The website is easy to use.</i>
		Loyalty	.64	11-stufig not at all likely – extremely likely	2	<i>How likely are you to recommend this website to a friend or colleague?</i>
User Engagement Scale (UES)	O'Brian & Toms (2009)	Focused attention	.92	5-stufig strongly disagree – strongly agree	7	<i>I blocked out things around me, when I was shopping on this website</i>
		Involvement	.71		3	<i>I was really drawn into my shopping task.</i>
		Perceived Usability	.91		8	<i>This shopping experience is demanding.</i>
Modularer Fragebogen zur Erfassung des Nutzungserleben (meCUE)	Minge & Riedel (2013)	Nützlichkeit	.83	5-stufig stimme völlig zu – lehne völlig ab	3	<i>Mithilfe des Produkts kann ich meine Ziele erreichen.</i>
		Benutzbarkeit	.89		3	<i>Die Bedienung des Produkts ist verständlich.</i>
Intrinsic Motivation	Isen & Reeve (2005)	Intrinsic Motivation	.92	7-stufig strongly disagree – strongly agree	8	<i>It makes me feel curious about it.</i>
Positive and Negative Affect Schedule (PANAS)	Krohne et al. (1996)	Positive Emotion	.85	5 stufig gar nicht - äusserst	10	<i>Wie fühlen Sie sich im Moment: begeistert.</i>
		Negative Emotion	.86		10	<i>Wie fühlen Sie sich im Moment: verärgert.</i>
Ars reception survey (ARS)	Hager et al. (2012)	Positive attraction	.83	5-stufig	5	<i>This painting thrills me.</i>
		Cognitive Stimulation	.90	completely disagree – completely agree	5	<i>This painting is thought-provoking.</i>
Achievment emotions questionnaire (AEQ)	Pekrun (2005)	Learning related boredom	.92	5-stufig strongly disagree – strongly agree	11	<i>The material bores me to death.</i>

## Übersetzung und Rückübersetzungen der Items

Subskalen	Originalitem	Übersetzung	Rück-Übersetzung
Expressive Aesthetics	Fascinating Design.	Faszinierendes Design	The website design is fascinating.
Expressive Aesthetics	Creative Design.	Kreatives Design.	The website design is creative.
Expressive Aesthetics	Use of special effects	Die Verwendung von Spezialeffekten...	Use of special effects...
Pleasurable Interaction	I feel joyful.	Ich fühle Freude.	I feel happy.
Usability	The website is easy to use.	Die Webseite ist einfach zu gebrauchen	The website is easy to use.
Usability	It's easy to navigate within the website.	Es ist einfach innerhalb der Webseite zu navigieren.	It is easy to navigate the website.
Loyalty	How likely are you to recommend this website to a friend or colleague?	Wie wahrscheinlich ist es, dass Sie diese Webseite einem Freund oder Kollegen empfehlen würdest.	How likely are you to recommend this website to a friend or colleague?
Focused attention	I blocked out things around me, when I was shopping on this website	Ich blendete Dinge um mich aus, als ich auf dieser Webseite eingekauft hatte.	I blanked out the world around me while shopping on this website.
Involvement	I was really drawn into my shopping task.	Ich wurde richtig in meine Einkaufsaufgabe hineingezogen.	I was really absorbed by my shopping task.
Perceived Usability	This shopping experience is demanding.	Diese Einkaufs-Erfahrung ist anspruchsvoll.	This is a high-quality shopping experience.
Intrinsic Motivation	It makes me feel curious about it.	Es macht mich neugierig	It makes me curious.
Intrinsic Motivation	It makes me want to explore it further.	Ich möchte es weiter erforschen.	I would like to investigate it further.
Intrinsic Motivation	It's interesting.	Es ist interessant	It is interesting.
Positive attraction	This painting thrills me.	Dieses Gemälde begeistert mich.	This painting is enthralling.
Cognitive Stimulation	This painting is thought-provoking.	Dieses Gemälde ist Gedanken anregend.	This painting is thought-provoking.
Learning related boredom	The material bores me to death.	Das Material langweilt mich zu Tode.	The material bores me to death.

## Printscreens Solar Walk



### Printscreens Physics Lite

Mechanische Schwingungen und Wellen

**IV. Molekulare Physik und Thermodynamik**

- ▶ 1. Brownsche Bewegung
- ▶ 2. Maxwell-Geschwindigkeitsverteilung
- ▶ 3. Dampflokomotive
- ▶ 4. Perpetuum mobile

**V. Elektrizität und Magnetismus**

- ▶ 1. Ladungen und Felder
- ▶ 2. Elektrische Schaltung
- ▶ 3. Elektronenstrahlung



## Skript Testdurchführung

### 1. Mich vorstellen

- ursprünglich Lehrperson Realschule, Schulleiter
- Studium Angewandte Psychologie, FHNW Olten
- Masterarbeit: Design von Lern-Apps
- auf Beziehung zu Lehrperson hinweisen

1 Minute

### 2. Ablauf des Testes erklären (Ablauf steht am Flip-Chart) (1 Minute)

- Allgemeine Infos und Infos zum Fragebogen
- Schlüsselwörter erklären
- Einführung I-Pad allgemein
- 1. App Solar Walk (ca. 15 min.)
- Lernfragen beantworten, eigenes Erforschen, Test ausfüllen
- 2. App Physics lite (ca. 15 min.)
- Lernfragen beantworten, eigenes Erforschen, Test ausfüllen

1 Minute

### 3. Schüler lesen 1. und 2. Seite der Testunterlage

- Hinweis zum Ausfüllen, Vorzeigen auf dem Visualizer
- Fragen zum Gelesenen beantworten

nach Bedarf

### 4. Schlüsselwörter im Fragebogen erklären (Liste auf Flip-Chart)

2 Minuten

Design	Wie sieht die App aus? Farben, Formen, Bilder, Schriften, Animationen
faszinierend	anziehend, fesselnd, bezaubernd
originell	einfallsreich, neuartig, aussergewöhnlich
Spezialeffekte	bewegte Bilder und Symbole, 3-D Ansichten,
navigieren	ein Programm oder einen Programmpunkt nach dem anderen aktivieren
anspruchsvoll	braucht volle Konzentration, schwierig zu verstehen

### 5. Tutorial I-Pad (vorzeigen mit I-Pad)

SchülerInnen machen gleich simultan mit

- ein/aus oben rechts
- Lautstärke rechts oben
- Hauptmenü unten Mitte
- Wischen um Apps zu finden

1 Minute

### 6. Tutorial Solar Walk (vorzeigen mit Beamer)

SchülerInnen machen gleich simultan mit

1 Minute

---

**Hauptfunktionen:**

- Zoomen mit zwei Fingern (Scherenbewegung)
- Ansicht drehen mit einem Finger
- verschieben mit 2 Fingern parallel

**Planeten ansteuern:**

- mit Lupe unten links kommt man auf das Menü
  - oder mit dem Steuern der Finger
  - Infos zum Planeten mit *i* –der betreffende Planet steht im Kästchen  
> Menü Allgemeine Infos, Fakten und Zahlen, Innere Struktur, Atlas
  - < Pfeil oben links, zurück zum Hauptmenü
- 

**7. Lernen und Fragebogen ausfüllen mit Solar Walk**

15-20 Minuten
---------------

---

**8. Tutorial Physics lite, Dampflokomotive (Affe ist gleich)  
SchülerInnen machen gleich simultan mit**

1 Minute
----------

**Funktionen:**

- Scrollen mit einem Finger im Hauptmenü
  - grüner Pfeil -> öffnen des Untermenüs
  - rot = Stopp Animation, gelb = Pause und Kurzvorlauf, grün = Start Animation
  - Symbol Ton ein, aus
  - Haken im Kästchen: Vorwärts, rückwärts
  - Schieberegler = Geschwindigkeit
  - schwarzes Viereck -> zurück zum Hauptmenü
- 

**9. Lernen und Fragebogen ausfüllen mit Physics lite**

15-20 Minuten
---------------

---

**10. Ende**

- Dank an Schüler und Lehrperson
- Klasse erhält eine kleine Belohnung (Sack Süßigkeiten)
- Einziehen der Testunterlagen

1 Minute
----------

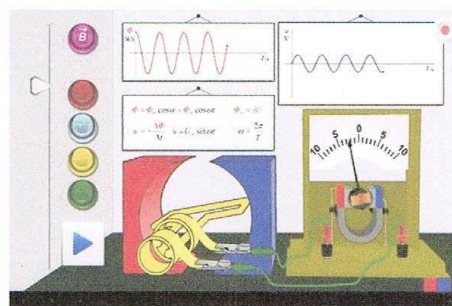
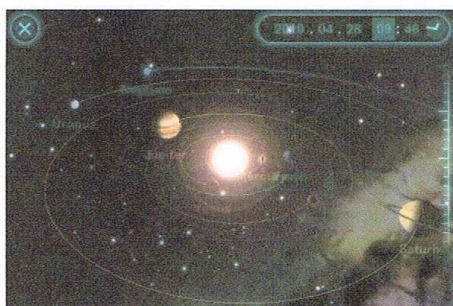


## Testunterlagen

FHNW/APS Olten Masterarbeit B. Röthlisberger	Fragebogen zur Beurteilung von Lern-Apps
---	--

ID	<input type="text"/>
----	----------------------

# Lern-App Test



### Allgemeine Informationen

Liebe Schülerin, lieber Schüler

Zuerst einmal herzlichen Dank, dass du am Test teilnimmst. Du hilfst mit, einen Fragebogen zu entwickeln, mit dem man eine Lern-App beurteilen kann.

Ein solcher Fragebogen hilft den App-Designern, gute und interessante Apps zu entwickeln. Diese Apps sollen die Schülerinnen und Schüler beim Lernen unterstützen und auch Spass machen. An diesem Projekt sind verschiedene Teams von der Fachhochschule Nordwestschweiz beteiligt.

Die Befragung ist anonym. Das heisst, du gibst deinen Namen nicht an. Bei der Auswertung weiss man nicht mehr, wer den Fragebogen ausgefüllt hat.

Vielen Dank

*Beat Röthlisberger*

<b>FHNW/APS Olten</b> Masterarbeit B. Röthlisberger	Fragebogen zur Beurteilung von Lern-Apps
--	--

### Anleitung zum Ausfüllen des Fragebogens

Im Fragebogen stehen verschiedene Aussagen zur App. Kreuze jeweils an, wie stark du dieser Aussage zustimmst. Wenn du gar nicht zustimmst, mache das Kreuz ganz links. Wenn du voll zustimmst, machst du das Kreuz ganz rechts. Du kannst natürlich auch die Zwischenabstufungen verwenden.

### Beispiel

Nr.	Aussage	stimme gar nicht zu	-	-/+	+	stimme voll zu
01	Die App ist nützlich.	<b>X</b>				

Du findest die App überhaupt nicht nützlich. Dann würdest du das Kreuz ganz links setzen.

- ☞ Bitte setze das Kreuz nicht zwischen die Häuschen. Entscheide dich für eine der 5 Abstufungen.
- ☞ Bitte kontrolliere am Schluss, ob du auch alle Aussagen angekreuzt hast.
- ☞ Ich bitte dich, den Fragebogen seriös und ehrlich auszufüllen.
- ☞ Ist soweit alles klar, oder hast du Fragen?

**FHNW/APS Olten**  
Masterarbeit B. Röthlisberger

Fragebogen zur Beurteilung von Lern-Apps

### 1. Teil Solar Walk

Öffne die App Solar Walk.

1. Fliege zum Jupiter und beantworte folgende Frage:

Wie gross ist der Radius des Jupiters? \_\_\_\_\_ km

2. Fliege zum Uranus und beantworte folgende Frage:

Wie hoch ist die Oberflächentemperatur des Uranus? \_\_\_\_\_ °C

3. Jetzt kannst du selbständig weitere Planeten erforschen und anschliessend den Fragebogen ausfüllen.

<b>FHNW/APS Olten</b> Masterarbeit B. Röthlisberger	Fragebogen zur Beurteilung von Lern-Apps
--	--

### Fragebogen zu Solar Walk

Nr.	Aussage	stimme gar nicht zu	-	-/+	+	stimme voll zu
01	Das Design der App ist faszinierend.					
02	Das Design der App ist originell.					
03	Die Spezialeffekte in der App sind gut gemacht.					
04	Das Benützen der App macht Spass.					
05	Die App ist einfach zu gebrauchen.					
06	Es ist einfach in der App zu navigieren.					
07	Ich werde diese App weiter empfehlen.					
08	Bei der Benützung der App vergesse ich meine Umgebung.					
09	Ich werde richtig in das Thema reingezogen.					
10	Die App langweilt mich.					
11	Die Benützung der App ist anspruchsvoll.					
12	Mit der App kann ich die Lernaufgabe lösen.					
13	Die Bedienung der App ist verständlich.					
14	Der Ton passt zum Design.					
15	Der Ton stört mich.					
16	Die App macht mich neugierig auf das Thema (Astronomie).					
17	Ich möchte das Thema am liebsten mit der App weiter erkunden.					
18	Das Thema (Astronomie) interessiert mich.					
19	Die App begeistert mich.					
20	Die App verärgert mich.					
21	Die Benützung der App ist aufregend.					
22	Das Benützen der App regt meine Gedanken zum Thema (Astronomie) an.					



☞ Bitte gehe nicht weiter. Erforsche nochmals das Weltall.

<b>FHNW/APS Olten</b> Masterarbeit B. Röthlisberger	Fragebogen zur Beurteilung von Lern-Apps
--	--

## 2. Teil Physics at school

Öffne die App Physics at school.

Gehe im Menü zu Teil:

IV Molekulare Physik und Thermodynamik

3. Dampflokomotive

1. Öffne die Dampflokomotive und befasse dich mit der Darstellung. Kannst du die Frage beantworten?

Warum fährt die Lokomotive schneller oder langsamer?

---

Gehe im Menü zu Teil:

II Gravitationsfeld

1. Horizontaler Wurf

1. Öffne „Horizontaler Wurf“ und befasse dich mit der Darstellung. Kannst du die Frage beantworten?

Bei welcher Beschleunigung (Kästchen unten links) wird der Affe von den Bananen getroffen?

$v_0 = \underline{\hspace{2cm}} \text{ m s}^{-1}$

2. Jetzt kannst du selbständig weitere Themen erforschen und anschliessend den Fragebogen ausfüllen.

<b>FHNW/APS Olten</b> Masterarbeit B. Röthlisberger	Fragebogen zur Beurteilung von Lern-Apps
--	--

**Fragebogen zu Physics at school**

Nr.	Aussage	stimme gar nicht zu	-	-/+	+	stimme voll zu
01	Das Design der App ist faszinierend.					
02	Das Design der App ist originell.					
03	Die Spezialeffekte in der App sind gut gemacht.					
04	Das Benützen der App macht Spass.					
05	Die App ist einfach zu gebrauchen.					
06	Es ist einfach in der App zu navigieren.					
07	Ich werde diese App weiter empfehlen.					
08	Bei der Benützung der App vergesse ich meine Umgebung.					
09	Ich werde richtig in das Thema reingezogen.					
10	Die App langweilt mich.					
11	Die Benützung der App ist anspruchsvoll.					
12	Mit der App kann ich die Lernaufgabe lösen.					
13	Die Bedienung der App ist verständlich.					
14	Der Ton passt zum Design.					
15	Der Ton stört mich.					
16	Die App macht mich neugierig auf das Thema (Physik).					
17	Ich möchte das Thema am liebsten mit der App weiter erkunden.					
18	Das Thema (Physik) interessiert mich.					
19	Die App begeistert mich.					
20	Die App verärgert mich.					
21	Die Benützung der App ist aufregend.					
22	Das Benützen der App regt meine Gedanken zum Thema (Physik) an.					

**FHNW/APS Olten**  
Masterarbeit B. Röthlisberger

Fragebogen zur Beurteilung von Lern-Apps

### 3. Teil: Fragen zu dir

Wie alt bist du? Ich bin \_\_\_\_\_ Jahre alt.

Ich bin:  männlich  
 weiblich

Wie oft arbeitest du mit einer Lern-App innerhalb einer Woche?  
Auf dem Handy, auf einem Tablet oder mit dem Computer?

nie  einmal in der Woche  mehrmals in der Woche

Jetzt hast du es geschafft!



**KSA-Test Solar Walk**

		AE.Expr.Ac: Das Design der App ist faszinierend.	AE.Expr.Ac: Das Design der App ist originell.	AE.Expr.Ac: Die Spezialeffekte in der App sind gut gemacht.	AE.Plea.Int: Das Benützen der App macht Spass.	SPQ.Usa: Die App ist einfach zu gebrauchen.	SPQ.Usa: Es ist einfach in der App zu navigieren.	SPQ.Loy: Ich werde diese App weiter empfehlen.	UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	UES.Inv: Ich werde richtig in das Thema reingezogen.	R-L.Bor: Die App langweilt mich.	R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	meCu.Nütz: Mit der App kann ich die Lernaufgabe lösen.	meCu.Benütz: Die Bedienung der App ist verständlich.	Ton: Der Ton passt zum Design.	R-Ton: Der Ton stört mich.	In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	In.Mot: Das Thema interessiert mich.	Pan.Emot: Die App begeistert mich.	R-Pan.Emot: Die App verärgert mich.	ARS.PAAtt: Die Benützung der App ist aufregend.	ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.
N		158	158	158	158	158	158	157	157	158	158	156	156	158	158	157	157	158	157	157	156	158	158
Parameter der Normalverteilung <sup>b,c</sup>	Mittelwert	4.42	4.17	4.36	4.25	4.35	4.13	4.00	3.25	3.96	4.34	3.58	4.49	4.36	4.16	4.05	3.80	3.98	3.57	4.04	4.65	3.80	3.66
	Standardabweichung	0.698	0.808	0.733	0.901	0.732	0.875	0.954	1.250	0.950	0.894	1.164	0.741	0.759	0.990	1.208	1.036	0.974	1.145	0.857	0.689	1.031	1.133
Extremste Differenzen	Absolut	0.336	0.245	0.308	0.292	0.298	0.238	0.252	0.158	0.217	0.332	0.185	0.362	0.300	0.256	0.294	0.189	0.255	0.193	0.218	0.446	0.221	0.213
	Positiv	0.202	0.204	0.192	0.202	0.199	0.162	0.162	0.148	0.167	0.231	0.147	0.247	0.200	0.199	0.216	0.174	0.150	0.144	0.177	0.304	0.137	0.124
	Negativ	-0.336	-0.245	-0.308	-0.292	-0.298	-0.238	-0.252	-0.158	-0.217	-0.332	-0.185	-0.362	-0.300	-0.256	-0.294	-0.189	-0.255	-0.193	-0.218	-0.446	-0.221	-0.213
Statistik für Test		0.336	0.245	0.308	0.292	0.298	0.238	0.252	0.158	0.217	0.332	0.185	0.362	0.300	0.256	0.294	0.189	0.255	0.193	0.218	0.446	0.221	0.213
Asymptotische Signifikanz (2-seitig)		.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>

a. Test-App = SolarWalk

b. Die zu testende Verteilung ist eine Normalverteilung.

c. Aus den Daten berechnet.

d. Signifikanzkorrektur nach Lilliefors.



**KSA-Test Physics Lite**

		AE.Expr.Ae: Das Design der App ist faszinierend.	AE.Expr.Ae: Das Design der App ist originell.	AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	AE.Plea.Int: Das Benützen der App macht Spass.	SPQ.Usa: Die App ist einfach zu gebrauchen.	SPQ.Usa: Es ist einfach in der App zu navigieren.	SPQ.Loy: Ich werde diese App weiter empfehlen.	UES.FA: Bei der Benutzung der App vergesse ich meine Umgebung.	UES.Inv: Ich werde richtig in das Thema reingezogen.	R-L.Bor: Die App langweilt mich.	R-UES.Perc:Us: Die Benutzung der App ist anspruchsvoll.	meCu.Nütz: Mit der App kann ich die Lernaufgabe lösen.	meCu.Benütz: Die Bedienung der App ist verständlich.	Ton: Der Ton passt zum Design.	R-Ton: Der Ton stört mich.	In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	In.Mot: Das Thema interessiert mich.	Pan.Emot: Die App begeistert mich.	R-Pan.Emot: Die App verärgert mich.	ARS.PAtt: Die Benutzung der App ist aufregend.	ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.
N		158	158	158	157	157	156	157	157	157	157	157	155	157	155	155	157	157	156	156	155	156	156
Parameter der Normalverteilung <sup>b,c</sup>	Mittelwert	3.04	3.19	3.15	3.25	3.50	3.53	2.85	2.01	2.42	3.11	3.48	3.55	3.55	3.52	3.79	2.66	2.76	3.12	2.73	3.71	2.71	2.64
	Standardabweichung	0.986	1.054	1.188	1.130	1.130	1.098	1.148	1.121	1.127	1.166	1.141	1.135	1.106	1.235	1.190	1.072	1.140	1.180	1.074	1.168	1.072	1.113
Extremste Differenzen	Absolut	0.216	0.190	0.182	0.180	0.207	0.178	0.182	0.243	0.231	0.182	0.171	0.175	0.175	0.181	0.220	0.179	0.183	0.174	0.188	0.201	0.208	0.184
	Positiv	0.214	0.160	0.155	0.154	0.135	0.178	0.182	0.243	0.231	0.182	0.167	0.165	0.174	0.133	0.154	0.179	0.174	0.174	0.188	0.174	0.208	0.168
	Negativ	-0.216	-0.190	-0.182	-0.180	-0.207	-0.174	-0.175	-0.183	-0.130	-0.161	-0.171	-0.175	-0.175	-0.181	-0.220	-0.178	-0.183	-0.160	-0.163	-0.201	-0.145	-0.184
Statistik für Test		0.216	0.190	0.182	0.180	0.207	0.178	0.182	0.243	0.231	0.182	0.171	0.175	0.175	0.181	0.220	0.179	0.183	0.174	0.188	0.201	0.208	0.184
Asymptotische Signifikanz (2-seitig)		.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>	.000 <sup>d</sup>

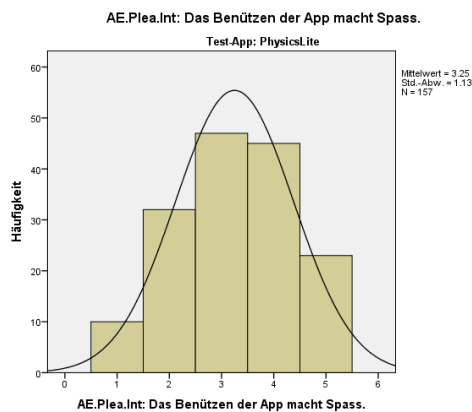
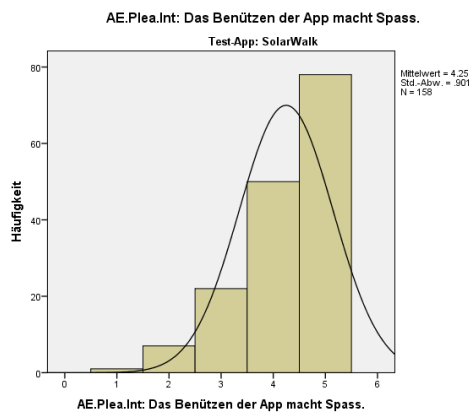
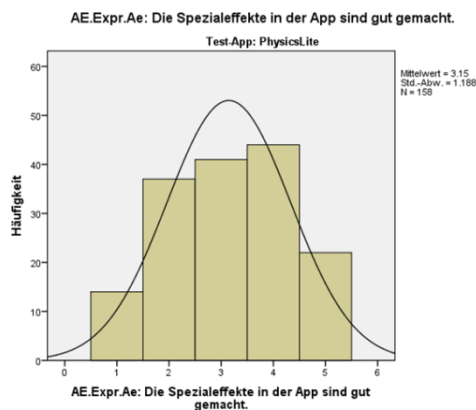
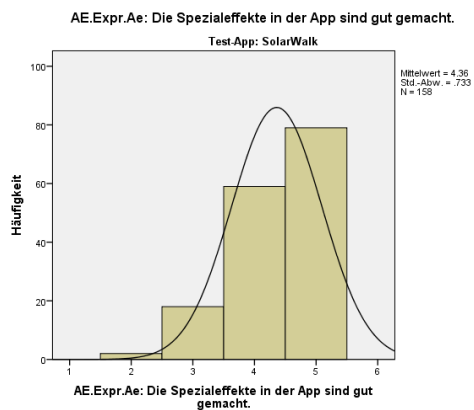
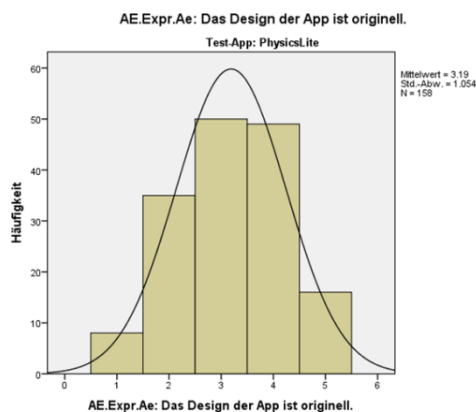
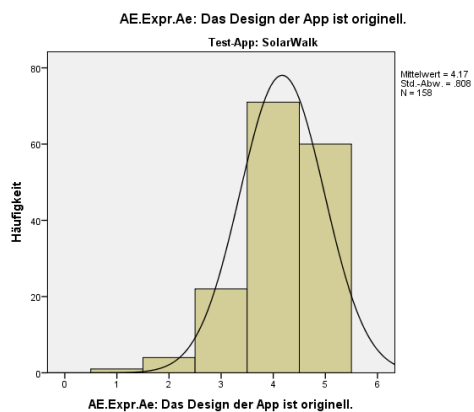
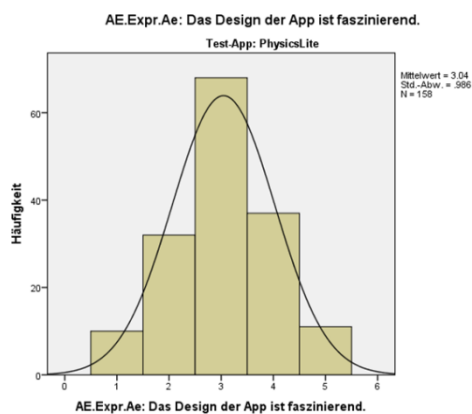
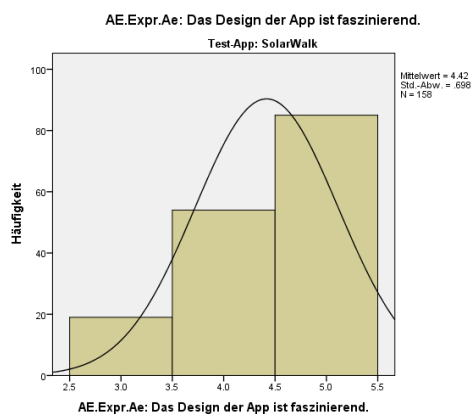
a. Test-App = PhysicsLite

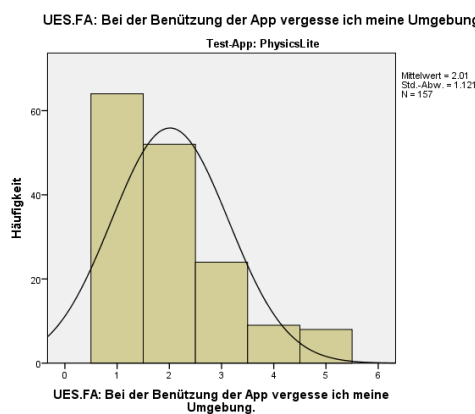
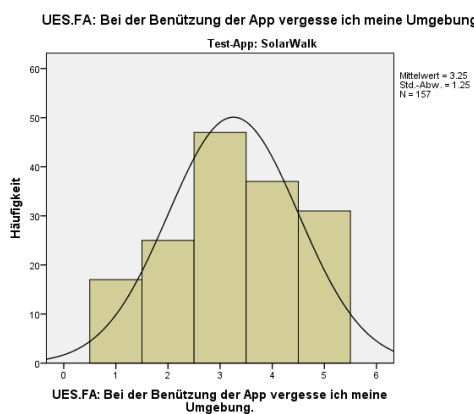
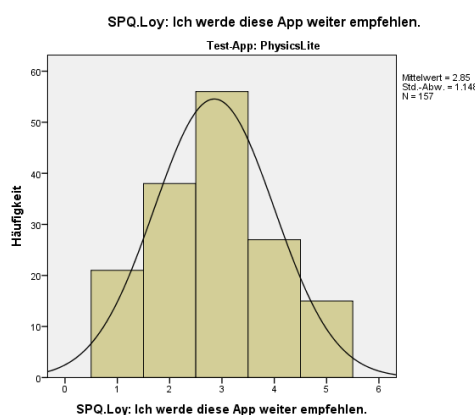
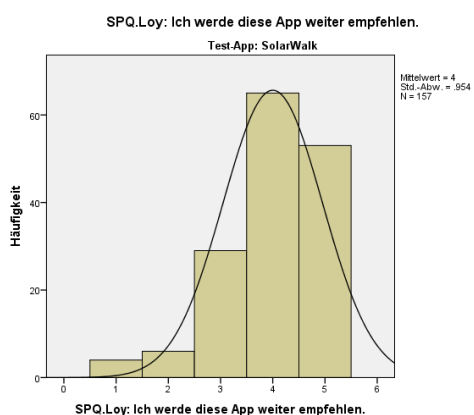
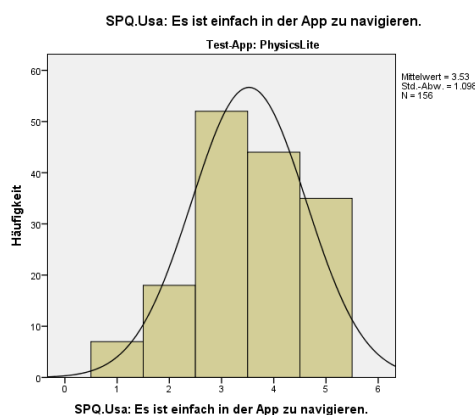
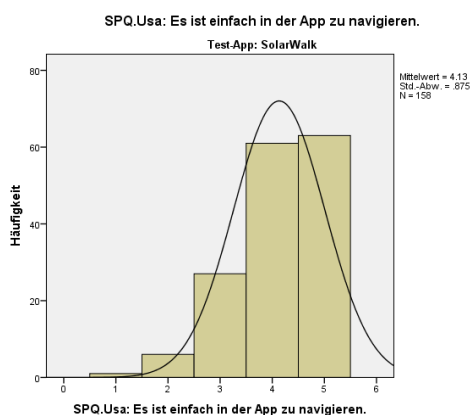
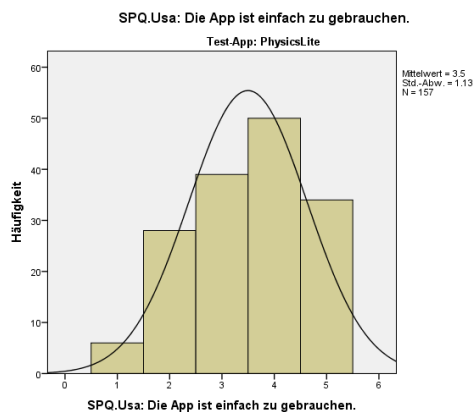
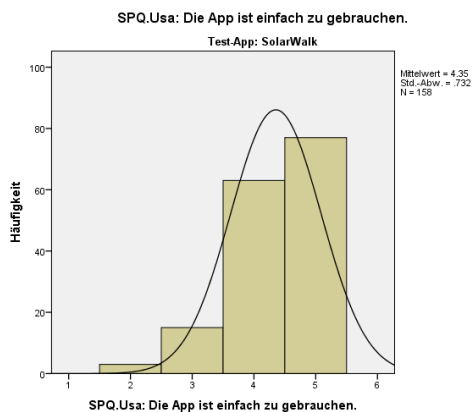
b. Die zu testende Verteilung ist eine Normalverteilung.

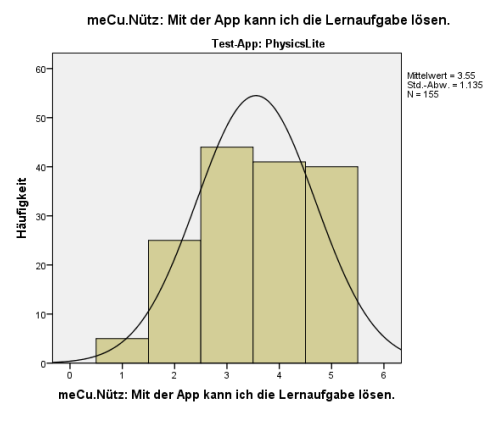
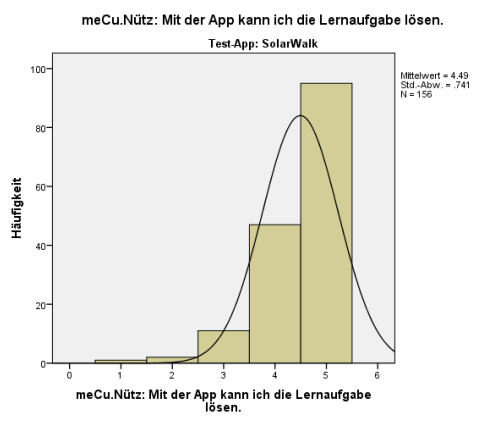
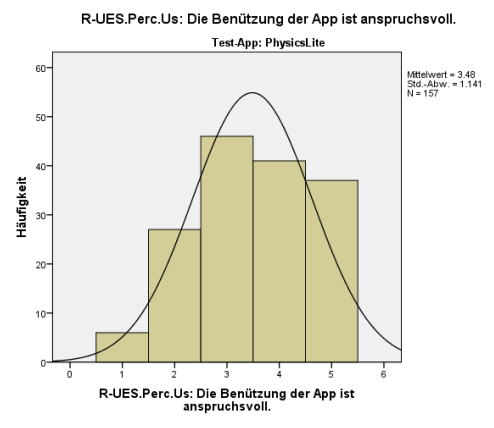
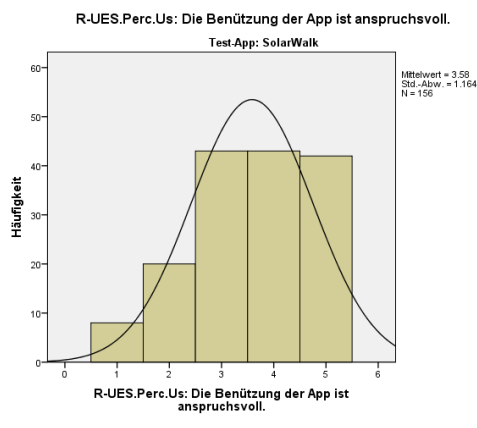
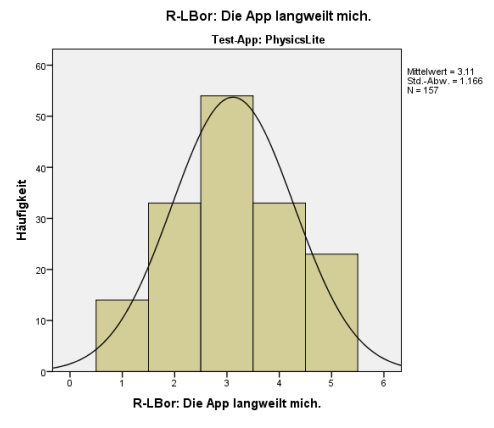
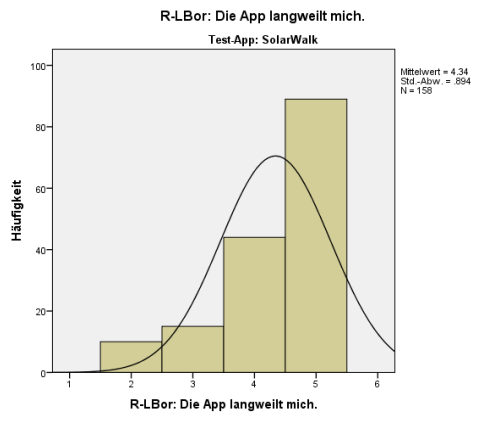
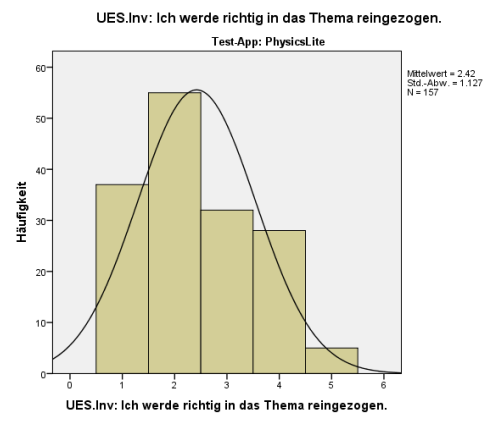
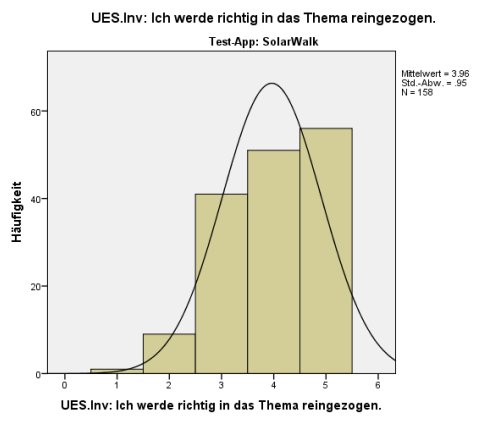
c. Aus den Daten berechnet.

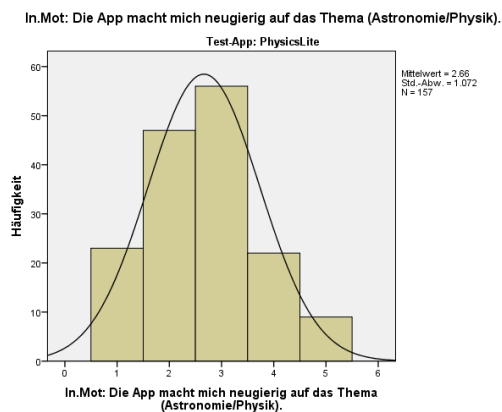
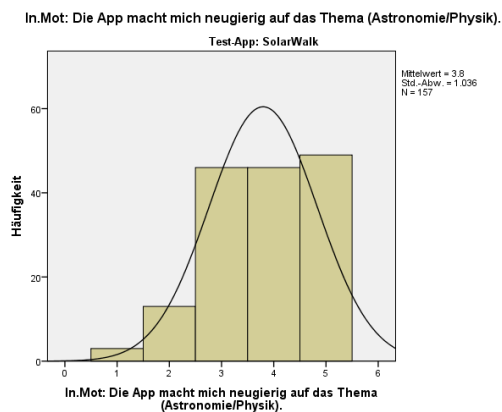
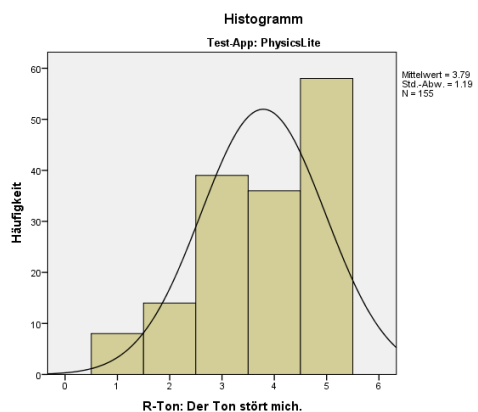
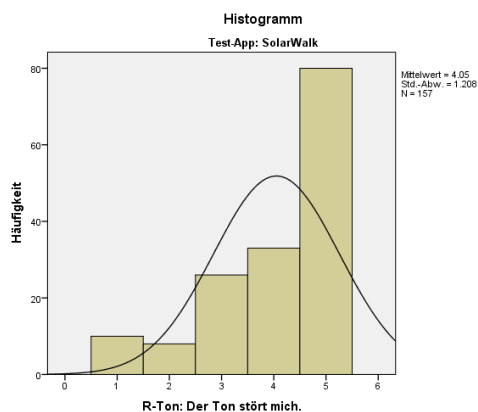
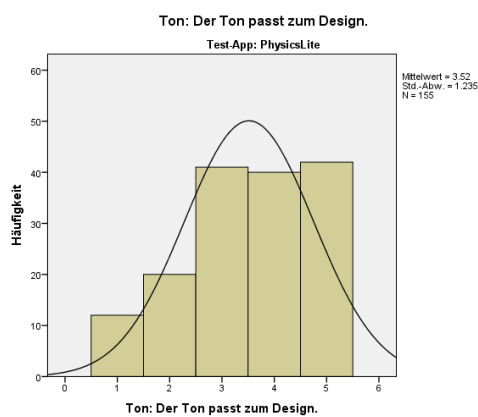
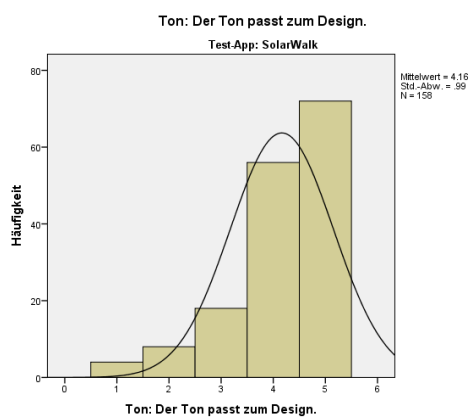
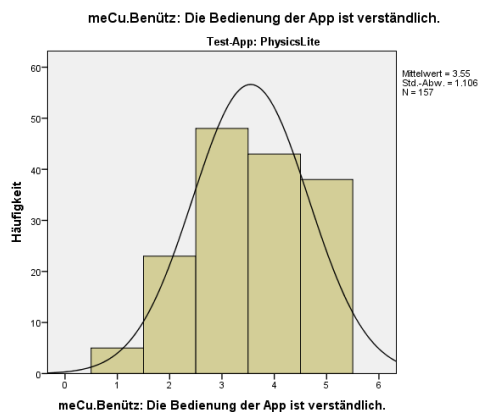
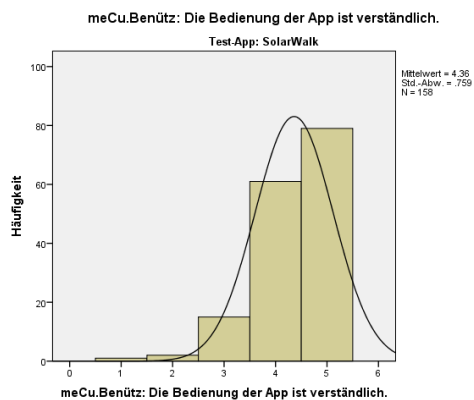
d. Signifikanzkorrektur nach Lilliefors.

## Histogramme

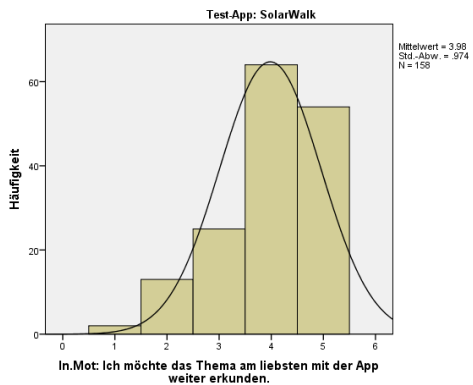




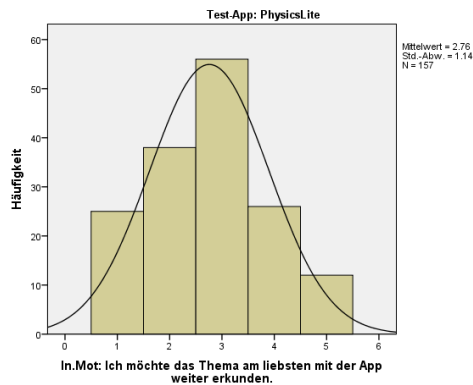




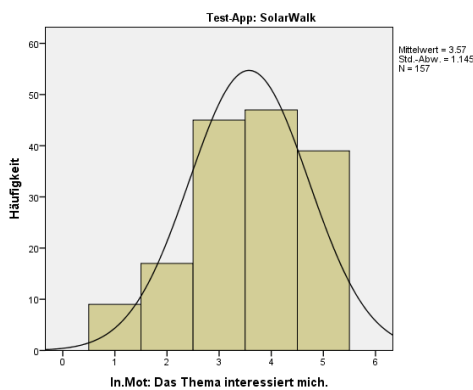
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.



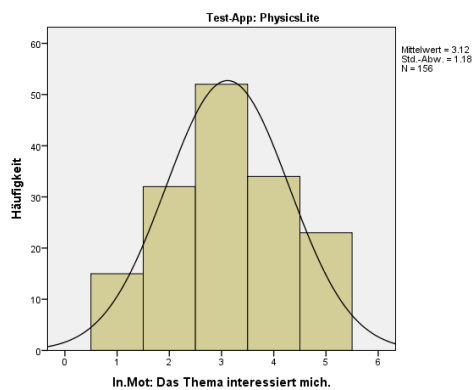
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.



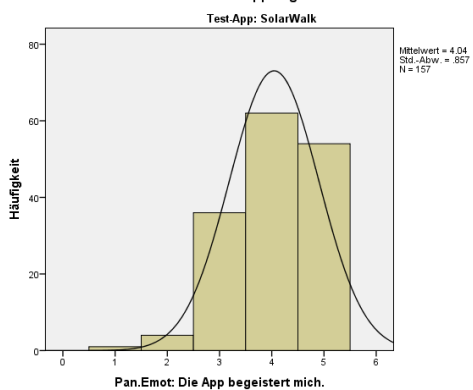
In.Mot: Das Thema interessiert mich.



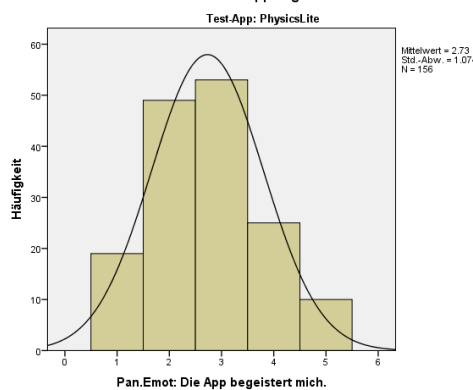
In.Mot: Das Thema interessiert mich.



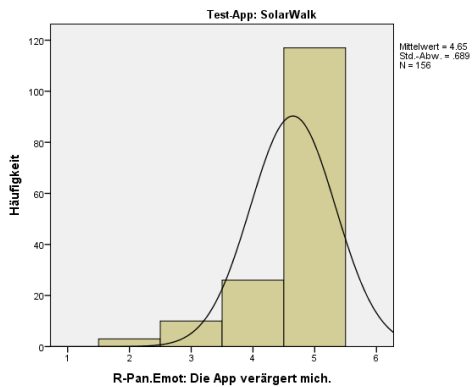
Pan.Emot: Die App begeistert mich.



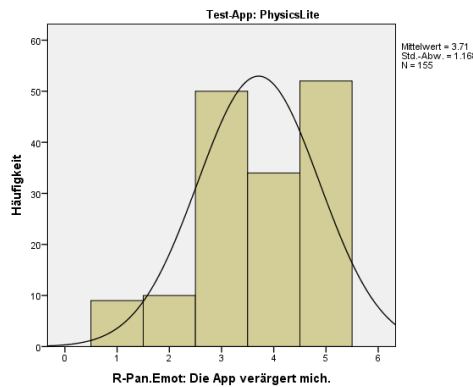
Pan.Emot: Die App begeistert mich.

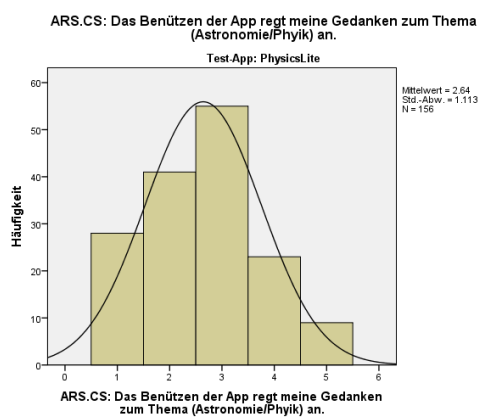
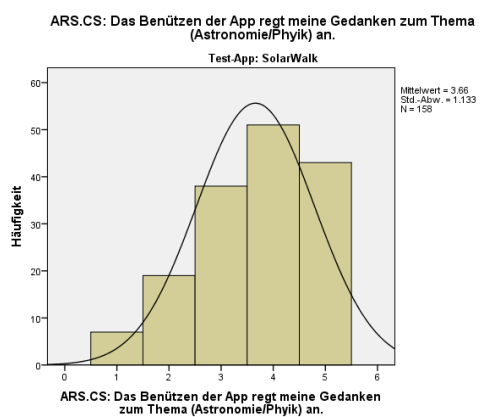
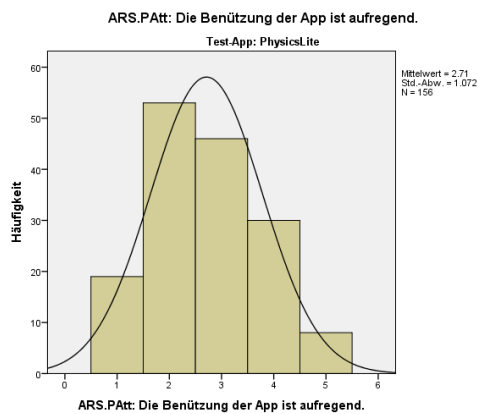
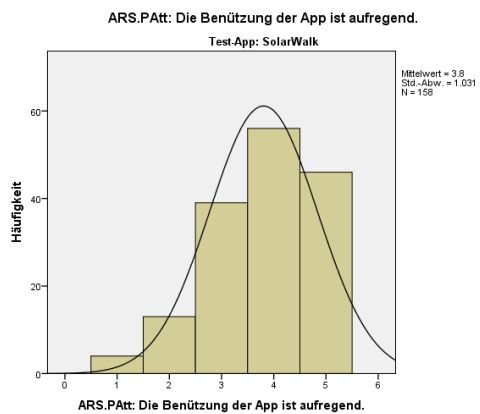


R-Pan.Emot: Die App verärgert mich.



R-Pan.Emot: Die App verärgert mich.





### Mittelwertvergleich Geschlecht Solar Walk

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit			Effektstärke r	Deskriptive Statistik			
		F	Signifikanz	T	df	Signifikanz (2-seitig)		N	MW	SD	
AE.Expr.Ae: Das Design der App ist faszinierend.	Varianzen sind gleich	5.531	0.020	-2.134	156	0.034	.168	männlich	89	4.31	0.748
	Varianzen sind nicht gleich			-2.190	155.642			weiblich	69	4.55	0.607
AE.Expr.Ae: Das Design der App ist originell.	Varianzen sind gleich	1.741	0.189	-2.675	156	0.008	.209	männlich	89	4.02	0.904
	Varianzen sind nicht gleich			-2.801	153.735			weiblich	69	4.36	0.618
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	Varianzen sind gleich	0.476	0.491	-1.562	156	0.120	.124	männlich	89	4.28	0.769
	Varianzen sind nicht gleich			-1.587	153.423			weiblich	69	4.46	0.677
AE.Plea.Int: Das Benützen der App macht Spass.	Varianzen sind gleich	0.046	0.831	-1.424	156	0.157	.113	männlich	89	4.16	0.891
	Varianzen sind nicht gleich			-1.420	145.033			weiblich	69	4.36	0.907
SPQ.Usa: Die App ist einfach zu gebrauchen.	Varianzen sind gleich	0.001	0.971	-1.217	156	0.226	.097	männlich	89	4.29	0.726
	Varianzen sind nicht gleich			-1.214	145.265			weiblich	69	4.43	0.737
SPQ.Usa: Es ist einfach in der App zu navigieren.	Varianzen sind gleich	0.240	0.625	0.580	156	0.563	.046	männlich	89	4.17	0.801
	Varianzen sind nicht gleich			0.567	131.036			weiblich	69	4.09	0.966
SPQ.Loy: Ich werde diese App weiter empfehlen.	Varianzen sind gleich	2.128	0.147	-2.395	155	0.018	.189	männlich	88	3.84	1.016
	Varianzen sind nicht gleich			-2.453	154.672			weiblich	69	4.20	0.833
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	Varianzen sind gleich	0.434	0.511	-0.182	155	0.856	.015	männlich	88	3.24	1.304
	Varianzen sind nicht gleich			-0.184	151.548			weiblich	69	3.28	1.187
UES.Inv: Ich werde richtig in das Thema reingezogen.	Varianzen sind gleich	2.406	0.123	-1.289	156	0.199	.103	männlich	89	3.88	1.009
	Varianzen sind nicht gleich			-1.315	154.466			weiblich	69	4.07	0.863
R-LBor: Die App langweilt mich.	Varianzen sind gleich	2.003	0.159	-1.334	156	0.184	.106	männlich	89	4.26	0.960
	Varianzen sind nicht gleich			-1.366	155.263			weiblich	69	4.45	0.796
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	Varianzen sind gleich	0.152	0.698	-1.495	154	0.137	.120	männlich	87	3.46	1.179
	Varianzen sind nicht gleich			-1.502	148.416			weiblich	69	3.74	1.133
meCu.Nütz: Mit der App kann ich die Lernaufgabe lösen.	Varianzen sind gleich	1.266	0.262	-0.857	154	0.393	.069	männlich	87	4.45	0.803
	Varianzen sind nicht gleich			-0.878	153.886			weiblich	69	4.55	0.654
meCu.Benütz: Die Bedienung der App ist verständlich.	Varianzen sind gleich	0.016	0.899	-1.508	156	0.134	.120	männlich	89	4.28	0.754
	Varianzen sind nicht gleich			-1.507	145.851			weiblich	69	4.46	0.759
Ton: Der Ton passt zum Design.	Varianzen sind gleich	0.182	0.670	-0.752	156	0.453	.060	männlich	89	4.11	1.038
	Varianzen sind nicht gleich			-0.763	152.917			weiblich	69	4.23	0.926
R-Ton: Der Ton stört mich.	Varianzen sind gleich	0.354	0.553	0.062	155	0.951	.005	männlich	89	4.06	1.246
	Varianzen sind nicht gleich			0.062	148.814			weiblich	68	4.04	1.165
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	Varianzen sind gleich	0.001	0.973	0.300	155	0.765	.024	männlich	88	3.82	1.056
	Varianzen sind nicht gleich			0.301	148.616			weiblich	69	3.77	1.017
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	Varianzen sind gleich	0.390	0.533	-1.372	156	0.172	.109	männlich	89	3.89	0.994
	Varianzen sind nicht gleich			-1.382	149.810			weiblich	69	4.10	0.942
In.Mot: Das Thema interessiert mich.	Varianzen sind gleich	0.372	0.543	1.203	155	0.231	.096	männlich	88	3.67	1.172
	Varianzen sind nicht gleich			1.212	149.754			weiblich	69	3.45	1.105
Pan.Emot: Die App begeistert mich.	Varianzen sind gleich	3.258	0.073	-1.695	155	0.092	.135	männlich	89	3.94	0.831
	Varianzen sind nicht gleich			-1.682	139.909			weiblich	68	4.18	0.880
R-Pan.Emot: Die App verärgert mich.	Varianzen sind gleich	4.512	0.035	-1.643	154	0.102	.131	männlich	88	4.57	0.691
	Varianzen sind nicht gleich			-1.647	145.612			weiblich	68	4.75	0.677
ARS.PAtt: Die Benützung der App ist aufregend.	Varianzen sind gleich	0.425	0.515	-0.861	156	0.391	.069	männlich	89	3.74	1.050
	Varianzen sind nicht gleich			-0.865	149.053			weiblich	69	3.88	1.008
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	Varianzen sind gleich	1.458	0.229	0.200	156	0.842	.016	männlich	89	3.67	1.185
	Varianzen sind nicht gleich			0.203	152.325			weiblich	69	3.64	1.071

Anmerkungen . F=Prüfgrösse F, T=Prüfgrösse T, df=Freiheitsgrade, N=Stichprobe, MW= Mittelwert, SD=Standardabweichung

$$\text{Effektstärke } r = \left| \sqrt{\frac{t^2}{t^2 + df}} \right|$$



### Mittelwertvergleich Geschlecht Physics Lite

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit			Effektstärke <i>r</i>	Deskriptive Statistik			
		<i>F</i>	Signifikanz	<i>T</i>	<i>df</i>	Signifikanz (2-seitig)		<i>N</i>	<i>MW</i>	<i>SD</i>	
AE.Expr.Ae: Das Design der App ist faszinierend.	Varianzen sind gleich	1.273	0.261	-2.630	156	0.009	.206	männlich	89	2.86	0.912
	Varianzen sind nicht gleich			-2.593	138.658			weiblich	69	3.27	1.034
AE.Expr.Ae: Das Design der App ist originell.	Varianzen sind gleich	0.825	0.365	-1.949	156	0.053	.154	männlich	89	3.05	1.027
	Varianzen sind nicht gleich			-1.941	145.619			weiblich	69	3.37	1.066
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	Varianzen sind gleich	0.147	0.702	-1.738	156	0.084	.138	männlich	89	3.00	1.174
	Varianzen sind nicht gleich			-1.735	147.333			weiblich	69	3.33	1.188
AE.Plea.Int: Das Benützen der App macht Spass.	Varianzen sind gleich	0.469	0.494	0.907	155	0.366	.073	männlich	89	3.32	1.126
	Varianzen sind nicht gleich			0.906	147.241			weiblich	69	3.16	1.137
SPQ.Usa: Die App ist einfach zu gebrauchen.	Varianzen sind gleich	0.360	0.549	0.252	155	0.802	.020	männlich	89	3.52	1.098
	Varianzen sind nicht gleich			0.250	143.220			weiblich	69	3.47	1.176
SPQ.Usa: Es ist einfach in der App zu navigieren.	Varianzen sind gleich	0.031	0.860	-0.909	154	0.365	.073	männlich	89	3.45	1.102
	Varianzen sind nicht gleich			-0.910	148.081			weiblich	69	3.61	1.094
SPQ.Loy: Ich werde diese App weiter empfehlen.	Varianzen sind gleich	0.250	0.618	-1.297	155	0.197	.104	männlich	88	2.75	1.102
	Varianzen sind nicht gleich			-1.285	142.108			weiblich	69	2.99	1.198
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	Varianzen sind gleich	0.163	0.687	0.843	155	0.400	.068	männlich	88	2.08	1.112
	Varianzen sind nicht gleich			0.842	146.705			weiblich	69	1.93	1.133
UES.Inv: Ich werde richtig in das Thema reingezogen.	Varianzen sind gleich	2.221	0.138	-0.081	155	0.935	.007	männlich	89	2.41	1.052
	Varianzen sind nicht gleich			-0.080	136.723			weiblich	69	2.43	1.223
R-L.Bor: Die App langweilt mich.	Varianzen sind gleich	0.189	0.664	0.416	155	0.678	.033	männlich	89	3.15	1.147
	Varianzen sind nicht gleich			0.414	145.132			weiblich	69	3.07	1.196
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	Varianzen sind gleich	0.772	0.381	-0.578	155	0.564	.046	männlich	87	3.44	1.097
	Varianzen sind nicht gleich			-0.572	141.578			weiblich	69	3.54	1.200
meCu.Nütz: Mit der App kann ich die Lernaufgabe lösen.	Varianzen sind gleich	0.190	0.664	-1.533	153	0.127	.123	männlich	87	3.43	1.133
	Varianzen sind nicht gleich			-1.534	146.175			weiblich	69	3.71	1.126
meCu.Benütz: Die Bedienung der App ist verständlich.	Varianzen sind gleich	0.011	0.917	0.050	155	0.960	.004	männlich	89	3.55	1.097
	Varianzen sind nicht gleich			0.050	146.264			weiblich	69	3.54	1.125
Ton: Der Ton passt zum Design.	Varianzen sind gleich	0.653	0.420	-0.505	153	0.614	.041	männlich	89	3.47	1.278
	Varianzen sind nicht gleich			-0.509	150.774			weiblich	69	3.57	1.187
R-Ton: Der Ton stört mich.	Varianzen sind gleich	1.102	0.295	-1.735	153	0.085	.139	männlich	89	3.64	1.126
	Varianzen sind nicht gleich			-1.716	138.559			weiblich	68	3.97	1.248
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	Varianzen sind gleich	0.010	0.919	-2.220	155	0.028	.176	männlich	88	2.49	1.010
	Varianzen sind nicht gleich			-2.196	140.887			weiblich	69	2.87	1.115
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	Varianzen sind gleich	0.000	0.992	-1.548	155	0.124	.123	männlich	89	2.63	1.111
	Varianzen sind nicht gleich			-1.540	144.831			weiblich	69	2.91	1.164
In.Mot: Das Thema interessiert mich.	Varianzen sind gleich	0.374	0.542	0.676	154	0.500	.054	männlich	88	3.17	1.193
	Varianzen sind nicht gleich			0.678	147.300			weiblich	69	3.04	1.169
Pan.Emot: Die App begeistert mich.	Varianzen sind gleich	0.041	0.839	-0.536	154	0.593	.043	männlich	89	2.69	1.038
	Varianzen sind nicht gleich			-0.531	140.394			weiblich	68	2.78	1.123
R-Pan.Emot: Die App verärgert mich.	Varianzen sind gleich	0.040	0.843	-1.675	153	0.096	.134	männlich	88	3.57	1.143
	Varianzen sind nicht gleich			-1.669	143.618			weiblich	68	3.88	1.182
ARS.PAtt: Die Benützung der App ist aufregend.	Varianzen sind gleich	1.324	0.252	-1.496	154	0.137	.120	männlich	89	2.60	1.017
	Varianzen sind nicht gleich			-1.478	138.417			weiblich	69	2.86	1.128
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	Varianzen sind gleich	0.006	0.940	-1.273	154	0.205	.102	männlich	89	2.54	1.087
	Varianzen sind nicht gleich			-1.266	142.771			weiblich	69	2.77	1.139

$$\text{Effektstärke } r = \left| \sqrt{\frac{t^2}{t^2 + df}} \right|$$

Anmerkungen. F=Prüfgrösse F, T=Prüfgrösse T, df=Freiheitsgrade, N=Stichprobe, MW= Mittelwert, SD=Standardabweichung

## Mittelwertvergleich Geschlecht weitere Variablen

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit			Deskriptive Statistik			
		<i>F</i>	Signifikanz	<i>T</i>	<i>df</i>	Signifikanz (2-seitig)	Geschl.	<i>N</i>	<i>MW</i>	<i>SD</i>
Lernaufgabe Solar Walk	Varianzen sind gleich	0.130	0.719	0.123	156	0.902	männlich	89	1.82	0.415
	Varianzen sind nicht gleich			0.122			137.771	0.903	weiblich	69
Lernaufgabe Physics Lite	Varianzen sind gleich	1.664	0.199	-1.157	156	0.249	männlich	89	1.44	0.641
	Varianzen sind nicht gleich			-1.170			153.312	0.244	weiblich	69
App-Nutzung pro Woche	Varianzen sind gleich	0.279	0.598	-0.235	154	0.815	männlich	88	1.55	0.757
	Varianzen sind nicht gleich			-0.236			147.542	0.813	weiblich	68

Anmerkungen. *F*=Prüfgrösse *F*, *T*=Prüfgrösse *T*, *df*=Freiheitsgrade, *N*=Stichprobe, *MW*= Mittelwert, *SD*=Standardabweichung

### Korrelationen Solar Walk

	AE.Expr.Ae: Das Design der App ist faszinierend.	AE.Expr.Ae: Das Design der App ist originell.	AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	AE.Plea.Int: Das Benützen der App macht Spass.	SPQ.Usa: Die App ist einfach zu gebrauchen.	SPQ.Usa: Es ist einfach in der App zu navigieren.	UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	UES.Inv: Ich werde richtig in das Thema reingezogen.	R-LBor: Die App langweilt mich.	R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	meCu.Benütz: Die Bedienung der App ist verständlich.	In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	R-Pan.Emot: Die App verärgert mich.	Pan.Emot: Die App begeistert mich.	ARS.PAtt: Die Benützung der App ist aufregend.	ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.
AE.Expr.Ae: Das Design der App ist faszinierend.	-																
AE.Expr.Ae: Das Design der App ist originell.	.573																
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	.351	.325															
AE.Plea.Int: Das Benützen der App macht Spass.	.392	.353	.260														
SPQ.Usa: Die App ist einfach zu gebrauchen.	.195	.285	.152	.330													
SPQ.Usa: Es ist einfach in der App zu navigieren.	.044	.247	.133	.289	.483												
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	.340	.324	.092	.276	.110	.067											
UES.Inv: Ich werde richtig in das Thema reingezogen.	.187	.216	.111	.428	.148	.129	.534										
R-LBor: Die App langweilt mich.	.270	.157	.180	.575	.164	.202	.233	.435									
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	-.020	.107	.025	.008	.176	.147	-.241	-.057	.060								
meCu.Benütz: Die Bedienung der App ist verständlich.	.111	.221	.062	.288	.514	.416	-.064	.107	.117	.180							
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	.179	.193	.247	.431	.164	.166	.262	.523	.437	-.047	.070						
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	.218	.231	.188	.441	.081	.227	.315	.474	.476	-.046	.113	.623					
R-Pan.Emot: Die App verärgert mich.	.181	.152	.183	.568	.225	.278	.102	.312	.517	.090	.312	.299	.381				
Pan.Emot: Die App begeistert mich.	.364	.441	.167	.490	.260	.247	.317	.355	.380	-.044	.161	.328	.444	.356			
ARS.PAtt: Die Benützung der App ist aufregend.	.336	.285	.246	.368	.008	.191	.390	.486	.350	.043	.181	.439	.504	.356	.429		
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	.190	.169	.203	.283	.070	.117	.418	.526	.292	-.004	-.041	.555	.439	.213	.266	.411	-

**Signifikanzen (1-seitig) der Korrelationen Solar Walk**

	AE.Expr.Ae: Das Design der App ist faszinierend.	AE.Expr.Ae: Das Design der App ist originell.	AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	AE.Plea.Int: Das Benützen der App macht Spass.	SPQ.Usa: Die App ist einfach zu gebrauchen.	SPQ.Usa: Es ist einfach in der App zu navigieren.	UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	UES.Inv: Ich werde richtig in das Thema reingezogen.	R-LBor: Die App langweilt mich.	R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	meCu.Benütz: Die Bedienung der App ist verständlich.	In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	R-Pan.Emot: Die App verärgert mich.	Pan.Emot: Die App begeistert mich.	ARS.PAtt: Die Benützung der App ist aufregend.	ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.
AE.Expr.Ae: Das Design der App ist faszinierend.	-																
AE.Expr.Ae: Das Design der App ist originell.	.000	-															
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	.000	.000	-														
AE.Plea.Int: Das Benützen der App macht Spass.	.000	.000	.000	-													
SPQ.Usa: Die App ist einfach zu gebrauchen.	.007	.000	.028	.000	-												
SPQ.Usa: Es ist einfach in der App zu navigieren.	.291	.001	.048	.000	.000	-											
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	.000	.000	.125	.000	.085	.203	-										
UES.Inv: Ich werde richtig in das Thema reingezogen.	.009	.003	.082	.000	.032	.054	.000	-									
R-LBor: Die App langweilt mich.	.000	.025	.012	.000	.020	.005	.002	.000	-								
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	.403	.091	.377	.462	.014	.033	.001	.237	.229	-							
meCu.Benütz: Die Bedienung der App ist verständlich.	.083	.003	.219	.000	.000	.000	.212	.090	.071	.012	-						
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	.012	.008	.001	.000	.020	.019	.000	.000	.000	.280	.191	-					
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	.003	.002	.009	.000	.156	.002	.000	.000	.000	.285	.079	.000	-				
R-Pan.Emot: Die App verärgert mich.	.011	.029	.011	.000	.002	.000	.102	.000	.000	.133	.000	.000	.000	-			
Pan.Emot: Die App begeistert mich.	.000	.000	.018	.000	.000	.001	.000	.000	.000	.291	.022	.000	.000	.000	-		
ARS.PAtt: Die Benützung der App ist aufregend.	.000	.000	.001	.000	.459	.008	.000	.000	.000	.296	.012	.000	.000	.000	.000	-	
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.	.008	.017	.005	.000	.190	.072	.000	.000	.000	.478	.305	.000	.000	.004	.000	.000	-

### Anti-Image Matrix Solar Walk

	AE.Expr.Ae: Das Design der App ist faszinierend.	AE.Expr.Ae: Das Design der App ist originell.	AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	AE.Plea.Int: Das Benützen der App macht Spass.	SPQ.Usa: Die App ist einfach zu gebrauchen.	SPQ.Usa: Es ist einfach in der App zu navigieren.	UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	UES.Inv: Ich werde richtig in das Thema reingezogen.	R-LBor: Die App langweilt mich.	R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	meCu.Benütz: Die Bedienung der App ist verständlich.	In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	R-Pan.Emot: Die App verärgert mich.	Pan.Emot: Die App begeistert mich.	ARS.PAtt: Die Benützung der App ist aufregend.	ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.
AE.Expr.Ae: Das Design der App ist faszinierend.	.766 <sup>b</sup>																
AE.Expr.Ae: Das Design der App ist originell.	-.421	.765 <sup>b</sup>															
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	-.170	-.167	.824 <sup>b</sup>														
AE.Plea.Int: Das Benützen der App macht Spass.	-.155	-.058	-.050	.902 <sup>b</sup>													
SPQ.Usa: Die App ist einfach zu gebrauchen.	-.107	.024	-.080	-.075	.651 <sup>b</sup>												
SPQ.Usa: Es ist einfach in der App zu navigieren.	.218	-.142	-.033	-.036	-.331	.788 <sup>b</sup>											
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	-.121	-.178	.114	-.018	-.164	.003	.745 <sup>b</sup>										
UES.Inv: Ich werde richtig in das Thema reingezogen.	.143	-.021	.073	-.098	-.038	.096	-.348	.874 <sup>b</sup>									
R-LBor: Die App langweilt mich.	-.136	.150	-.007	-.254	.029	-.055	-.016	-.147	.879 <sup>b</sup>								
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	.066	-.195	.049	.055	-.166	-.015	.298	-.007	-.121	.418 <sup>b</sup>							
meCu.Benütz: Die Bedienung der App ist verständlich.	.009	-.102	.097	-.085	-.406	-.151	.200	-.080	.089	.000	.689 <sup>b</sup>						
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	.081	-.042	-.090	-.115	-.162	.061	.180	-.172	-.105	.110	.080	.826 <sup>b</sup>					
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	.019	-.011	.015	-.015	.140	-.109	-.053	-.026	-.136	.030	-.038	-.378	.894 <sup>b</sup>				
R-Pan.Emot: Die App verärgert mich.	.050	.078	-.052	-.307	.005	-.043	.088	-.015	-.241	-.029	-.148	.078	-.072	.873 <sup>b</sup>			
Pan.Emot: Die App begeistert mich.	-.021	-.259	.102	-.152	-.161	-.014	.023	-.009	-.066	.124	.106	.062	-.170	-.065	.885 <sup>b</sup>		
ARS.PAtt: Die Benützung der App ist aufregend.	-.174	.067	-.121	.059	.325	-.112	-.190	-.159	.045	-.160	-.219	-.123	-.129	-.114	-.199	.827 <sup>b</sup>	
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Phyik) an.	-.035	.061	-.084	.036	.019	-.056	-.197	-.203	.060	-.117	.111	-.318	-.044	-.036	-.007	-.054	.867 <sup>b</sup>

b. Maß der Stichprobeneignung=MSA-Koeffizient

### Korrelationen Physics Lite

	AE.Expr.Ae: Das Design der App ist faszinierend.	AE.Expr.Ae: Das Design der App ist originell.	AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	AE.Plea.Int: Das Benützen der App macht Spass.	SPQ.Usa: Die App ist einfach zu gebrauchen.	SPQ.Usa: Es ist einfach in der App zu navigieren.	UES.FA: Bei der Benutzung der App vergesse ich meine Umgebung.	UES.Inv: Ich werde richtig in das Thema reingezogen.	R-LBor: Die App langweilt mich.	R-UES.Perc.Us: Die Benutzung der App ist anspruchsvoll.	meCu.Benütz: Die Bedienung der App ist verständlich.	In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	R-Pan.Emot: Die App verärgert mich.	Pan.Emot: Die App begeistert mich.	ARS.PAtt: Die Benutzung der App ist aufregend.	ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.
AE.Expr.Ae: Das Design der App ist faszinierend.	-																
AE.Expr.Ae: Das Design der App ist originell.	.635																
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	.582	.548															
AE.Plea.Int: Das Benützen der App macht Spass.	.467	.348	.514														
SPQ.Usa: Die App ist einfach zu gebrauchen.	.235	.196	.262	.259													
SPQ.Usa: Es ist einfach in der App zu navigieren.	.286	.281	.336	.394	.684												
UES.FA: Bei der Benutzung der App vergesse ich meine Umgebung.	.370	.274	.320	.367	.197	.227											
UES.Inv: Ich werde richtig in das Thema reingezogen.	.468	.370	.453	.521	.293	.334	.503										
R-LBor: Die App langweilt mich.	.385	.326	.541	.440	.146	.214	.185	.358									
R-UES.Perc.Us: Die Benutzung der App ist anspruchsvoll.	.125	.058	.096	.050	.409	.362	-.010	.055	-.003								
meCu.Benütz: Die Bedienung der App ist verständlich.	.279	.231	.306	.316	.617	.619	.103	.277	.244	.388							
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	.423	.384	.487	.493	.245	.283	.292	.569	.447	.061	.271						
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	.469	.404	.480	.584	.238	.401	.424	.558	.421	.150	.345	.646					
R-Pan.Emot: Die App verärgert mich.	.335	.261	.340	.328	.251	.210	.078	.356	.410	.132	.301	.395	.365				
Pan.Emot: Die App begeistert mich.	.549	.431	.512	.600	.236	.310	.382	.582	.491	.069	.307	.506	.638	.500			
ARS.PAtt: Die Benutzung der App ist aufregend.	.516	.425	.389	.407	.201	.270	.385	.559	.353	.191	.289	.523	.571	.338	.538		
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Phyik) an.	.336	.354	.383	.359	.237	.285	.305	.598	.266	.053	.262	.580	.606	.171	.474	.535	-

**Signifikanzen (1-seitig) der Korrelationen Physics Lite**

	AE.Expr.Ae: Das Design der App ist faszinierend.	AE.Expr.Ae: Das Design der App ist originell.	AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	AE.Plea.Int: Das Benützen der App macht Spass.	SPQ.Usa: Die App ist einfach zu gebrauchen.	SPQ.Usa: Es ist einfach in der App zu navigieren.	UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	UES.Inv: Ich werde richtig in das Thema reingezogen.	R-LBor: Die App langweilt mich.	R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	meCu.Benütz: Die Bedienung der App ist verständlich.	In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	R-Pan.Emot: Die App verärgert mich.	Pan.Emot: Die App begeistert mich.	ARS.PAtt: Die Benützung der App ist aufregend.	ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.
AE.Expr.Ae: Das Design der App ist faszinierend.	-																
AE.Expr.Ae: Das Design der App ist originell.	.000	-															
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	.000	.000	-														
AE.Plea.Int: Das Benützen der App macht Spass.	.000	.000	.000	-													
SPQ.Usa: Die App ist einfach zu gebrauchen.	.001	.007	.000	.001	-												
SPQ.Usa: Es ist einfach in der App zu navigieren.	.000	.000	.000	.000	.000	-											
UES.FA: Bei der Benützung der App vergesse ich meine Umgebung.	.000	.000	.000	.000	.007	.002	-										
UES.Inv: Ich werde richtig in das Thema reingezogen.	.000	.000	.000	.000	.000	.000	.000	-									
R-LBor: Die App langweilt mich.	.000	.000	.000	.000	.034	.004	.010	.000	-								
R-UES.Perc.Us: Die Benützung der App ist anspruchsvoll.	.058	.234	.115	.266	.000	.000	.451	.247	.483	-							
meCu.Benütz: Die Bedienung der App ist verständlich.	.000	.002	.000	.000	.000	.000	.100	.000	.001	.000	-						
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	.000	.000	.000	.000	.001	.000	.000	.000	.000	.224	.000	-					
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	.000	.000	.000	.000	.001	.000	.000	.000	.000	.031	.000	.000	-				
R-Pan.Emot: Die App verärgert mich.	.000	.000	.000	.000	.001	.004	.165	.000	.000	.049	.000	.000	.000	-			
Pan.Emot: Die App begeistert mich.	.000	.000	.000	.000	.001	.000	.000	.000	.000	.194	.000	.000	.000	.000	-		
ARS.PAtt: Die Benützung der App ist aufregend.	.000	.000	.000	.000	.006	.000	.000	.000	.000	.008	.000	.000	.000	.000	.000	-	
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Phyik) an.	.000	.000	.000	.000	.001	.000	.000	.000	.000	.253	.000	.000	.000	.017	.000	.000	-

### Anti-Image Matrix Physics Lite

	AE.Expr.Ae: Das Design der App ist faszinierend.	AE.Expr.Ae: Das Design der App ist originell.	AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	AE.Plea.Int: Das Benützen der App macht Spass.	SPQ.Usa: Die App ist einfach zu gebrauchen.	SPQ.Usa: Es ist einfach in der App zu navigieren.	UES.FA: Bei der Benutzung der App vergesse ich meine Umgebung.	UES.Inv: Ich werde richtig in das Thema reingezogen.	R-LBor: Die App langweilt mich.	R-UES.Perc.Us: Die Benutzung der App ist anspruchsvoll.	meCu.Benütz: Die Bedienung der App ist verständlich.	In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	R-Pan.Emot: Die App verärgert mich.	Pan.Emot: Die App begeistert mich.	ARS.PAtt: Die Benutzung der App ist aufregend.	ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Physik) an.
AE.Expr.Ae: Das Design der App ist faszinierend.	.907 <sup>b</sup>																
AE.Expr.Ae: Das Design der App ist originell.	-0.407	.895 <sup>b</sup>															
AE.Expr.Ae: Die Spezialeffekte in der App sind gut gemacht.	-0.205	-0.236	.922 <sup>b</sup>														
AE.Plea.Int: Das Benützen der App macht Spass.	-0.074	0.070	-0.140	.933 <sup>b</sup>													
SPQ.Usa: Die App ist einfach zu gebrauchen.	-0.021	0.021	-0.004	0.024	.793 <sup>b</sup>												
SPQ.Usa: Es ist einfach in der App zu navigieren.	0.051	-0.096	-0.034	-0.149	-0.462	.853 <sup>b</sup>											
UES.FA: Bei der Benutzung der App vergesse ich meine Umgebung.	-0.081	0.011	-0.066	-0.031	-0.146	0.000	.845 <sup>b</sup>										
UES.Inv: Ich werde richtig in das Thema reingezogen.	-0.058	0.047	-0.014	-0.139	-0.044	-0.042	-0.298	.917 <sup>b</sup>									
R-LBor: Die App langweilt mich.	0.014	0.028	-0.306	-0.072	0.040	0.002	0.059	0.016	.917 <sup>b</sup>								
R-UES.Perc.Us: Die Benutzung der App ist anspruchsvoll.	-0.053	0.075	-0.049	0.090	-0.204	-0.086	0.108	0.048	0.104	.767 <sup>b</sup>							
meCu.Benütz: Die Bedienung der App ist verständlich.	-0.029	0.019	-0.023	-0.032	-0.303	-0.264	0.128	0.022	-0.059	-0.133	.897 <sup>b</sup>						
In.Mot: Die App macht mich neugierig auf das Thema (Astronomie/Physik).	0.000	-0.021	-0.094	-0.092	-0.088	0.066	0.083	-0.141	-0.122	0.057	0.042	.927 <sup>b</sup>					
In.Mot: Ich möchte das Thema am liebsten mit der App weiter erkunden.	0.001	-0.006	0.020	-0.189	0.190	-0.172	-0.198	0.087	-0.030	-0.118	-0.061	-0.280	.906 <sup>b</sup>				
R-Pan.Emot: Die App verärgert mich.	-0.007	-0.012	-0.012	0.057	-0.123	0.093	0.198	-0.136	-0.141	-0.022	-0.081	-0.145	-0.063	.853 <sup>b</sup>			
Pan.Emot: Die App begeistert mich.	-0.142	-0.020	-0.019	-0.225	0.004	0.032	-0.055	-0.114	-0.129	0.043	-0.010	0.106	-0.207	-0.279	.932 <sup>b</sup>		
ARS.PAtt: Die Benutzung der App ist aufregend.	-0.180	-0.080	0.112	0.033	0.099	0.024	-0.126	-0.129	-0.062	-0.187	-0.062	-0.102	-0.083	-0.067	-0.080	.934 <sup>b</sup>	
ARS.CS: Das Benützen der App regt meine Gedanken zum Thema (Astronomie/Phyik) an.	0.112	-0.075	-0.071	0.138	-0.088	0.027	0.127	-0.312	0.080	0.091	-0.035	-0.214	-0.280	0.240	-0.102	-0.193	.859 <sup>b</sup>

b. Maß der Stichprobeneignung=MSA-Koeffizient



---

**Wilcoxon-Test für abhängige Stichproben**


---

	<b>Z</b>	<b>Asymptotische Signifikanz (2-seitig)</b>
1-SW - 1-PHL	-9.565 <sup>b</sup>	0.000
2-SW - 2-PHL	-7.592 <sup>b</sup>	0.000
3-SW - 3-PHL	-8.796 <sup>b</sup>	0.000
4-SW - 4-PHL	-7.574 <sup>b</sup>	0.000
5-SW - 5-PHL	-7.094 <sup>b</sup>	0.000
6-SW - 6-PHL	-5.171 <sup>b</sup>	0.000
7-SW - 7-PHL	-8.437 <sup>b</sup>	0.000
8-SW - 8-PHL	-8.092 <sup>b</sup>	0.000
9-SW - 9-PHL	-9.118 <sup>b</sup>	0.000
Rec10SW - Rec10PHL	-7.454 <sup>b</sup>	0.000
rec11SW - rec11PHL	-.702 <sup>b</sup>	0.483
12-SW - 12-PHL	-7.677 <sup>b</sup>	0.000
13-SW - 13-PHL	-7.098 <sup>b</sup>	0.000
14-SW - 14-PHL	-5.146 <sup>b</sup>	0.000
Rec15SW - Rec15PHL	-3.015 <sup>b</sup>	0.003
16-SW - 16-PHL	-8.001 <sup>b</sup>	0.000
17-SW - 17-PHL	-8.791 <sup>b</sup>	0.000
18-SW - 18-PHL	-3.882 <sup>b</sup>	0.000
19-SW - 19-PHL	-8.991 <sup>b</sup>	0.000
Rec20SW - Rec20PHL	-7.030 <sup>b</sup>	0.000
21-SW - 21-PHL	-8.182 <sup>b</sup>	0.000
22-SW - 22-PHL	-7.363 <sup>b</sup>	0.000

---

*Anmerkungen.* SW=Solar Walk, PHL=Physics Lite, Z=Prüfgrösse Z

b. Basiert auf negativen Rängen.

## Fragebogen zur Erfassung des ästhetischen Empfindens

Items	Cronbachs Alpha	Item Nr.
<b>Ästhetisches Empfinden (Gesamte Skala)</b>	<b>(<math>\alpha=.86-.91</math>)</b>	
<b>Dimension Kognitive Stimulation</b>	<b>(<math>\alpha=.83-.86</math>)</b>	
Bei der Benützung der App vergesse ich meine Umgebung.		1
Ich werde richtig in das Thema reingezogen.		2
Die App macht mich neugierig auf das Thema (xxxxx).		3
Ich möchte das Thema am liebsten mit der App weiter erkunden.		4
Das Benützen der App regt meine Gedanken zum Thema (xxxxx) an.		5
Die Benützung der App ist aufregend.		6
<b>Dimension Expressive Ästhetik</b>	<b>(<math>\alpha=.68-.86</math>)</b>	
Das Design der App ist faszinierend.		7
Das Design der App ist originell.		8
Die Spezialeffekte in der App sind gut gemacht.		9
<b>Dimension Emotion</b>	<b>(<math>\alpha=.77-.79</math>)</b>	
Das Benützen der App macht Spass.		10
Die App langweilt mich.		11
Die App begeistert mich.		12
Die App verärgert mich.		13
<b>Dimension Usability</b>	<b>(<math>\alpha=.74-.84</math>)</b>	
Die App ist einfach zu gebrauchen.		14
Es ist einfach in der App zu navigieren.		15
Die Bedienung der App ist verständlich.		16
<b>Kontrollskala</b>		
Ich werde diese App weiter empfehlen.		17
Mit der App kann ich die Lernaufgabe lösen.		18
Das Thema interessiert mich.		19

*Anmerkungen.* (xxxxx) = Thema, das in der App behandelt wird.