Universität Bremen

– Fachbereich 1 –

Erklär’s doch mal richtig! – Validierung eines Qualitätsmaß-stabs für Erklärvideos

Masterarbeit

Abgabetermin
xx.02.2016

vorgelegt von:
Johannes Block
2549885

betreut von:
ErstgutachterIn:
Herr Dr. rer. nat. Christoph Kulgemeyer
ZweitgutachterIn:
Herr Prof. Dr. rer. nat. habil. Horst Schecker

Zusammenfassung (Abstract)

Das Ziel der vorliegenden Arbeit bestand darin den Qualitätsmaßstab für „gutes Erklären“ von Kulgemeyer und Tomczyszyn (2015) zu validieren. Hierzu wurden zwei Erklärvideos zum Thema Kurvenfahrt erstellt. Eines der beiden Videos war gemäß des Qualitätsmaßstabs von guter, das andere von schwacher Qualität. Der Einfluss dieser qualitativ unterschiedlichen Erklärvideos auf den Lernzuwachs wurde mittels einer weitestgehend quantitativen Fragebogenerhebung im Prä-/Posttestdesign mit 176 Bremer Schülerinnen und Schülern untersucht. Die auf der Empirischen Erhebung basierende Analyse der hake-Faktoren ergab, dass durch das Erklärvideo guter Qualität ein höherer Lernzuwachs (mittlerer Lernzuwachs) erreicht wurde, als durch das Video schwacher Qualität (geringer Lernzuwachs). Dabei war der erhobene Lernzuwachs weitgehend altersunabhängig. Jedoch wurde festgestellt, dass die weiblichen Probanden im Prätest signifikant schlechtere Leistungen als die männlichen Probanden erbracht haben. Der Posttest ergab hingegen unabhängig vom Geschlecht und vorgeführtem Video gleiche Leistungen. Aus den Ergebnissen der Empirischen Erhebung folgt, dass zu Gunsten der Adressatengerechtheit das Vorwissen und Interesse der Probanden berücksichtigt werden sollten. Demzufolge konnte der Qualitätsmaßstab nur bedingt validiert werden, und eine Validierung mittels realer Erklärsituationen wird empfohlen.

Abstract

The aim of the present thesis was to validate the quality standard for „good explaining“ by Kulgemeyer and Tomczyszyn. For that purpose two illustrating videos explaining the issue of driving along curves were created. While one of those clips is of good quality, the quality of the second video is rather low. To measure the influence of those clips on the learning progress, a quantitative questionnaire survey by the means of pre/post design tests was conducted. The questionnaire was given to 176 students who are all attending schools in Bremen. The outcomes of the empirical research show that the higher quality video generates a higher learning progress (medium learning progress according to Hake) than the lower quality video (low learning progress according to Hake). The results also reveal that that the learning progress does not depend on the students’ age. However, the female participants achieved significantly worse results in the pretest than their male counterparts, whereas the posttest results do not indicate that neither gender nor the quality of the shown video had any significant influence on the test’s outcome. Based on the findings, it can be pointed out that the foreknowledge and the personal interests of the participants should be considered at all times. Therefore this scale for quality is only partly valid. For a complete validation a field test in real explaining situations is recommended.

Inhaltsverzeichnis

Zusammenfassung (Abstract) I
Abstract I
1 Einleitung 1
2 Theoretischer Rahmen 2
2.1 Fachliche Theorie zur Kurvenfahrt 2
2.1.1 Trägheit 3
2.1.2 Kreisbewegung 3
2.1.3 Reibung 5
2.2 Fachdidaktische Theorie 7
2.2.1 Erklären 7
2.2.1.1 Bedeutung des Erklärens im Physikunterricht 8
2.2.1.2 Erklärfähigkeiten 8
2.2.1.3 Qualitätsmaßstab für „gutes Erklären“ 10
2.2.1.4 Erklärvideos 12
2.2.2 Interesse 12
2.2.3 Vorwissen 13
3 Empirische Erhebung 15
3.1 Fragestellung und Hypothesen 15
3.2 Probanden 16
3.3 Methode der Datenerhebung 16
3.4 Methode der Datenanalyse 19
4 Ergebnisse 23
5 Diskussion 34
6 Fazit und Ausblick 41
7 Literaturverzeichnis 43
Anhang 47
Eidesstattliche Erklärung 52

1 Einleitung

„Erklären ist eine Standardsituation im Physikunterricht“ (vgl. Kulgemeyer 2013). Gut erklären zu können ist somit eine der grundlegenden Kompetenzen für alle Physiklehrerinnen und -lehrer. Umso erstaunlicher ist es, dass sich nur wenige Studien mit dem Thema des Erklärens beschäftigen (vgl. Kulgemeyer/Tomczyszyn 2015). Daher widmet sich das Projekt ProfiLe-P, unterteilt in drei Teilprojekte, unter anderem das Teilprojekt EWis, dem Professionswissen angehender Lehrkräfte. Das für die Erklärfähigkeit zuständige EWis beschäftigt sich mit einer standardisierten Erklärsituation, in welcher ein Erklärer einem geschulten Adressaten ein vorgegebenes physikalisches Phänomen erklärt (vgl. Kulgemeyer/Tomczyszyn 2015). Neben Fachwissen und Fachdidaktik ermöglicht es dieser dritte Teilbereich des Erklärens somit die geforderten Inhalte entsprechend zu vermitteln. Im Rahmen dieses Teilprojekts wurde ein Qualitätsmaßstab zur Bewertung von „gutem Erklären“ entwickelt. Dieser soll künftig ein wichtiges Hilfsmittel in der Ausbildung angehender Physiklehrerinnen und -lehrer darstellen.

Das Ziel dieser Arbeit besteht darin, den entwickelten Qualitätsmaßstab mit Hilfe von Erklärvideos zu validieren. Diese Untersuchungen sollen speziell für das Phänomen der Kurvenfahrt als exemplarischem Teilgebiet der Physik erfolgen. Aus diesem Grund wurden zwei Videos zum Thema Kurvenfahrt erstellt, die zwar inhaltlich gleichwertig, aber laut Maßstab von deutlich unterschiedlicher Qualität sind. Durch ein Fragebogenverfahren im Prä-/Posttestdesign kann der Einfluss der Videoqualität auf den jeweiligen Lernzuwachs untersucht werden.

Um dieses Ziel erreichen zu können, wurde die vorliegende Arbeit folgendermaßen strukturiert: Zunächst wird die fachliche Theorie zum Thema Kurvenfahrt erläutert. Anschließend wird die entsprechende Fachdidaktik beleuchtet, indem die Begriffe „Erklären“, „Interesse“ und „Vorwissen“ genauer analysiert werden. Im Anschluss daran werden Fragestellung und Hypothesen der Fragebogenerhebung, die Probanden und die Methode der Datenerhebung sowie -analyse dargestellt. Nach dem Ergebnissteil folgt die Diskussion sowie abschließend ein Fazit mit Ausblick.

2 Theoretischer Rahmen

2.1 Fachliche Theorie zur Kurvenfahrt

Das Ziel der vorliegenden Arbeit besteht darin, den Qualitätsmaßstab für Erklärvideos am Beispiel des physikalischen Themas „Kurvenfahrt“ zu validieren. Dafür sollen zunächst die theoretischen Grundlagen zur Kurvenfahrt dargestellt werden.
Bei diesem Szenario handelt es sich um ein alltägliches Phänomen beziehungsweise Problem, welches auftritt, sobald ein Fahrzeug schnell in eine Kurve mit „feuchtem“ Untergrund fährt. Exemplarisch ist eine solche Kurvenfahrt am Beispiel eines Motorrads in Abbildung 1 dargestellt. Fährt das Motorrad ausgehend vom „festen“ plötzlich auf „feuchten“ Untergrund, so verringert sich die Reibungskraft. Das liegt daran, dass der Reibungskoeffizient zwischen Reifen und Asphalt auf „feuchtem“ Untergrund geringer ist, als auf „festem“ Untergrund. Durch die Verringerung der Reibungskraft in der Kurve wird auch die Zentripetalkraft kleiner. Hierbei handelt es sich um die zum Mittelpunkt hin gerichtete Kraft. Aus diesem Grund fährt das Motorrad im geschilderten Szenario einen größeren Kurvenradius. Das Fahrzeug weist nicht mehr die gleiche Kraft in Richtung Kreismittelpunkt auf, die es auf „festem“ Untergrund hatte und „fliegt aus der Kurve“.

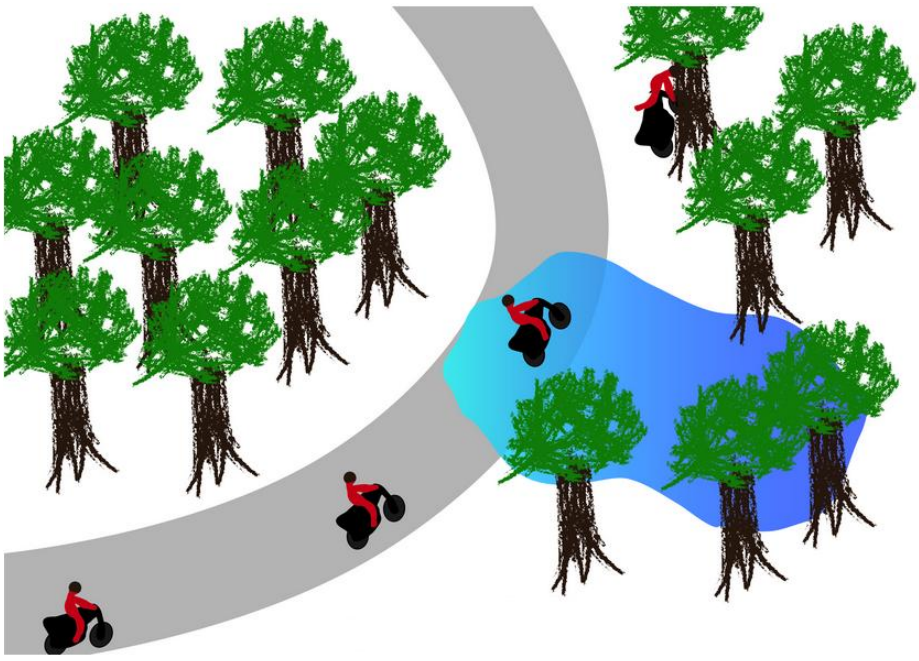

Abbildung 1: Szenario der Kurvenfahrt für ein Motorrad zu verschiedenen Zeitpunkten (aus Kulgemeyer/Schecker 2015).

Um das Gesamtkonzept des Phänomens der Kurvenfahrt erschließen zu können, müssen die zentralen Einflussfaktoren bekannt sein. Hierbei handelt es sich um die Trägheit, die Kreisbewegung und die Reibungskraft. Diese werden im Folgenden näher beleuchtet.

2.1.1 Trägheit

Bei der Trägheit handelt es sich um die Eigenschaft physikalischer Körper, in ihrem Bewegungszustand zu verharren, wenn keine Kraft auf sie wirkt (vgl. Demtröder 2015).
Demzufolge bleibt ein ruhender Körper ohne Krafteinwirkung in Ruhe, während ein bewegter Körper mit gleichbleibender Geschwindigkeit geradlinig weiterläuft (vgl. Hammer/Hammer 1995). Das bedeutet wiederum, wenn ein ruhender Körper in Bewegung gebracht werden soll, so muss eine Kraft auf ihn wirken. Um jedoch die Geschwindigkeit eines bewegten Körpers zu steigern, ist ebenfalls eine Kraft nötig, desgleichen um einen bewegten Körper abzubremsen. Auch zur bloßen Richtungsänderung der Geschwindigkeit muss eine Kraft aufgewendet werden.
Das beschriebene Phänomen der Trägheit wurde von Newton im Beharrungs- oder Trägheitsprinzip, auch erstes Newtonsches „Gesetz“ genannt, definiert als: *„Jeder Körper behält seine Geschwindigkeit nach Betrag und Richtung bei, wenn keine Kraft auf ihn wirkt“* (Heintze/Bock 2014). Genauer gesagt muss die resultierende Kraft auf den physikalischen Körper Null sein.

2.1.2 Kreisbewegung

„Die einfachste krummlinige Bewegung ist die eines Körpers, der mit konstanter Geschwindigkeit einen Punkt umkreist“ (Bolte/Meldau 1925). Sie wird auch als Zentral- oder Kreisbewegung bezeichnet.
Die in Richtung des Kreismittelpunktes gerichtete Kraft wird Zentripetalkraft F_Z genannt. Diese kann mit Hilfe von Formel (1) beschrieben werden:

$$\mathbf{F_Z} = \frac{\mathbf{m} \cdot \mathbf{v^2}}{\mathbf{r}} \tag{1}$$

Hierbei ist *m* die Masse des Körpers, der die Kreisbewegung vollführt, *v* seine Geschwindigkeit und *r* der Radius des Kreises.

Ist die Geschwindigkeit, mit welcher besagter Körper einen Punkt umkreist, nicht konstant, so handelt es sich nicht um eine Kreisbewegung, sondern um eine andere krummlinige Bewegung. Verantwortlich für die Veränderung der Geschwindigkeit sind die Tangentialkraft und die Zentripetalkraft. Letztere ist zuständig für die Richtungsänderung, während die Tangentialkraft die Änderung des Betrags der Geschwindigkeit, also des Tempos, bewirken kann (vgl. Alonso/Finn 2000).

Die Tangentialkraft F_T kann mit Hilfe von Formel (2) beschrieben werden und ist die in Richtung der Tangente an die Bahnkurve wirkende Kraft:

$$\mathbf{F_T} = \mathbf{m} \cdot \mathbf{a_T} \tag{2}$$

Hierbei entspricht *m* ebenfalls der Masse des Körpers und a_T ist die Beschleunigung in Tangentialrichtung.

Das bedeutet, dass die Tangentialkraft bei Kreisbewegungen Null ist. Außerdem wird dann jeweils die gleiche Strecke in der gleichen Zeit zurückgelegt (vgl. Demtröder 2015). Sollte jedoch die Zentripetalkraft Null sein, so würde es sich um eine geradlinige Bewegung handeln.

Die zugehörigen Vektoren sind in Abbildung 2 dargestellt: Die Tangential- und Zentripetalvektoren sind an den Punkten P, P_1 und P_2 des Kreisausschnittes mit dem Mittelpunkt *M* eingezeichnet.

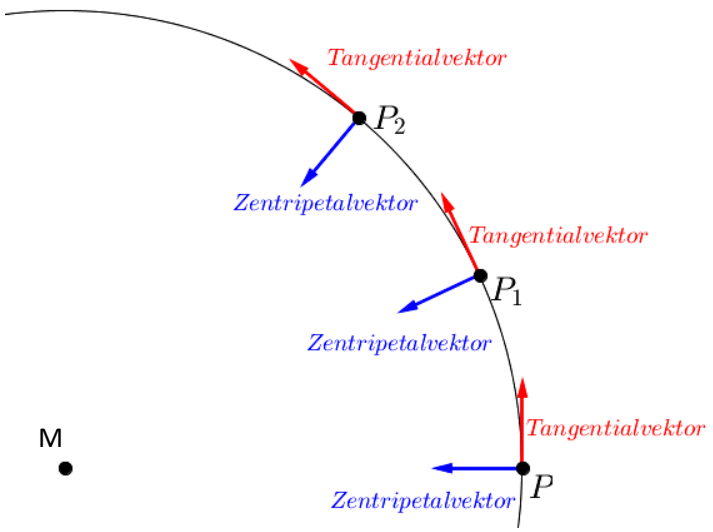

Abbildung 2: Tangential- und Zentripetalvektoren bei einer Kreisbewegung.

2.1.3 Reibung

Bei der Reibung zwischen zwei Festkörpern handelt es sich, anders als bei den vorherigen Phänomenen, nicht um wohldefinierte, mathematisch formulierbare Naturgesetze, sondern um eine empirische Gesetzmäßigkeit (vgl. Heintze/Bock 2014).
Die Reibungskraft tritt auf, sobald sich zwei Körper berühren und relativ zueinander bewegen, zum Beispiel wenn ein Holzklotz über den Tisch gleitet oder ein Rad sich über den Asphalt dreht. Verstärkt wird diese Kraft durch Unebenheiten der Oberflächen und Deformationen bei der Berührung. Der wichtigste Aspekt, warum die Reibungskraft von besonderer Wichtigkeit ist, ist die Tatsache, dass sich ohne Reibung Bewegungen nicht verändern könnten (vgl. Demtröder 2015). Unterschieden wird bei der Reibungskraft zwischen der Haft-, Gleit- und Rollreibung, auf die im Folgenden genauer eingegangen wird.
Die Haftreibung tritt auf, wenn sich der Körper in Ruhe befindet und in Bewegung gesetzt werden soll. Damit sich der Körper bewegt, muss die auf den Körper wirkende Kraft einen Grenzwert F_H überschreiten. Ist diese Kraft F geringer als die Haftreibungskraft F_H ($F < F_H$), so bewegt sich der Körper nicht. Ist diese Kraft jedoch mindestens genauso groß ($F \geq F_H$), so bewegt sich der Körper mit der Richtung der wirkenden Kraft. Diese Haftreibung ist proportional zu der auf den Körper wirkenden Normalkraft F_N. Dieser Zusammenhang ist in Formel (3) dargestellt:

$$\mathbf{F_H} = \boldsymbol{\mu_H} \cdot \mathbf{F_N} \tag{3}$$

Wobei μ_H der Haftreibungskoeffizient ist, welcher vom Material der sich berührenden Körper und von der Beschaffenheit ihrer Oberflächen abhängt (vgl. Demtröder 2015). Dieser Zusammenhang ist in Abbildung 3 dargestellt:

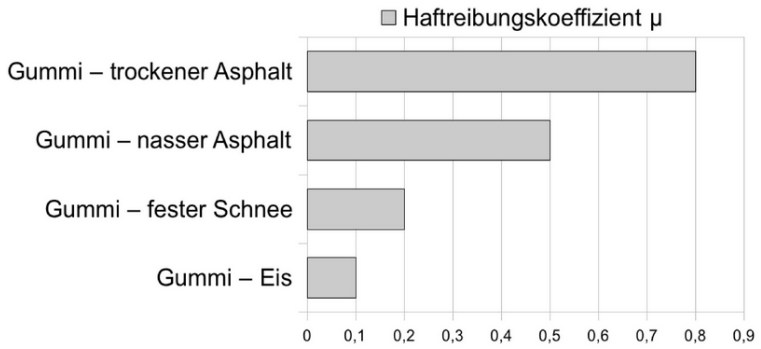

Abbildung 3: Haftreibungskoeffizient für verschiedene Untergrunde (aus Kulgemeyer/Schecker 2015).

Je größer der Haftreibungskoeffizient ist, desto stärker ist die Kraft, die zwischen den beiden Körper wirkt. Wurde ein Körper aber bereits in Bewegung gebracht, also hat man eine Kraft $F \geq F_H$ aufgebracht, so wird besagter Körper beschleunigt.
Ist der Körper einmal in Bewegung, wirkt nicht mehr der Haftreibungs-, sondern der Gleitreibungskoeffizient μ_G. Das bedeutet, dass zum Erhalt einer gleichmäßigen Bewegung eine Gleitreibungskraft $F_G \leq F_H$ ausreicht, welche wiederum proportional zur Normalkraft F_N ist (vgl. Demtröder 2015). Dies kann Formel (4) entnommen werden.

$$\mathbf{F_G} = \boldsymbol{\mu_G} \cdot \mathbf{F_N} \tag{4}$$

Analog zum Haftreibungskoeffizienten μ_H ist auch dieser Gleitreibungskoeffizient μ_G abhängig vom Material der sich berührenden Körper und der Beschaffenheit ihrer Oberflächen.
Abschließend soll das Phänomen der Rollreibung beschrieben werden. Diese tritt nur dann auf, wenn ein runder Körper auf einer ebenen Unterlage rollt.
Zurückzuführen ist diese Reibungskraft auf die Anziehungskräfte zwischen den Atomen entlang der Berührungslinie zwischen Unterlage und rollendem Körper. Um den rollenden Körper abzubremsen, muss das entgegengesetzte Drehmoment D_R das Drehmoment des Körpers kompensieren (vgl. Demtröder 2015). Analog zu den anderen Reibungskräften ist auch die Rollreibungskraft des Drehmomentes D_R proportional zur Normalkraft F_N (vgl. Formel (5)):

$$\mathbf{D_R} = \boldsymbol{\mu_R} \cdot \mathbf{F_N} \tag{5}$$

Allerdings ist der Rollreibungskoeffizient μ_R im Gegensatz zu den anderen beiden Reibungskoeffizienten nicht dimensionslos, da er die Längendimension aufweist (vgl. Demtröder 2015).

Nachdem die zentralen Einflussfaktoren Trägheit, Kreisbewegung und Reibung beleuchtet wurden, soll nun ein Rückbezug zum dargestellten Szenario einer Kurvenfahrt erfolgen.
Beim beschriebenen Körper handelt es sich in diesem Beispiel um ein Motorrad. Fährt dieses, wie bereits beschrieben, schnell um eine Kurve, wirkt die dafür notwendige Zentripetalkraft durch die Reibung zwischen Reifen und Asphalt. Solange das

Motorrad in der Kurve fährt, wirkt die bereits erläuterte Haftreibung in Richtung eines gedachten Kreismittelpunktes und es gilt: Die Haftreibungskraft ist gleich der Zentripetalkraft ($F_H = F_Z$). Somit ergibt sich aus den Formeln (1) und (3) die Formel (6) für die Kurvenfahrt:

$$\mathbf{\mu_H \cdot F_N = \frac{m \cdot v^2}{r}} \tag{6}$$

Der Reibungskoeffizient μ_H ist abhängig von der Beschaffenheit der Oberflächen, also Reifen und Straße. Fährt das Fahrzeug nun auf eine feuchte Stelle innerhalb der Kurve, so verringert sich der Reibungskoeffizient zwischen Reifen und Straße: Der Reibungskoeffizient auf der feuchten Fahrbahn ist geringer als auf der trockenen ($\mu_{H(feucht)} < \mu_{H(trocken)}$). Somit ist für den „feuchten“ Untergrund auch die Zentripetalkraft geringer. Nach Formel (6) kann sich lediglich der Krümmungsradius *r* der Bahn ändern, da die Masse *m* des Fahrzeuges konstant bleibt und die Geschwindigkeit *v* des Fahrzeugs sich ebenfalls nicht ändert. Gemäß Formel muss ein größerer Kurvenradius gefahren werden, weil das Fahrzeug nicht mehr genügend „Halt“ auf der feuchten Straße hat.
Da dies bei einer asphaltierten Straße nur bedingt möglich ist, „fliegt“ das Motorrad bedingt durch die Trägheit, aus der Kurve.

2.2 Fachdidaktische Theorie

Nachdem die fachlichen Grundlagen dargestellt wurden, sollen im Folgenden die für die empirische Erhebung relevanten Theorien erörtert werden. Im Zuge dessen werden die Begrifflichkeiten „Erklären“, „Interesse“ und „Vorwissen“ – stets bezogen auf die entsprechenden Grundlagen der Physikdidaktik – beleuchtet.

2.2.1 Erklären

Beim Erklären geht es darum, ein bestimmtes Wort, ein Phänomen, einen Sachinhalt oder einen Textsinn verständlich darzulegen (vgl. Spreckels 2009). Dabei erzeugt

„gutes Erklären“ allerdings nicht automatisch Verständnis. Jedoch bildet es die Grundlage dafür einen Sachverhalt verstehbar zu machen (vgl. Gage 1968).

2.2.1.1 Bedeutung des Erklärens im Physikunterricht

„Explanation is at the core of science education“ (Baid 1988). Wie Baid festgestellt hat, stellt das Erklären die Grundlage für den naturwissenschaftlichen Unterricht dar. Dies gilt somit auch und im Besonderen für den Physikunterricht. Ergänzend formulierte Kulgemeyer: „*Das Erklären ist eine Standardsituation im Physikunterricht*“ (Kulgemeyer 2013) beziehungsweise „*gut erklären zu können wird als wichtigste Eigenschaft von Lehrkräften oft genannt*“ (vgl. Kulgemeyer/Tomczyszyn 2015).
Das bedeutet, dass die Schülerinnen und Schüler auf die Fähigkeit ihrer Lehrkraft, eine gute Erklärung geben zu können, angewiesen sind. Kann diese den Sachverhalt gut darstellen, so haben die Schülerinnen und Schüler eine höhere Chance das Gesagte zu verstehen (vgl. Leisen 2013). Somit muss dieser Kompetenz bereits in der Lehrerausbildung eine besondere Bedeutung zugeschrieben werden. Dennoch beschäftigt sich die aktuelle Forschung höchstens am Rand mit der Thematik des guten Erklärens. Die wenigen Untersuchungen auf diesem Gebiet fokussierten zudem eher die Rolle des Erklärens im Unterricht, statt der Qualität verschiedener Erklärungen (vgl. Kulgemeyer/Tomczyszyn 2015).

2.2.1.2 Erklärfähigkeiten

Nachdem der Begriff des „Erklärens“ einleitend definiert wurde, soll im Folgenden auf die sogenannte „Erklärfähigkeit“ von Personen näher eingegangen werden.
Nach Leisen muss das Anspruchsniveau einer physikalischen Erklärung je nach Adressat angepasst werden. Hierbei sind folgende Fälle zu unterscheiden. Zum einen das Erklären zwischen Mitschülerinnen und -schülern, bei denen der gleiche Verstehenshorizont vorliegt. Zum anderen das Erklären gegenüber Laien, wobei sich eine möglichst anschauliche Art der Argumentation empfiehlt. Letztlich das Erklären gegenüber Experten. In diesem Fall soll eine fachlich exakte, knappe Darstellung des Inhaltes erfolgen (vgl. Leisen 2008).

Diese Anpassung des Erklärens an den Adressaten ist die Voraussetzung für eine „gute Erklärung“. Dabei erzeugt diese zwar nicht automatisch Verständnis für den Sachverhalt, stellt aber wiederum die Grundlage für dessen Verständnis dar (vgl. Gage 1968).

Daraus folgt, dass die Erklärfähigkeit die Kompetenz ist, Fachinformationen fach- und adressatengerecht aufzubereiten, sodass der Adressat daraus selbst die gewünschte Bedeutung konstruieren kann (vgl. Kulgemeyer 2013).

Dieser Ansatz basiert auf einem konstruktivistischen Modell für Kommunikationskompetenz, welches zwischen der Sach- und Adressatengerechtheit unterscheidet (vgl. Kulgemeyer/Schecker 2009). Die Sachgerechtheit bedeutet, dass der Erklärer über das nötige Physikwissen verfügt und in der Lage ist, eine entsprechende fachliche Korrektheit der Erklärung zu leisten. Die Adressatengerechtheit bedeutet hingegen die Berücksichtigung von Interesse und Vorwissen des Gegenübers.

Bei besagtem konstruktivistischem Modell wird beschrieben, dass die physikalische Kommunikationskompetenz abhängig von den vier Komponenten „Kontext“, „Sprachebene“, „Sachinhalt“ und „Darstellungsform“ des gewählten Sachinhaltes ist (vgl Kulgemeyer 2010). Dieses Modell wurde von Kulgemeyer und Tomczyszyn weiterentwickelt und ist in Abbildung 4 dargestellt.

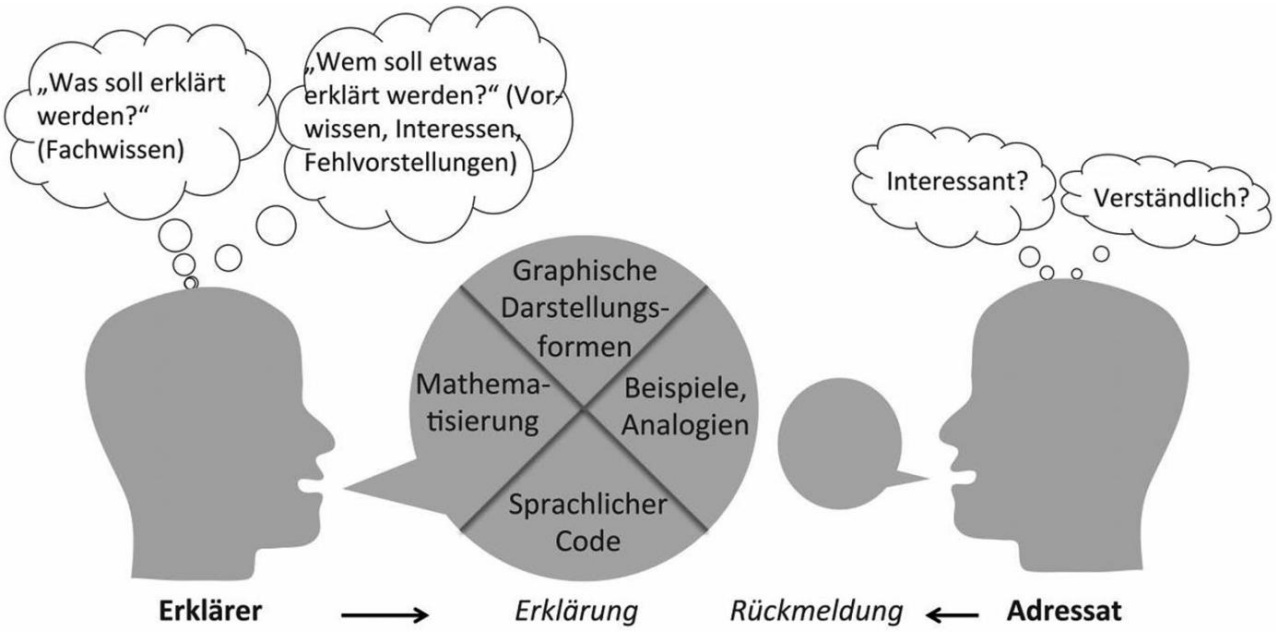

Abbildung 4: Der Prozess des Erklärens mit Vermittlungscharakter von Physik (aus Kulgemeyer und Tomcyszyn 2015).

In diesem erweiterten Modell werden Erklärungen als sprachlicher Prozess aufgefasst, bei welchem sich die Adressaten- und Sachgerechtheit aus vier „Variablen“ zusammensetzt. Hierzu gehören graphische Darstellungsformen, Beispiele und Analogien, der sprachliche Code und die Mathematisierung. Um dem Adressaten die Erklärung verständlicher zu machen, hat der Erklärer somit die Möglichkeit die sachgerechte („Was soll erklärt werden?“) und die adressatengemäße Perspektive („Wem soll etwas erklärt werden?“) in diesen vier Variablen entsprechend anzupassen (vgl. Kulgemeyer/Tomczyszyn 2015).
Abgesehen davon wurden bereits 2013 sowohl von Kulgemeyer als auch von Leisen die zehn Kriterien für „gutes Erklären“ zusammengetragen (vgl. Kulgemeyer 2013; Leisen 2013). Diese Kriterien sind in Tabelle 1 dargestellt.

Tabelle 1: Die zehn Kriterien "guten Erklärens".

1	Vorbereitung auf eine Erklärung
2	Klarheit der Ziele zu Beginn der Erklärung
3	Orientierung über den Ablauf zu Beginn der Erklärung
4	Strukturiertheit
5	Wiederholung von Erklärungen nach Elaboration
6	Kontakt mit dem Adressaten aufnehmen
7	Regelmäßige Verständnissicherung
8	Anpassung der sprachlichen Komplexität
9	Einsatz visueller Medien
10	Verwendung von Beispielen

2.2.1.3 Qualitätsmaßstab für „gutes Erklären“

Basierend auf dem Kommunikationsmodell von Kulgemeyer und Tomczyszyn (vgl. Abbildung 4) sowie den zehn Kriterien für „gutes Erklären“ (vgl. Tabelle 1) wurde im Rahmen des Projekts Profile-P ein Maß zur Messung der adressatengemäßen Erklärfähigkeit entworfen (vgl. Kulgemeyer 2013).
Diese Grundlagen überschneiden sich bereits inhaltlich: Die letzten drei Kriterien „guten Erklärens“, die Anpassung der sprachlichen Komplexität, der Einsatz visueller Medien und die Verwendung von Beispielen entsprechen drei der Variablen des be-

schriebenen Modells, nämlich dem sprachlichen Code, der graphischen Darstellung sowie den Beispielen und Analogien (vgl. Kulgemeyer 2013).
Besagtes Maß zur Messung von Erklärfähigkeiten ist in Tabelle 2 dargestellt:

Tabelle 2: Qualitätsmaßstab für „gutes Erklären“ (aus Kulgemeyer/Tomczyszyn 2015).

	Kategorie	**Beschreibung**
1	Rückblick	Der Proband sagt explizit, dass er etwas vorher gezeigt oder erklärt hat.
2	Zusammenfassung / Wiederholung	Der Proband gibt eine Zusammenfassung oder macht eine Wiederholung eines Sachverhalts, der schon erklärt wurde.
3	Ermunterung	Der Proband macht eine positive Äußerung bezüglich des Adressaten
4	Umschreibung von Fachbegriffen	Der Proband vermeidet Fachbegriffe, indem er diese durch Alltagsbegriffe ersetzt.
5	Unangemessene Beispiele	Die Situation wird falsch beschrieben oder die Situation passt nicht zum physikalischen Inhalt
6	Zahlenbeispiele	Eine physikalische Größe wird durch eine Zahl ersetzt.
7	Verständnisversicherung	Der Proband erfragt den Erfolg seiner Erklärung, indem er eine Wissensfrage oder Aufgabe stellt.
8	Handlungsaufforderung	Der Proband fordert den Adressaten auf etwas zu tun.
9	Gegenständliche Hilfsmittel	Gegenstände werden zum Erklären genutzt.
10	Experimentieren	Gegenständliches Hilfsmittel verlässt die Hand oder der Gegenstand wird variiert.
11	Handlung / Gegenstand mit Szenario verknüpfen	Der Proband benennt den Gegenstand mit einem Objekt aus dem Szenario.
12	Verknüpfen von nonverbalen Elementen	Der Proband sagt, was an zwei nonverbalen Elementen gleich ist.

Der in Tabelle 2 dargestellte von Kulgemeyer und Tomczysyn entwickelte Qualitätsmaßstab für „gutes Erklären“ wurde bereits auf seine Validität und Objektivität überprüft (vgl. Bartels 2015) und fokussiert die Interaktion zwischen Erklärer (Proband) und Adressat. Somit wird auf die Betrachtung der physikalischen Korrektheit der Erklärung (Sachgerechtheit) verzichtet (vgl. Kulgemeyer/Tomczyszyn 2015).
Desweiteren umfasst der in Tabelle 2 dargestellte Qualitätsmaßstab zur Messung von Erklärfähigkeiten zwölf verschiedene Kategorien. Bei der Einschätzung, ob es sich bei einer bestimmten Erklärung um eine „gute Erklärung“ handelt, wird ein Punk-

tesystem verwendet: Wird das jeweilige Kriterium vom Erklärer (Proband) erfüllt, erhält die Erklärung jeweils einen Punkt. Dies gilt nicht für die Kategorie 5 („Unangemessene Beispiele“) – hier verhält es sich andersherum. Somit kann eine Erklärung zwischen null und zwölf Punkten erreichen. Je höher die Punktzahl, desto besser ist die jeweilige Erklärung beziehungsweise die Erklärfähigkeit des Probanden.

2.2.1.4 Erklärvideos

Um eine Erklärung mit Hilfe des vorgestellten Qualitätsmaßstabs korrekt bewerten zu können, eignet sich der Einsatz von sogenannten Erklärvideos. Dabei handelt es sich in diesem Kontext um ein Instrument, das die Aufzeichnung und somit die wiederholte Wiedergabe (zur Bewertung) einer Erklärung ermöglicht.
Zudem sind Erklärvideos von besonderer didaktischer Bedeutung, weil sie neben Vorteilen wie gesteigerter Anschaulichkeit und Aktualität (vgl. Bickel-Sandkötter 2003), die einfache und verständliche Vermittlung komplexer Themen und Inhalte ermöglichen. Verschiedene Vorgänge und Abläufe können mit Hilfe von Erklärvideos wirklichkeitsnah veranschaulicht werden. Zum einen ist es möglich, dass mittels Zeitlupe und Zeitraffer entsprechende Abläufe beschleunigt oder verlangsamt dargestellt werden (z.B. der freie Fall). Zum anderen ermöglichen Erklärvideos Zoomeffekte, sodass zum Beispiel modellhafte Elektronenbewegungen betrachtet werden können. Somit führt eine didaktische Reduktion zur Fokussierung auf wesentliche Elemente (vgl. Killermann/Hiering/Starosta 2005).
Der ausschlaggebende Punkt für die Nutzung von Erklärvideos zur Validierung des vorgestellten Qualitätsmaßstabs ist die Tatsache, dass die Erklärung beim wiederholten Zeigen nicht variiert.

2.2.2 Interesse

Zwar geht es in der vorliegenden Arbeit überwiegend um „gutes Erklären“, dennoch spielt auch das Interesse von Schülerinnen und Schüler, wie bereits im Kapitel 2.2.1.2 beschrieben und in Abbildung 4 dargestellt, eine wichtige Rolle. Aus diesem Grund wird der Begriff des Interesses im Folgenden kurz definiert.

Interesse ist eine spezifische Relation zwischen einer Person und einem Gegenstand. Dieses Interesse entwickelt sich erst in der Auseinandersetzung mit dem Gegenstand selbst (vgl. Schiefele 2008). Laut Schiefele muss zwischen zwei Arten von Interesse unterschieden werden; dem individuellen Interesse und dem situationalen Interesse. Letzteres kennzeichnet einen durch äußere Umstände hervorgerufenen, handlungsbegleitenden Zustand des Interessiertseins, der sich beispielsweise durch erhöhte Aufmerksamkeit auszeichnet. Dementgegen handelt es sich beim individuellen Interesse um ein dauerhaftes Merkmal einer Person, welches als eine Wertschätzung eines spezifischen Gegenstandes verstanden werden kann. Dies äußert sich in gefühls- und wertbezogenen Valenzüberzeugungen (vgl. Schiefele 2008). Unter gefühlsbezogenen Valenzüberzeugungen versteht man, dass ein Gegenstand für eine Person mit positiven Gefühlen verbunden ist. Zweiteres besagt, dass der Gegenstand mit einer persönlichen Bedeutsamkeit verknüpft ist (vgl. Krapp/Schiefele/ Wild/Winteler 1993).

Das wohl wichtigste Merkmal von Interesse ist jedoch sein intrinsischer Charakter, beziehungsweise die Selbstintentionalität (vgl. Krapp 1992). Somit kann nur von individuellem Interesse gesprochen werden, wenn die Auseinandersetzung mit einem (Lern-)Gegenstand primär aus „sachimmanenten" Gründen erfolgt. Das bedeutet, bezogen auf die beschriebene Konzeption des individuellen Interesses als Kombination aus gefühls- und wertbezogenen Valenzen, dass sich die Gefühls- und Wertzuschreibungen direkt auf einen Gegenstand oder eine Handlung beziehen müssen und nicht nur aufgrund extrinsischer Valenzen auftreten. Diese extrinsischen Valenzen liegen vor, wenn der Gegenstand mit anderen Sachverhalten in Verbindung steht, denen der eigentliche Wert und die positiven Gefühle zukommen (vgl. Krapp/Schiefele/Wild/Winteler 1993).

2.2.3 Vorwissen

Neben dem bereits beschriebenen Interesse spielt auch das Vorwissen eine große Rolle hinsichtlich des Erklärens und dessen Gelingen (vgl. Abbildung 4) und soll im Folgenden kurz thematisiert werden.

Unter Vorwissen beziehungsweise Vorkenntnissen einer Person versteht man das Wissen, über das bereits vor Beginn des Lernprozesses verfügt wird. Dieses wird bei

Beginn des Lernprozesses aufgabenspezifisch abgerufen (vgl. Wirth 2004). Aufgrund ihrer lebensweltlichen Erfahrung und Bewährung haben die Vorstellungen der Schülerinnen und Schüler einen Eigenwert. Sie sind persönliche Konstrukte mit emotionalen und sozialen Komponenten, die Ausgangspunkte und Hilfsmittel des Lernens darstellen (vgl. Kattmann 2007).

Es gilt auf der Basis von Vorwissen (und Interesse) durch entsprechend angepasstes Erklären (vgl. Kapitel 2.2.1) für die Schülerinnen und Schüler einen möglichst hohen Lernzuwachs zu erzielen.

3 Empirische Erhebung

3.1 Fragestellung und Hypothesen

Ziel der Arbeit ist die Validierung des in Kapitel 2.2.1.3 vorgestellten Qualitätsmaßstabs für „gutes Erklären“, beziehungsweise für Erklärvideos am Beispiel der Kurvenfahrt.
Es ergibt sich dementsprechend folgende Forschungsfrage: Erzielt ein Erklärvideo zur Kurvenfahrt, das gemäß Maßstab von guter Qualität ist, einen höheren Lernzuwachs als ein entsprechendes Erklärvideo von minderer Qualität?
Die dazugehörige Hypothese lautet:

Hypothese: *Wenn ein Erklärvideo zur Kurvenfahrt gemäß des Qualitätsmaßstabs von guter Qualität ist, dann ist der Lernzuwachs der Probanden höher, als wenn das Erklärvideo von schwacher Qualität ist.*

Diese Hypothese kann durch die beiden folgenden untergeordneten Hypothesen noch spezifiziert werden:

Unterhypothese 1: *Wenn die Probanden ein Erklärvideo zur Kurvenfahrt mit guter Qualität sehen, dann ist ihr Lernzuwachs im Durchschnitt sehr hoch.*

Unterhypothese 2: *Wenn die Probanden ein Erklävideo zur Kurvenfahrt mit schwacher Qualität sehen, dann ist ihr Lernzuwachs im Durchschnitt nicht vorhanden.*

Neben dieser hypothesengeleiteten Untersuchung sollen auch im explorativen Verfahren bestimmte Sachverhalte festgehalten und untersucht werden. Dazu zählt, ob signifikante Unterschiede zwischen den beiden Geschlechtern und in Abhängigkeit vom Alter der Probanden auftreten.

3.2 Probanden

Im Rahmen dieser Arbeit erfolgte eine Befragung von Schülerinnen und Schüler, die die Oberschule Rockwinkel besuchen. Diese Schule befindet sich in einem der „gehobeneren“ Stadtteile Bremens (Oberneuland). Es handelt sich sowohl um Schülerinnen und Schüler der neunten Klasse, die kurz vor dem Erwerb des mittleren Schulabschlusses stehen, als auch um Schülerinnen und Schüler der Einführungsphase der gymnasialen Oberstufe, welche Physik im Grundkurs belegen.
Diese Zielgruppe ergab sich, da das Beispielthema der Kurvenfahrt zum Gebiet der Mechanik gehört, welches alle Probanden bereits im Physikunterricht behandelt haben sollten. Entweder beim Rahmenthema „Alles in Bewegung“ der Oberschule (vgl. Die Senatorin für Bildung und Wissenschaft 2010) oder bei „Kräfte und Bewegungen“ am Gymnasium (vgl. Die Senatorin für Bildung und Wissenschaft 2006).
Somit sind die Schülerinnen und Schüler zwischen 13 und 17 Jahren alt, das ermittelte Durchschnittsalter beträgt 15,02 Jahre (SD = 0,88).
Insgesamt wurden 176 Schülerinnen und Schüler befragt, wobei 80 Probanden männlich und 96 weiblich waren und 90 Probanden das Erklärvideo guter Qualität und 86 das Erklärvideo schwacher Qualität zu sehen bekamen.

3.3 Methode der Datenerhebung

In den fachdidaktischen Theorien wurde bereits beschrieben, dass die Erklärfähigkeiten des Erklärers das Verständnis und somit den Lernzuwachs des Adressaten beeinflussen (vgl. Kapitel 2.2.1.2).
Um die Abhängigkeit des Verständnisses der Schülerinnen und Schüler in Form eines messbaren Lernzuwachses von der Qualität der Erklärung untersuchen zu können, wurden zwei Erklärvideos konzipiert. Diese sind gemäß des Qualitätsmaßstabs für „gutes Erklären“ qualitativ unterschiedlich, behandeln inhaltlich aber beide das Phänomen der Kurvenfahrt im gleichen zeitlichen Umfang (ca. 6,5 Minuten). Eines der beiden Erklärvideos (Video 1) beinhaltet viele der in Tabelle 2 dargestellten Kategorien und erhält auf der Bewertungsskala 8 von 12 Punkten. Die Kategorien 3, 5, 7 und 8 (vgl. Tabelle 2) konnten aufgrund dieses gewählten Mediums nicht erfüllt

werden, sodass Video 1 die Maximalpunktzahl des Qualitätsmaßstabs für Erklärvideos aufweist und somit von „guter Qualität“ ist. Das andere der beiden Erklärvideos (Video 2) weist null Punkte bezüglich des Qualitätsmaßstabs auf und ist somit von „schwacher Qualität“.

Zur Überprüfung des eventuellen Lernzuwachses der Schülerinnen und Schüler bezüglich der Kurvenfahrt durch Video 1 beziehungsweise Video 2 wurde ein Fragebogen entwickelt, der im Prä- / Posttest-Design ausgerichtet ist. Dabei wurde durch den Prätest zunächst das Vorwissen und nach dem Input durch Erklärvideo 1 oder 2 mit Hilfe des Posttests der Lernzuwachs erhoben (vgl. Bortz/Döring 2006).

Allgemein stellt ein Fragebogen ein gutes methodisches Verfahren zur Erfassung der Variable Fachwissen zu einem bestimmten Sachverhalt, hier der Kurvenfahrt, dar. Für die Bearbeitung des jeweiligen Fragebogens stand den Probanden jeweils eine Bearbeitungszeit von etwa 15 Minuten zur Verfügung.

Der erstellte Fragebogen wurde nach bestimmten Prinzipien konzipiert. Um einer möglichen Verfälschung des Antwortverhaltens entgegenzuwirken, wird den Probanden im Eingangstext zunächst Anonymität zugesichert (vgl. Bortz/Döring 2006). Desweiteren befasst sich dieser Text mit einer kurzen Vorstellung der Person und die Ziele der Untersuchung werden aufgeklärt (vgl. Neuhaus/Braun 2007). Anschließend werden personenbezogene Daten zu Geschlecht und Alter der Probanden erhoben und ein Code gemäß Riese/Reinhold generiert. Dieser Code dient dazu, jedem Prätest einen Posttest eindeutig zuweisen zu können (vgl. Riese/Reinhold 2014). Im Hauptteil wurde mit Hilfe von sogenannten Items das Fachwissen überprüft, da dieses als einziger der vier Kompetenzbereiche im Fach Physik (vgl. Die Senatorin für Bildung und Wissenschaft 2010) eindeutig und damit auch objektiv messbar ist.

Diese Items sind Aufgabenstellungen mit zugehörigen Lösungsmöglichkeiten. Im vorliegenden Fragebogen wurden acht Items auf der Grundlage der von Schecker und Gerdes übersetzten Version des Force Concept Inventory (FCI) von Hestenes, Wells und Swackhammer zusammengestellt (vgl. Schecker/Gerdes 1999).

Dabei ist bei den ersten sechs Aufgaben jeweils eine der Antwortmöglichkeiten korrekt, während die Schülerinnen und Schüler bei Aufgabe sieben zwei Antworten ankreuzen mussten, um die Aufgabe richtig zu beantworten. Lediglich Aufgabe acht unterscheidet sich im Aufgabenformat von den vorherigen. Dabei müssen die Schülerinnen und Schüler in dieser Aufgabe ihre Antwort in eigene Worte fassen. Wäh-

rend beim Ankreuzen der Zufall eine Rolle spielen kann, haben Freitextantworten den Vorteil, konkrete Zusammenhänge und Vorstellungen der Schülerinnen und Schüler erfassen zu können.

Dennoch wurde der Fragebogen insgesamt so konzipiert, dass er die drei Gütekriterien Objektivität, Reliabilität und Validität (vgl. Schmiemann/Lücken 2014) so gut wie möglich erfüllt.

Objektivität bedeutet, dass eine wissenschaftliche Aussage unabhängig von einer subjektiven Einschätzung und Bewertung ist (vgl. Ramb/Wübbenhorst 2013). Das bedeutet, dass unterschiedliche Forscher mit derselben Messung dasselbe Ergebnis erhalten (vgl. Bortz/Döring 2006). Um dies zu gewährleisten, bestand der Fragebogen vorwiegend aus geschlossenen Aufgaben (vgl. Neuhaus/Braun 2007). Darunter versteht man Aufgaben mit vorgegebenen Antwortmöglichkeiten. Dies ermöglicht eine verhältnismäßig eindeutige Auswertung unabhängig davon, wer diese vornimmt (vgl. Riese/Reinhold 2014).

Reliabilität beschreibt die Wiederholbarkeit einer Messung. Ein Messergebnis muss bei einer Wiederholung unter gleichen Bedingungen mit identischen Messverfahren mit dem vorherigen übereinstimmen (vgl. Wübbenhorst/Maier 2013). Die Reliabilität gibt also an, inwieweit der wahre Wert mit dem gemessenen Wert übereinstimmt (vgl. Danner 2015).

Die Validität ist die wohl fundamentalste und zugleich komplexeste der drei Güterkriterien (vgl. Jenßen/Dunekacke/Blömeke 2015). Sie beschreibt, inwieweit die Messmethode tatsächlich das zu messende Konstrukt misst (vgl. Wübbenhorst 2013). Sie wird demnach als Angemessenheit, Aussagekraft und Nützlichkeit der spezifischen Rückschlüsse aus den Testergebnissen definiert (vgl. Jenßen/Dunekacke/Blömeke 2015). Somit wird die Validität als die Eigenschaft von Messwerten verstanden, die angibt, inwieweit die Interpretation von Testergebnissen durch Empirie und Theorie gestützt wird (vgl. Braun 2015). Basierend auf der Tatsache, dass die Antwortmöglichkeiten des Fragebogens im Anhang entweder richtig oder falsch sein können, ist dieser durch die Theorie gestützt und somit valide.

3.4 Methode der Datenanalyse

Die bei der Fragebogenerhebung ermittelten Daten wurden mit dem Computerprogramm PSPP ausgewertet. Dazu wurden den Antwortitems Punkte zugeordnet und in PSPP eingegeben. Wurde die korrekte Antwortmöglichkeit angekreuzt, so erhielt der Proband beim jeweiligen Item einen Punkt. Wurde eine der nicht korrekten Antwortmöglichkeiten ausgewählt, so wurde das Item mit null Punkten bewertet. Die letzten beiden Items stellen hierbei Ausnahmen dar. Da bei Aufgabe sieben zwei Antwortmöglichkeiten korrekt sind, wird hier pro richtiger Antwort ein halber Punkt gegeben. Somit kann auch bei diesem Item insgesamt maximal ein Punkt erreicht werden. Aufgrund der zentralen Stellung der Aufgabe acht zur Ursache des „Rausfliegens“ aus der Kurve, wurde diese stärker in die Wertung eingebracht. Sollten alle drei Phänomene, Trägheit, Kreisbewegung und Reibung, benannt und in den richtigen Zusammenhang gebracht worden sein (vgl. Braun 2015), so konnten hier drei Punkte erreicht werden. Wurde das Phänomen der Reibung durch die Formulierung „*Wenn es nass ist*“ (Proband CL10RE72) oder „*wenn wenig Reibung zwischen den Reifen und der Straße besteht*“ (Proband SU10RE78) lediglich erwähnt, so gab es dafür einen halben Punkt. Sollte das Phänomen allerdings mit einer entsprechenden Begründung angesprochen werden „*Ein Motorradfahrer fliegt aus der Kurve, wenn die Reibung zwischen Reifen und Asphalt kleiner ist, weil er plötzlich auf nasser Straße ist*“ (Proband RE25ST51), so wurde hier ein Punkt vergeben. Einen weiteren halben Punkt konnte man erreichen, wenn man die Trägheit nur benannt hatte. Wurde zusätzlich der Zusammenhang erläutert, zum Beispiel das „*Motorrad fährt normalerweise geradeaus, also muss Kraft wirken*“ (Proband MA28RE90), gab es einen ganzen Punkt für diesen Antwortteil. Analog wurde mit der Ursache der Kreisbewegung verfahren. Erwähnte man nur „*niedrige Zentripetalkraft*“ (Proband SI04RE82) ohne eine zugehörige Erklärung, so gab es dafür einen halben Punkt. Einen ganzen Punkt erhielt der Proband beispielsweise für: „*Wenn die Geschwindigkeit so groß ist, dass die Kraft nach innen nicht mehr reicht*“ (Proband HA18RE47) oder „*kleine Kurve und niedrige Zentripetalkraft*“ (Proband AK07RE78).

Insgesamt konnten maximal 10 Punkte bei der richtigen Beantwortung aller Items des Fragebogens erreicht werden.

Zu Gunsten einer einfacheren Auswertung wurden die halben Punkte auf ganze Punkte normiert. Somit sind in den ersten sechs Aufgaben entweder null oder zwei Punkte erreichbar. In der siebten null, ein oder zwei Punkte und in der letzten Aufgabe waren zwischen null und sechs Punkten möglich. Auf Grundlage dieser Normierung beträgt die Maximalpunktzahl 20.

Zur weiteren Auswertung wurden die Mittelwerte der einzelnen Items ermittelt. Um zu bestimmen, ob sich bestimmte Werte signifikant voneinander unterscheiden, wurde eine Reihe von T-Tests durchgeführt. Der T-Test ist eine auf mathematischen Grundlagen basierende Entscheidungsregel, mit deren Hilfe ein Unterschied zwischen den empirisch gefunden Mittelwerten zweier Gruppen näher analysiert werden kann (vgl. Rasch/Hoffmann/Friese/Naumann 2010). Für den T-Test gibt es drei mathematische Voraussetzungen (vgl. Rasch/Hoffmann/Friese/Naumann 2010): Zunächst muss das untersuchte Merkmal intervallskaliert, also die Abstufungen der Antwortmöglichkeiten äquidistant, sein. Zudem sollte das untersuchte Merkmal in der Population normalverteilt sein und die Populationsvarianzen, aus denen die beiden Stichproben stammen, gleich sein. Diese Varianzhomogenität kann bei unabhängigen Stichproben durch den Levene-Test geklärt werden. Sind diese Voraussetzungen erfüllt, ist das Verfahren des T-Tests exakt mathematisch ableitbar. Doch Simulationsstudien zeigen, dass der T-Test auch bei einer Verletzung dieser Voraussetzungen zuverlässige Informationen liefert und somit robust gegenüber Verletzungen reagiert. Dies gilt, wenn die Stichproben nicht zu klein gewählt sind ($N > 30$) und beim Vergleich zweier Gruppen annähernd dieselbe Größe vorliegt (vgl. Rasch/Hoffmann/Friese/Naumann 2010). Dies trifft auf die durchgeführte Erhebung zu.

Dabei überprüft der T-Test mit welcher Wahrscheinlichkeit die Nullhypothese (es besteht kein signifikanter Unterschied zwischen den beiden Gruppen) abgelehnt werden kann. Das dazugehörige Signifikanzniveau liegt üblicherweise bei $\alpha = 0{,}05$ ($\alpha = 5\%$). Für die Interpretation der in dieser Arbeit ermittelten Daten soll aber die Kategorisierung nach Töpfer dienen (vgl.Tabelle 3).

Tabelle 3: Abstufung zur Interpretation des Signifikanzwerts p und entsprechende Kennzeichnung (aus Töpfer 2012)

p		Interpretation	Kennzeichnung
> 0,1	(> 10 %)	nicht signifikant	n.s.
≤ 0,1	(≤ 10 %)	gering signifikant	g.s.
≤ 0,05	(≤ 5 %)	signifikant	*
≤ 0,01	(≤ 1 %)	sehr signifikant	**
≤ 0,001	(≤ 0,1 %)	höchst signifikant	***

Zumeist erfolgte ein Vergleich der einzelnen Ergebnisse zwischen Prä- und Posttest, um den Lernzuwachs erheben zu können. Um Unterschiede im Lernzuwachs zwischen den einzelnen Items erkennen zu können, wurde der hake-Faktor berechnet. Dieser Faktor betrachtet zum einen die durchschnittlichen Punktzahlen im Prä- sowie Posttest und zum anderen die maximal erreichbare Anzahl an Punkten der jeweiligen Items. Er gibt den relativen Zuwachs zwischen Prä- und Posttest an und dient als Maß für den normalisierten mittleren Lernzuwachs. Dieser hake-Faktor g lässt sich mit dem folgenden Algorithmus, welcher in Formel (7) dargestellt wurde, bestimmen (vgl. Kempin 2015):

$$\mathbf{g} = \frac{\textbf{erreichter Zuwachs}}{\textbf{maximal möglicher Zuwachs}} = \frac{\%\textbf{Posttest} - \%\textbf{Prätest}}{\mathbf{100}\% - \%\textbf{Prätest}} \tag{7}$$

Für die Interpretation des hake-Faktors soll die Kategorisierung nach Grüß-Niehaus dienen (vgl. Grüß-Niehaus 2010), welche in Tabelle 4 dargestellt wurde.

Tabelle 4: Abstufung zur Interpretation des hake-Faktors g und entsprechende Kennzeichnung (aus Grüß-Niehaus 2010)

Wertebereich	Beurteilung
g < 0,3	geringer Lernzuwachs
0,3 ≤ g < 0,7	mittlerer Lernzuwachs
0,7 ≤ g	hoher Lernzuwachs

Aufgrund der Tatsache, dass es Items gibt, bei denen jedoch im Durchschnitt zwischen Prä- und Posttest ein Punkteverlust zu verzeichnen ist, muss dieser Algorithmus zur Berechnung des hake-Faktors g angepasst werden. Die wurde in Formel (8) dargestellt (vgl. Kempin 2015):

$$\boldsymbol{g} = \begin{cases} \dfrac{\%\boldsymbol{Posttest} - \%\boldsymbol{Pr}\text{ä}\boldsymbol{test}}{\boldsymbol{100}\% - \%\boldsymbol{Pr}\text{ä}\boldsymbol{test}}, \boldsymbol{P}_{\boldsymbol{Pr}\text{ä}} < \boldsymbol{P}_{\boldsymbol{Post}} \\ \dfrac{\%\boldsymbol{Posttest} - \%\boldsymbol{Pr}\text{ä}\boldsymbol{test}}{\%\boldsymbol{Pr}\text{ä}\boldsymbol{test}}, \boldsymbol{P}_{\boldsymbol{Pr}\text{ä}} > \boldsymbol{P}_{\boldsymbol{Post}} \end{cases} \tag{8}$$

4 Ergebnisse

In diesem Kapitel werden die Ergebnisse der empirischen Erhebung präsentiert. Anknüpfend an die in Kapitel 3.4 beschriebene Methode der Datenanalyse, werden hierzu die Resultate der Fragebogenauswertung dargestellt.

Zudem hat sich gezeigt, dass die Ergebnisse normalverteilt sind. Mithilfe des Kolmogorov-Smirnov-Verfahrens ergab sich eine Asymptotische Signifikanz von 0,098 für den Prätest und von 0,066 für den Posttest. Somit konnten T-Tests zum Vergleich der jeweils gebildeten Mittelwerte erstellt werden.

Zunächst soll der Vergleich der Ergebnisse zwischen Prä- und Posttest bezogen auf die vorgeführten Erklärvideos erfolgen. Hierfür wurden die von den Probanden im Durchschnitt erreichten Punktzahlen bezogen auf das Erklärvideo guter Qualität (Video 1) und dasjenige schwacher Qualität (Video 2) in Abbildung 5 dargestellt.

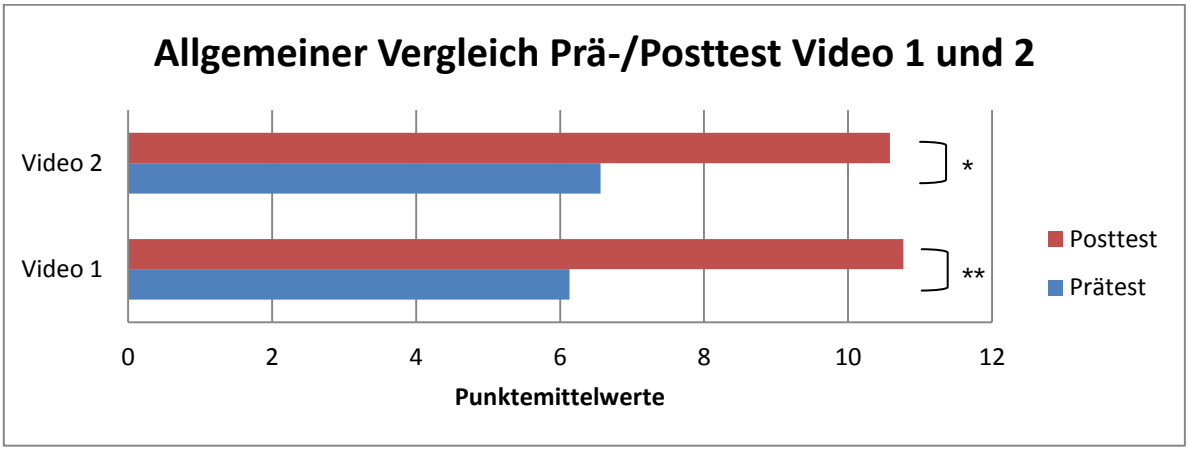

Abbildung 5: Mittelwertvergleich der arithmetischen Mittel der erreichten Gesamtpunktzahlen beim Prä- und Posttest zur Kurvenfahrt bezogen auf Video 1 und 2 unter Angabe der Signifikanz (2-seitig), auf einer Punkteskala von 0 bis 20; bei einer Gesamtprobandenzahl von N = 176.

Der Vergleich der erreichten Punktedifferenz zwischen Prä- und Posttest ergibt den jeweiligen Lernzuwachs bezogen auf Video 1 und 2.

Abbildung 5 ist zunächst zu entnehmen, dass sich die Leistungen im Prätest mit jeweils etwas über 6 Punkten nicht signifikant voneinander unterscheiden (p = 0,351). Davon ausgehend hat bei den Probanden, denen Video 1 vorgeführt wurde, ein deutlicher Lernzuwachs stattgefunden, da sich die Gesamtpunktzahl im Posttest höchst signifikant von derjenigen im Prätest unterscheidet. Allerdings konnte auch bei den Probanden, denen Video 2 vorgeführt wurde, ein signifikanter Lernzuwachs festge-

stellt werden. Untereinander unterscheiden sich die Leistungen im Posttest nicht signifikant voneinander (p = 0,71).

Konkreter wurde der Vergleich der Ergebnisse zu Video 1 und 2 in Abbildung 6 und 7 dargestellt, da hier die im Durchschnitt erreichten Punktzahlen aller Items aufgeführt sind. Dabei ist zu erwähnen, dass zwar bei den ersten sieben Items die Maximalpunktzahl 2 beträgt, bei Item 8 aber sogar 6 Punkte erzielt werden konnten.

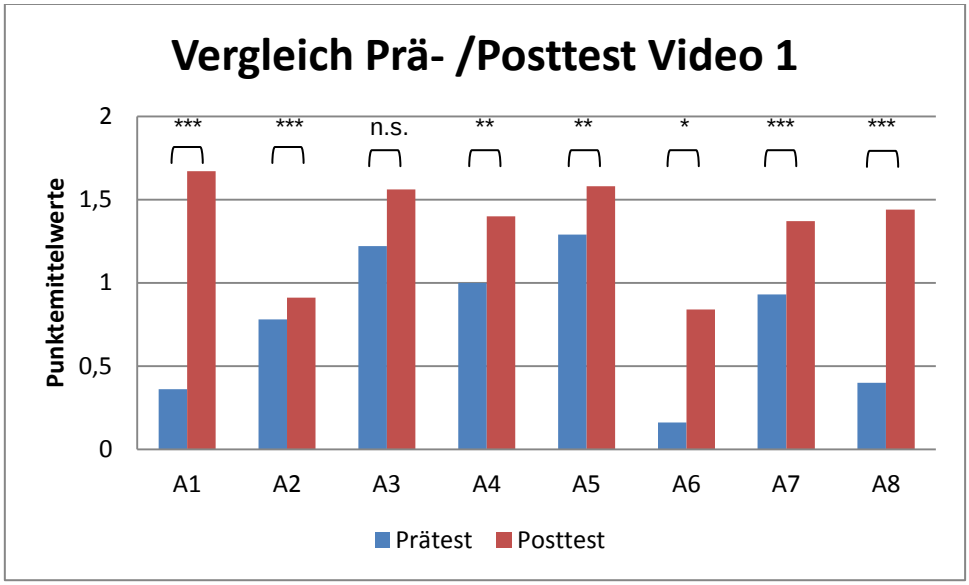

Abbildung 6: Mittelwertvergleich der arithmetischen Mittel der erreichten Punktzahlen der jeweiligen Aufgaben (Item A1 bis A8) beim Prä- und Posttest zur Kurvenfahrt bezogen auf Video 1 unter Angabe der Signifikanz (2-seitig); bei einer Probandenzahl von N = 90.

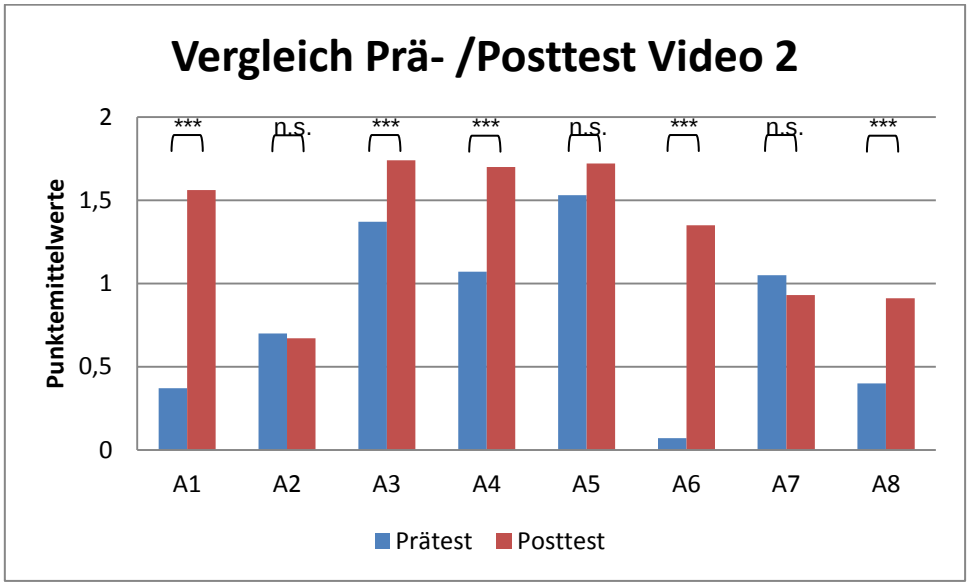

Abbildung 7: Mittelwertvergleich der arithmetischen Mittel der erreichten Punktzahlen der jeweiligen Aufgaben (Item A1 bis A8) beim Prä- und Posttest zur Kurvenfahrt bezogen auf Video 2 unter Angabe der Signifikanz (2-seitig); bei einer Probandenzahl von N = 86.

Abbildung 6 kann man entnehmen, dass der Punktezuwachs (Lernzuwachs) bei den Probanden, denen Video 1 vorgeführt wurde, bei fast allen Antworten sehr oder höchst signifikant war. Lediglich bei Item A3 ist der Unterschied zwischen Prä- und Posttest nicht signifikant.

In Abbildung 7 wird deutlich, dass es sich bei Video 2 anders verhält. Der Lernzuwachs ist bei fünf der acht Items höchst signifikant, bei drei hingegen nicht signifikant. Dabei handelt es sich um Item A2, A5 und A7. Bei Item A2 und A7 ist der Punktezuwachs zwischen Prä- und Posttest sogar negativ (vgl. Abbildung 7).

Zur Verdeutlichung und besseren Veranschaulichung des konkreten Lernzuwachses insgesamt wurde dieser bezogen auf Video 1 und 2 in Abbildung 8 vergleichend dargestellt.

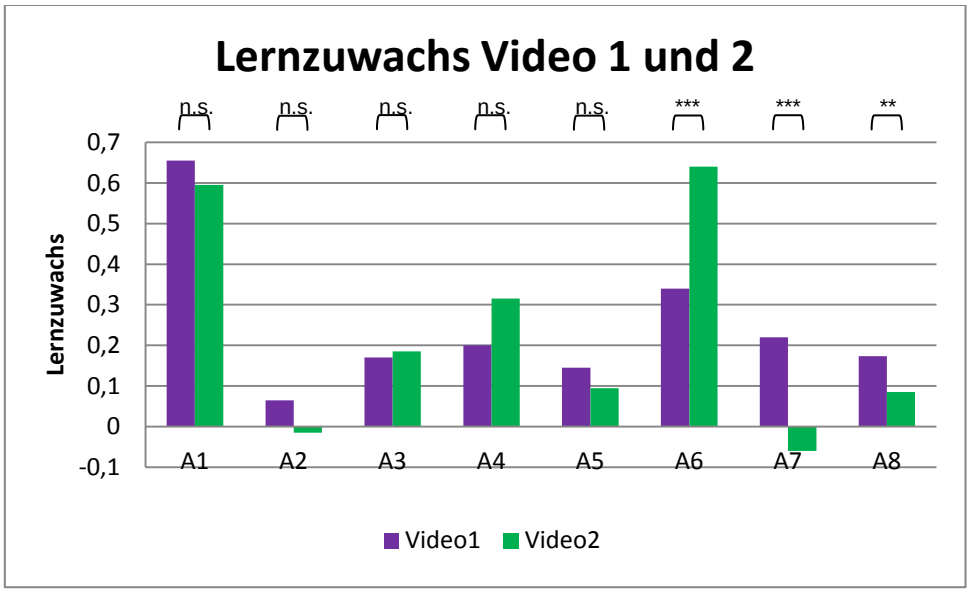

Abbildung 8: Normierter Mittelwertvergleich des Lernzuwachses zwischen Prä- und Posttest der jeweiligen Aufgaben zur Kurvenfahrt (Item A1 bis A8) bezogen auf Video 1 und 2 unter Angabe der Signifikanz (2-seitig); bei einer Gesamtprobandenzahl von N = 176.

Abbildung 8 kann man entnehmen, dass der Lernzuwachs der Probanden, bei den Items A1 (Video 1: 66%, Video 2: 60%) und A6 (Video 1: 34%, Video 2: 64 %) am größten ist. Besonders gering ist der Lernzuwachs hingegen bei Item A2 (Video 1: 7%, Video 2: (-2)%).

Allerdings wird auch deutlich, dass der Lernzuwachs bei den einzelnen Items zwischen Video 1 und 2 zumeist etwa gleich ist, da sich die Werte nicht signifikant voneinander unterscheiden. Dies gilt allerdings nicht für die letzten drei Items.

Bezogen auf Item A8 ist zu sagen, dass der Lernzuwachs als Punktedifferenz (Video 1: 1,0 Punkte, Video 2: 0,5 Punkte) zwar eigentlich recht hoch erscheint (vgl. Abbildung 6 und 7), doch prozentual gesehen eher gering ist (vgl. Abbildung 8), da bei dieser Aufgabe insgesamt 6 Punkte hätten erreicht werden können.

Im Folgenden wird dieser sich aus den bloßen Punktedifferenzen ergebende Lernzuwachs mit dem über den hake-Faktor errechneten Lernzuwachs verglichen (vgl. Formel (8), Abbildung 9).

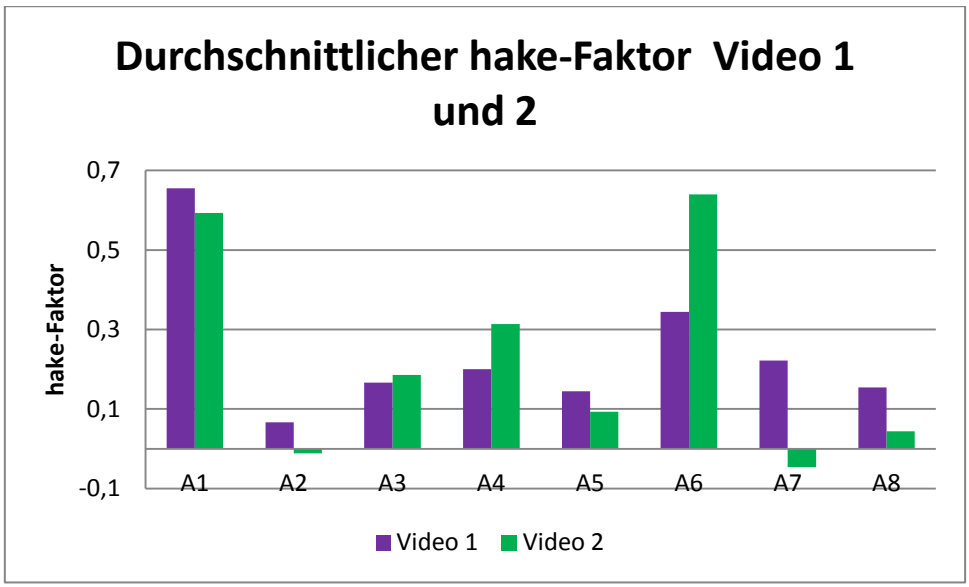

Abbildung 9: Normierter Mittelwertvergleich des Lernzuwachses über den hake-Faktors zwischen Prä- und Posttest der jeweiligen Aufgaben zur Kurvenfahrt (Item A1 bis A8) bezogen auf Video 1 und 2; bei einer Gesamtprobandenzahl von N = 176.

Vergleicht man den direkt aus den Punktedifferenzen ermittelten Lernzuwachs (vgl. Abbildung 8) mit dem berechneten Lernzuwachs über den hake-Faktor (vgl. Abbildung 9) wird deutlich, dass sich beide Diagramme kaum voneinander unterscheiden. Allerdings ermöglicht der hake-Faktor eine Klassifizierung des Lernzuwachses gemäß Tabelle 4.

Bei Item A1 konnte bei beiden Videos ein mittlerer Lernzuwachs erreicht werden (Video 1: g = 0,66, Video 2: g = 0,59). Dies gilt auch für Item A6 (Video 1: g = 0,34, Video 2: g = 0,64). Desweiteren konnte für Item A4 bei Video 2 mit einem hake-Faktor von g = 0,31 ebenfalls ein mittlerer Lernzuwachs verzeichnet werden. Dies gilt nicht für den entsprechenden Lernzuwachs bei Video 1, bei welchem der hake-Faktor g = 0,20 beträgt. Ein solch geringer Lernzuwachs ist auch bei den anderen dargestellten

Items zu verzeichnen. Ausnahme bilden diejenigen mit negativen hake-Faktoren. Dazu zählen Item A2 und A7 bei Video 2. Hier hat bei den Probanden im Gegensatz zu Video 1 kein nachweisbarer Lernzuwachs stattgefunden.

Vergleicht man allerdings den Lernzuwachs insgesamt zwischen Video 1 und Video 2 mit Hilfe des hake-Faktors, wird deutlich, dass die Probanden, denen Video 1 vorgeführt wurde, insgesamt einen mittleren Lernzuwachs erfahren (g = 0,31), wohingegen bei den Probanden, die Video 2 gesehen haben, sich lediglich ein geringer Lernzuwachs ergibt (g = 0,27). Dies ist in Abbildung 10 dargestellt:

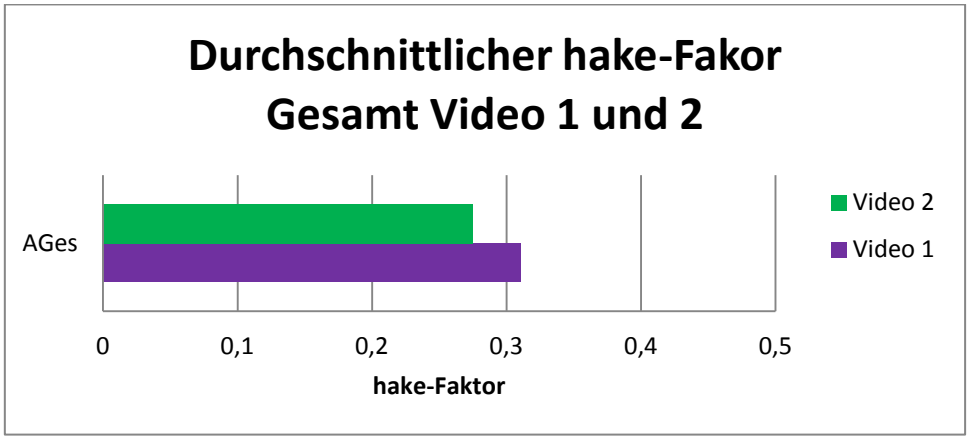

Abbildung 10: Normierter Mittelwertvergleich des Lernzuwachses über den hake-Faktors zwischen Prä- und Posttest aller Items (AGes) zur Kurvenfahrt bezogen auf Video 1 und 2; bei einer Gesamtprobandenzahl von N = 176.

Vertiefend soll eine geschlechterbezogene Überprüfung des Lernzuwachses erfolgen. Bereits in Kapitel 3.2 wurde geschildert, dass von den 176 Probanden 80 männlich und 96 weiblich sind. Zunächst ist die im Durchschnitt erreichte Punktzahl beim Prätest, unabhängig vom Video, in Abbildung 11 dargestellt.

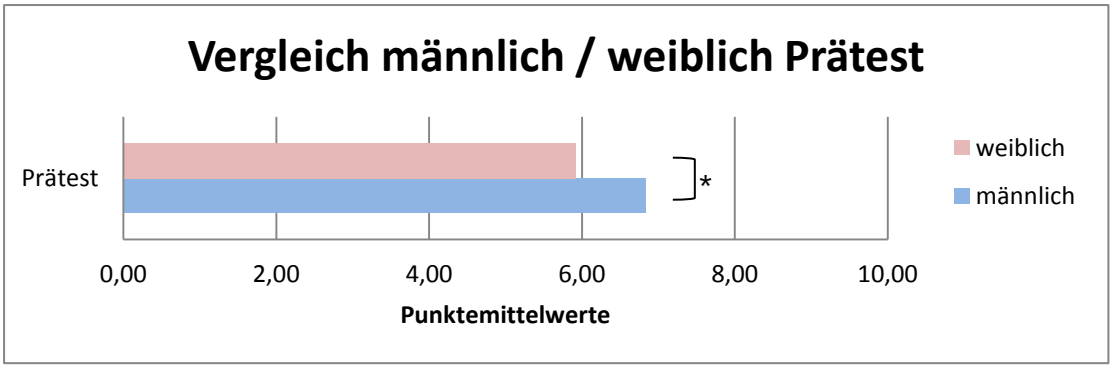

Abbildung 11: Mittelwertvergleich der arithmetischen Mittel der erreichten Gesamtpunktzahlen beim Prätest bezüglich des Geschlechts unter Angabe der Signifikanz (2-seitig), auf einer Punkteskala von 0 bis 20; bei einer Gesamtprobandenzahl von N = 176.

Abbildung 11 ist zu entnehmen, dass ein signifikanter Unterschied zwischen den Leistungen der Geschlechter im Prätest vorherrscht, die männlichen Probanden erreichten im Mittel (signifikant) höhere Punktzahlen als die weiblichen.
Konkreter wurde der Vergleich der Ergebnisse zum Prätest in Abbildung 12 dargestellt, da hier die von den männlichen und weiblichen Probanden im Durchschnitt erreichten Punktzahlen aller Items aufgeführt sind. Anzumerken ist, dass die Maximalpunktzahl der Items A1 bis A7 zwei Punkte betrug, bei A8 hingegen sechs.

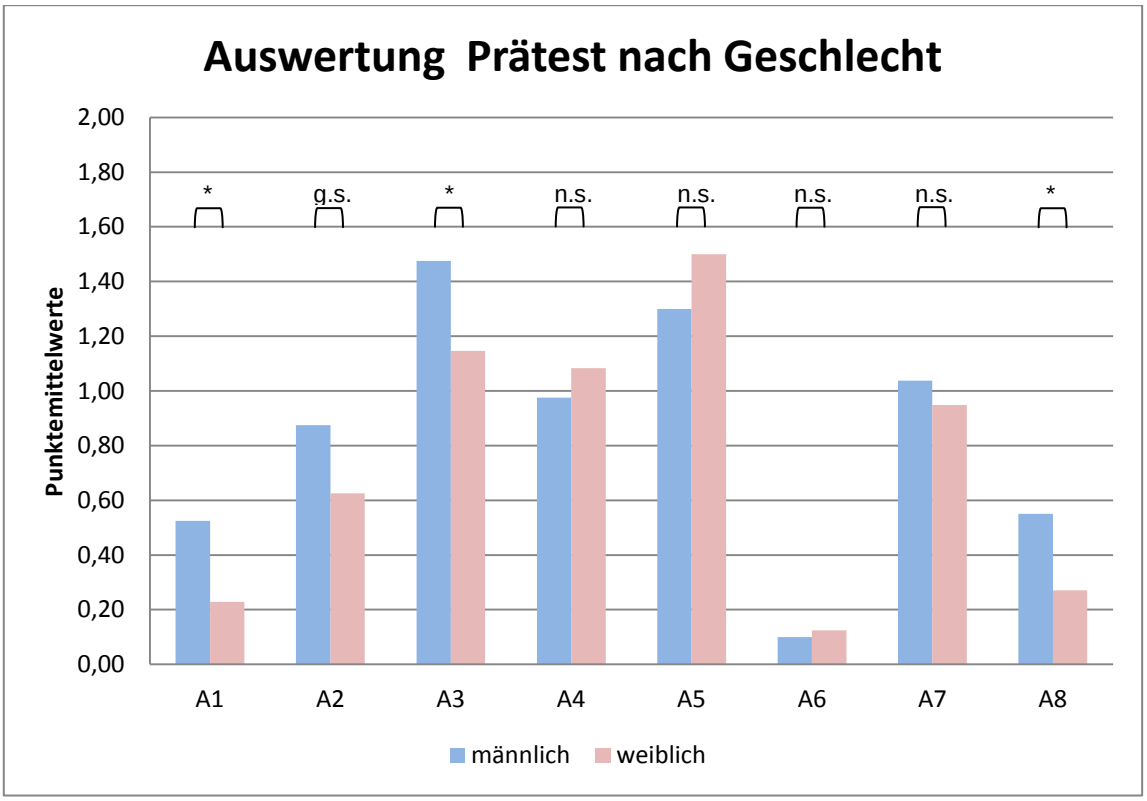

Abbildung 12: Mittelwertvergleich der arithmetischen Mittel der bei den einzelnen Aufgaben (Item A1 bis A8) erreichten Punktzahlen bezüglich des Geschlechts unter Angabe der Signifikanz (2-seitig); bei einer Gesamtprobandenzahl von N = 176.

Abbildung 12 kann man entnehmen, dass sich die Leistung der männlichen und weiblichen Probanden im Prätest lediglich bei vier der acht Items signifikant beziehungsweise gering signifikant (vgl. Item A2) voneinander unterscheidet. Bei Aufgabe A1, A2, A3 und A8 haben die Schüler (gering) signifikant mehr Punkte als die Schülerinnen erreicht. Bei den anderen Items sind die Unterschiede nicht signifikant.
In Abbildung 13 sind die im Durchschnitt erreichten Punkte der männlichen und weiblichen Probanden im Posttest vergleichend dargestellt. Hierbei wurde ebenfalls zwi-

schen den Leistungen der Probanden, denen Video 1 vorgeführt wurde und den Probanden, denen Video 2 vorgeführt wurde, unterschieden.

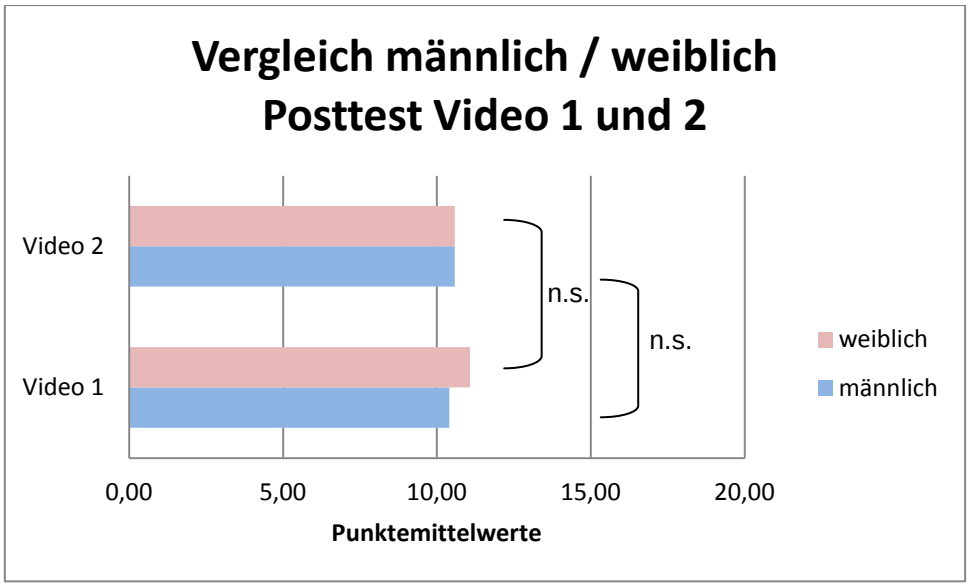

Abbildung 13: Mittelwertvergleich der arithmetischen Mittel der erreichten Gesamtpunktzahlen beim Posttest bezüglich des Geschlechts bezogen auf Video 1 und 2 unter Angabe der Signifikanz (2-seitig), auf einer Punkteskala von 0 bis 20; bei einer Gesamtprobandenzahl von N = 176.

Abbildung 13 kann man entnehmen, dass die erreichten Punktzahlen im Posttest bei beiden Geschlechtern und auch bei beiden Videos ähnlich sind. Signifikante Unterschiede gibt es nicht. Unabhängig vom Video waren die weiblichen Probanden etwa gleich stark. Dies gilt auch für den Vergleich zwischen den männlichen Probanden. Unabhängig vom Geschlecht war die Leistung der männlichen und weiblichen Probanden bei Video 1 annähernd gleich (kein signifikanter Unterschied). Dies gilt auch für die Probanden, denen Video 2 vorgeführt wurde.

Vergleicht man die Ergebnisse aus Abbildung 13 mit denen aus Abbildung 11 wird deutlich, dass die weiblichen Probanden im Prätest signifikant schlechtere Punktzahlen erreicht haben als die männlichen, im Posttest aber gleich gute Werte aufweisen. Demzufolge muss der Lernzuwachs bei den weiblichen Probanden höher sein als bei den männlichen. Dies konnte mit Hilfe der Berechnung der zugehörigen hake-Faktoren überprüft werden und ist in Abbildung 14 dargestellt.

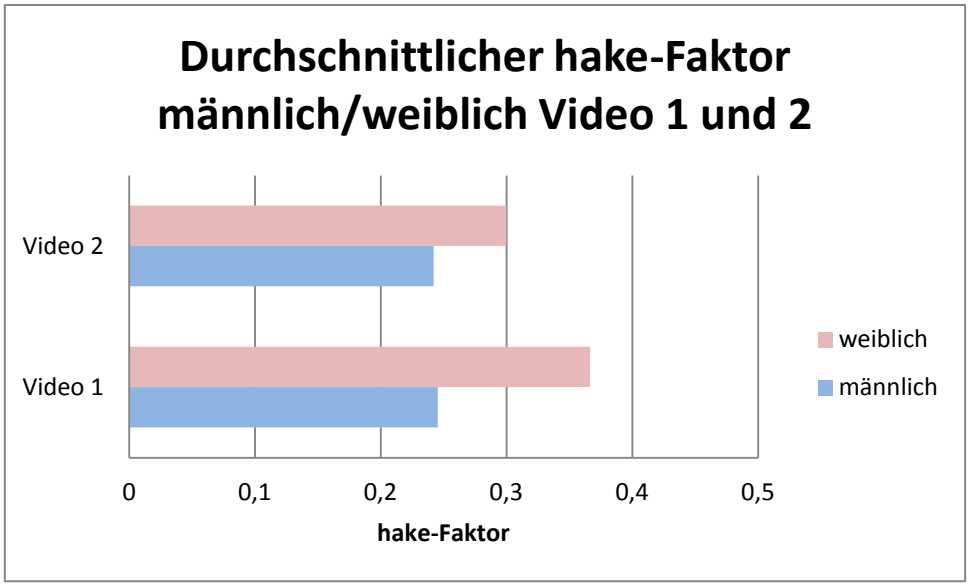

Abbildung 14: Normierter Mittelwertvergleich des Lernzuwachses über den hake-Faktors zwischen Prä- und Posttest zur Kurvenfahrt bezüglich des Geschlechts und bezogen auf Video 1 und 2; bei einer Gesamt-probandenzahl von N = 176.

Mit Hilfe von Abbildung 14 wird zunächst deutlich, dass der Lernzuwachs der männlichen Probanden bei Video 1 und 2 etwa gleich hoch ist. Es handelt sich jeweils um einen geringen Lernzuwachs (Video 1: g = 0,25, Video 2: g = 0,24). Der Unterschied im Lernzuwachs zwischen den Geschlechtern fällt besonders bei Video 1 auf. Hier weisen die weiblichen Probanden im Unterschied zu den männlichen einen mittleren Lernzuwachs auf (g = 0,37). Dies gilt auch für Video 2, wenngleich der Unterschied dort weniger deutlich ausfällt (g = 0,30).

Ebenfalls vertiefend soll eine altersbezogene Überprüfung des Lernzuwachses erfolgen. Die 176 Probanden waren zwischen 13 und 17 Jahren alt, das ermittelte Durchschnittsalter beträgt 15,02 Jahre (SD = 0,88). Die genaue Altersverteilung ist in Tabelle 5 dargestellt.

Tabelle 5: Übersicht zum Alter der Probanden

Alter in Jahren	13	14	15	16	17
Anzahl der Probanden	2	52	70	44	8

Tabelle 5 kann man entnehmen, dass nur zwei Schülerinnen und Schüler 13 und acht Schülerinnen und Schüler 17 Jahre alt waren. Daher sind die im weiteren Ver-

lauf getroffenen Aussagen nicht repräsentativ und stellen lediglich tendenzielle Aussagen dar.

Zunächst ist die im Durchschnitt erreichte Punktzahl beim Prätest der verschiedenen Altersstufen, unabhängig vom Video, in Abbildung 15 dargestellt.

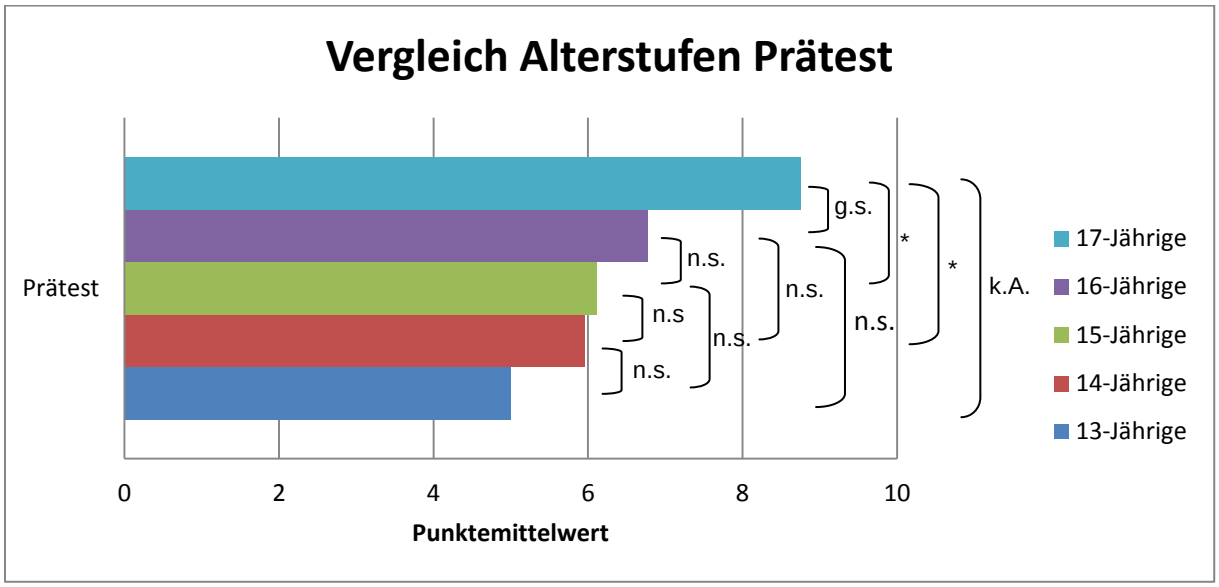

Abbildung 15: Mittelwertvergleich der arithmetischen Mittel der erreichten Gesamtpunktzahlen beim Prätest bezüglich der Altersstufe unter Angabe der Signifikanz (2-seitig), auf einer Punkteskala von 0 bis 20; bei einer Gesamtprobandenzahl von N = 176.

Abbildung 15 kann man entnehmen, dass sich die erreichten Punktzahlen im Posttest bei den jüngeren Probanden nicht signifikant unterscheiden. Lediglich die durchschnittliche Leistung der 17-Jährigen ist gering signifikant höher als die der 16-Jährigen und signifikant höher als die der 14- und 15-Jährigen. Bezogen auf die 13-Jährigen kann aus den bereits genannten Gründen keine Angabe erfolgen. Insgesamt wird eine Staffelung der Leistung nach Alter ersichtlich, auch wenn diese zumeist nicht signifikant ist.

In Abbildung 16 sind die im Durchschnitt erreichten Punkte der Probanden in Abhängigkeit vom jeweiligen Alter im Posttest vergleichend dargestellt. Hierbei wurde ebenfalls zwischen den Leistungen der Probanden, denen Video 1 vorgeführt wurde und den Probanden, denen Video 2 vorgeführt wurde, unterschieden.

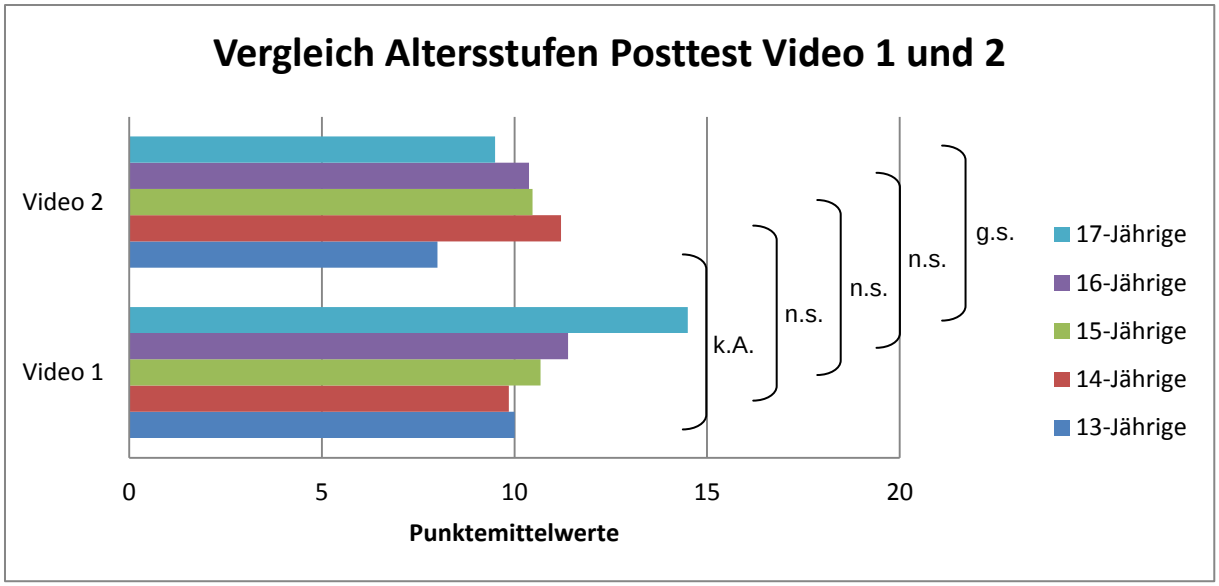

Abbildung 16: Mittelwertvergleich der arithmetischen Mittel der erreichten Gesamtpunktzahlen beim Posttest bezüglich der Altersstufe bezogen auf Video 1 und 2 unter Angabe der Signifikanz (2-seitig), auf einer Punkteskala von 0 bis 20; bei einer Gesamtprobandenzahl von N = 176.

In Abbildung 16 wird deutlich, dass sich die Leistungen der Probanden im Posttest in der Regel nicht signifikant voneinander unterscheiden – unabhängig davon, welches Video vorgeführt wurde. Lediglich bei den 17-Jährigen wird deutlich, dass die erreichte Punktzahl bei den Probanden, denen Video 1 vorgeführt wurde, zumindest gering signifikant höher ist als bei den Probanden, denen Video 2 vorgeführt wurde.

Zusätzlich können Aussagen zu den nicht dargestellten Vergleichen der Leistungen im Posttest der verschiedenen Altersstufen, bezogen auf das jeweilige Video, getroffen werden: Betrachtet man die erreichten Punkte im Posttest bei Video 1 wird deutlich, dass sich auch hier die Leistungen nicht signifikant voneinander unterscheiden. Lediglich die Leistungen der 17-Jährigen sind gering signifikant höher. Bei Video 2 unterscheiden sich alle erreichten Punktzahlen der verschiedenen Altersstufen nicht signifikant voneinander.

Um den konkreten Lernzuwachs vergleichen zu können, wurden die einzelnen hake-Faktoren der verschiedenen Altersstufen bezogen auf Video 1 und 2 berechnet und in Abbildung 17 dargestellt.

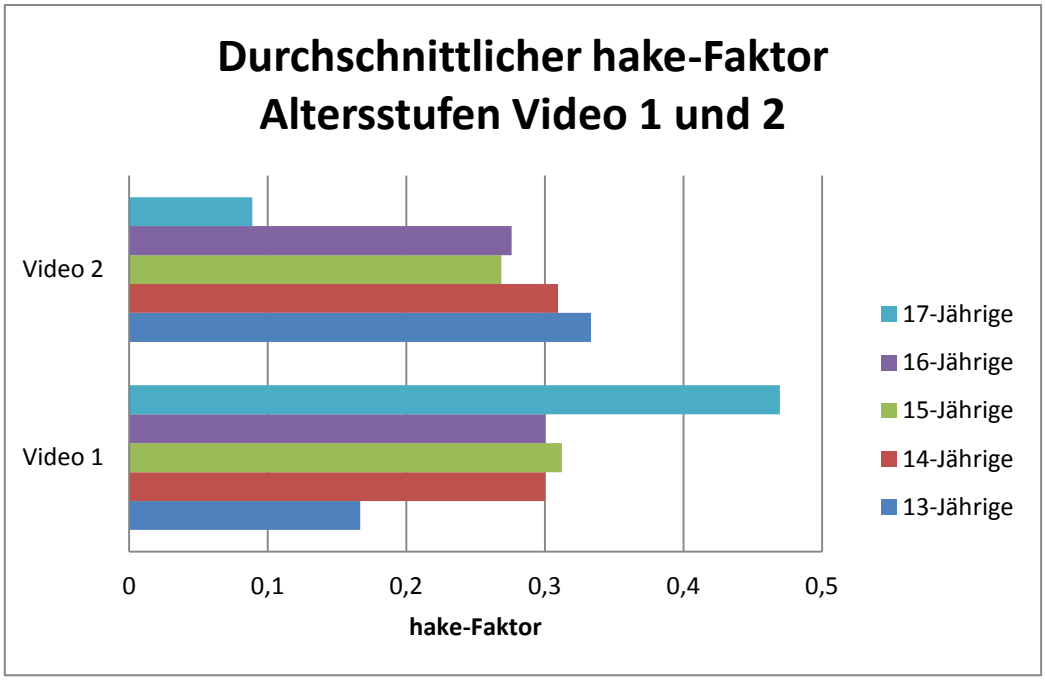

Abbildung 17: Normierter Mittelwertvergleich des Lernzuwachses über den hake-Faktors zwischen Prä- und Posttest zur Kurvenfahrt bezüglich der Altersstufe und bezogen auf Video 1 und 2; bei einer Gesamtprobandenzahl von N = 176.

Abbildung 17 kann man entnehmen, dass bei Video 1, unabhängig von der Altersstufe, ein mittlerer Lernzuwachs zu verzeichnen ist. Dies gilt nicht für die Gruppe der 13-Jährigen. Da es sich bezogen auf Video 1 aber nur um einen Probanden handelt, wird der entsprechende hake-Faktor nicht in die Betrachtungen einbezogen. Dieser mittlere Lernzuwachs ist bei der Gruppe der 17-Jährigen, denen Video 1 vorgeführt wurde, besonders deutlich.

Bezogen auf Video 2 haben die Probanden in der Regel nur einen geringen Lernzuwachs erfahren. Lediglich die Gruppe der 14-Jährigen (sowie der nicht betrachtete 13-jährige Proband) zeigen einen mittleren Lernzuwachs.

5 Diskussion

Im vorangegangenen Kapitel wurden die Ergebnisse der Empirischen Erhebung dargestellt. Diese sollen im aktuellen Kapitel diskutiert werden.
Zunächst soll Bezug zu den aufgestellten Hypothesen genommen werden (vgl. Kapitel 3.1). Hierzu wird untersucht, ob diese auf Grundlage der Ergebnisse der durchgeführten Erhebung verifiziert oder falsifiziert werden. Die zu Grunde liegende Forschungsfrage lautete, ob ein Erklärvideo zur Kurvenfahrt, das gemäß des zu validierenden Qualitätsmaßstabs von guter Qualität ist (Video 1), einen höheren Lernzuwachs als ein entsprechendes Erklärvideo von minderer Qualität (Video 2) erzielt. Die zugehörige Hypothese – zunächst ohne die entsprechenden Unterhypothesen – lautete:

Hypothese: *Wenn ein Erklärvideo zur Kurvenfahrt gemäß des Qualitätsmaßstabs von guter Qualität ist, dann ist der Lernzuwachs der Probanden höher, als wenn das Erklärvideo von schwacher Qualität ist.*

Abbildung 10 kann man entnehmen, dass die Probanden, denen das Video guter Qualität vorgeführt wurde, insgesamt einen mittleren Lernzuwachs aufweisen (g = 0,31). Die Probanden, denen das Video schwacher Qualität vorgeführt wurde, erzielten hingegen nur einen geringen Lernzuwachs (g = 0,27). Somit ist der mit Hilfe des hake-Faktors ermittelte Lernzuwachs durch das Video guter Qualität insgesamt höher als derjenige durch das Video von schwacher Qualität. Aus diesem Grund konnte diese Hypothese verifiziert werden.

Diese Ergebnisse ermöglichen auch die Einschätzung der zugehörigen Unterhypothesen:

Unterhypothese 1: *Wenn die Probanden ein Erklärvideo zur Kurvenfahrt mit guter Qualität sehen, dann ist ihr Lernzuwachs im Durchschnitt sehr hoch.*

Unterhypothese 2: *Wenn die Probanden ein Erklärvideo zur Kurvenfahrt mit schwacher Qualität sehen, dann ist ihr Lernzuwachs im Durchschnitt nicht vorhanden.*

Da beim Video von guter Qualität nur ein mittlerer und kein hoher Lernzuwachs ($g = 0{,}31 < 0{,}7$) erzielt werden konnte, muss die Unterhypothese 1 falsifiziert werden.
Da beim Video von schwacher Qualität ein geringer, aber dennoch nachweisbarer Lernzuwachs ($g = 0{,}27 > 0$) erzielt werden konnte, muss die Unterhypothese 2 ebenfalls falsifiziert werden.
Obwohl beide Unterhypothesen falsifiziert werden mussten, ist die ursprüngliche Hypothese dennoch korrekt: Das Video guter Qualität hat auf Grundlage der Klassifizierung nach hake-Faktor einen höheren Lernzuwachs als dasjenige schwacher Qualität bewirkt. Somit gilt die Forschungsfrage als beantwortet.

Betrachtet man die zu Grunde liegenden Ergebnisse wird deutlich, dass der Unterschied zwischen dem Lernzuwachs durch Video 1 und dem Lernzuwachs durch Video 2 nicht besonders groß ist. Dies soll im Folgenden genauer analysiert werden.
Aus Abbildung 5 kann man zunächst entnehmen, dass sich die Leistungen der Probanden zwischen Video 1 und Video 2 im Prätest und auch im Posttest nicht signifikant voneinander unterscheiden. Allerdings sind die erreichten Punkte im Posttest, auch wenn sie nur im ausreichenden Bereich liegen, bei beiden Videos signifikant beziehungsweise sehr signifikant höher als diejenigen im Prätest. Die Differenz beider Werte ergibt den jeweiligen Lernzuwachs. Statistisch ist dieser Lernzuwachs zwischen Video 1 und Video 2 nicht signifikant. Eine auf die verschiedenen Items anteilig verteilte Darstellung des Lernzuwachses erfolgte in Abbildung 8. Ein Vergleich mit den entsprechenden Lernzuwächsen, welche mit Hilfe des hake-Faktors gewonnen wurden (vgl. Abbildung 9), ergab nahezu identische Ergebnisse. Diese konnten mit Hilfe von Tabelle 4 interpretiert und der Lernzuwachs klassifiziert werden. Das gilt auch für den Lernzuwachs insgesamt (vgl. Abbildung 10), welcher bereits hypothesenbezogen diskutiert wurde.

Um den Lernzuwachs genauer zu untersuchen, wurden zudem geschlechtsspezifische Analysen durchgeführt. Abbildung 11 ist zu entnehmen, dass die männlichen Probanden signifikant höhere Punktzahlen im Prätest erreicht haben als die weiblichen. Im Posttest war der Unterschied zwischen den Geschlechtern bei beiden Videos nicht signifikant. Zudem waren die weiblichen Probanden unabhängig vom Video etwa gleich stark. Dies gilt auch für den Vergleich zwischen den männlichen Probanden (vgl. Abbildung 13). Für den Lernzuwachs bedeutet dies, dass die männ-

lichen Probanden unabhängig vom Video einen geringen Lernzuwachs, die weiblichen einen mittleren Lernzuwachs, erworben haben. Dabei ist dieser deutliche Lernzuwachs der weiblichen Probanden bei Video 1 höher als bei Video 2 (vgl. Abbildung 14). Somit begründet sich der bereits festgestellte mittlere Lernzuwachs aller Probanden, denen Video 1 vorgeführt wurde im Gegensatz zum geringen Lernzuwachs bei Video 2 auf diesen geschlechtsbezogenen Ergebnissen.

Bezüglich der Altersstufen lässt sich ergänzen, dass die Probandengruppen beim jeweiligen Video in ihren Leistungen recht homogen waren. Betrachtet man diejenigen Altersstufen mit besonders wenig Probanden nicht (vgl. Tabelle 5), wird deutlich, dass alle Probanden denen Video 1 vorgeführt wurde, einen mittleren Lernzuwachs und nahezu alle denen Video 2 vorgeführt wurde einen geringen Lernzuwachs, erfahren haben.

Im Folgenden soll der Versuch einer möglichen Erklärung der analysierten Ergebnisse erfolgen. Hierbei soll der in Abbildung 4 dargestellte Prozess des Erklärens mit Vermittlungscharakter von Physik (vgl. Kulgemeyer/Tomcyszyn 2015) die Grundlage darstellen. Das Vorwissen und das Interesse des Adressaten am Lerngegenstand stellen, neben der Erklärung an sich, wichtige Einflussfaktoren für einen optimalen Lernzuwachs dar.
Die Empirische Erhebung hat ergeben, dass das Vorwissen der Schülerinnen und Schüler zur Kurvenfahrt im Durchschnitt mangelhaft ist (vgl. Abbildung 5). Besonders die weiblichen Probanden haben vergleichsweise wenig Punkte im Prätest erreicht (vgl. Abbildung 11). Studien haben ergeben, dass nicht nur das Vorwissen (vgl. Kröll 2010) im Fach Physik bei den Mädchen schlechter ausgeprägt ist als bei den Jungen, sondern auch das Interesse am Fach und dessen Inhalten (vgl. Elster 2007; vgl. Holstermann/Bögeholz 2007). Das Schulfach Physik gilt sogar als unbeliebtestes Fach bei den Mädchen (vgl. Muckenfuss 2006). Dies könnte damit begründet werden, dass Physik als männliche Domäne gilt, denn die typischen Eigenschaften dieser harten Naturwissenschaft werden auch den Jungs zugeschrieben (vgl. Merzyn 2008). Dementgegen beschäftigen sich die Mädchen von Natur aus weniger mit dem Themenbereich der Physik und weisen somit auch ein vergleichsweise geringes Vorwissen im Gegensatz zu den Jungen auf. Auch entspricht der gewählte Kontext einer Kurvenfahrt mit einem Motorrad eher männlichen Interessen.

Den höheren Lernzuwachs seitens der weiblichen Probanden könnte man damit erklären, dass sich die männlichen Probanden der Erhebung möglicherweise weniger intensiv gewidmet haben. Denn im Gegensatz zu Jungen, sind Mädchen häufig zuverlässiger und fleißiger (vgl. Elster 2007). Insgesamt ist festzustellen, dass beide Videos, aber besonders das Video von guter Qualität, dazu geführt haben, dass die Mädchen ihr anfängliches Wissensdefizit überwinden und im Posttest genauso gute Ergebnisse erlangen konnten wie die männlichen Probanden (vgl. Abbildung 13).

Zwar ergab die durchgeführte Erhebung altersunabhängige Leistungen im Prä- und Posttest, doch bezieht man die vergleichsweise überdurchschnittlichen Ergebnisse der Gruppe der 17-Jährigen mit ein, könnte eine mögliche Erklärung die nahe Verwandtschaft des Themas mit der Problematik zur Erlangung des Führerscheins sein. Dies ist für diese Altersgruppe von besonderer Bedeutung, somit weisen sie ein vergleichsweise höheres Interesse und somit Vorwissen auf (vgl. Abbildung 15). Desweiteren haben sich diese Probanden, die die Einführungsphase der Oberstufe besuchen, bewusst für Physik als weiterführendes Unterrichtsfach entschieden.

Neben diesen sich wechselseitig beeinflussenden Faktoren des Vorwissens und Interesses spielt auch das Erklären im jeweiligen Video eine bedeutende Rolle. Im Folgenden soll analysiert werden, warum die Probanden im Posttest im Durchschnitt nicht über ausreichende Leistungen hinauskamen – unabhängig von Alter und Geschlecht. Berücksichtigt man zunächst das für die Erklärung gewählte Medium, das Erklärvideo, muss beachtet werden, dass es sich hierbei in beiden Fällen um eine 6,5-minütige Intervention und keine über mehrere Stunden angelegte Unterrichtseinheit handelt. Dies stellt eine erste Ursache für das immer noch verhältnismäßig schlechte Abschneiden aller Probanden im Posttest dar. Setzt man das Medium des Erklärvideos in Bezug zu den 10 Kriterien „guten Erklärens“ (vgl. Kulgemeyer 2013; Leisen 2013) so fällt auf, dass letztere auf die Interaktion zwischen Erklärer und Adressat abzielen. Bei dem gewählten Medium ist es aber nicht möglich, aktiv auf die Adressaten einzugehen. Demzufolge konnte keine optimale Adressatengerechtheit gewährleistet werden. Aus diesem Grund konnten auch nur 8 der 12 Punkte auf der Skala des Qualitätsmaßstabs für „gutes Erklären“ bei Video 1 erreicht werden. Somit handelt es sich nicht um das optimale Medium, um besagten Qualitätsmaßstab zu validieren. Da beide Videos sich nur bezüglich der Adressatengerechtheit und nicht

bezüglich der Sachgemäßheit unterscheiden, lässt sich somit auch der geringe, aber vorhandene Lernzuwachs der Probanden erklären, denen Video 2 vorgeführt wurde. Berücksichtigt man neben dem Medium des Erklärvideos zudem das Erhebungsinstrument des Fragebogens, sind auch hier einige Fehlerquellen zu diskutieren. Zwar stellt die Fragebogenmethode ein geeignetes Messinstrument zur Erhebung des Fachwissens dar, dennoch birgt er beispielsweise die Gefahr des Abschreibens. Außerdem sind möglicherweise nicht alle Schülerinnen und Schüler gleichermaßen motiviert einen solchen Bogen zu bearbeiten. Abgesehen davon könnte der Schwierigkeitsgrad des Fragebogens insgesamt oder einzelner Items für diese Lerngruppe nicht optimal sein. Aus diesem Grund werden im Folgenden abschließend die verwendeten Items bezogen auf die entsprechenden Formulierungen im Erklärvideo 1 und 2 diskutiert. Dabei liegt der Fokus auf den Items, bei denen signifikante Unterschiede in den Ergebnissen zwischen den Probanden, denen Video 1 vorgeführt wurde und denjenigen, denen Video 2 vorgeführt wurde, auftraten (vgl. Abbildung 8). Dies sind, wie bereits in Kapitel 4 beschrieben, die Items A6, A7 und A8. Desweiteren werden die Items mit einem negativen Lernzuwachs betrachtet (A2, A7 Video 2).

Der größte Unterschied im Lernzuwachs der Probanden zwischen denjenigen, denen Video 1 und denjenigen denen Video 2 vorgeführt wurde, trat bei Item A6 auf. Obwohl als Erklärung im Video guter Qualität ein Experiment durchgeführt und anhand dessen die Zentripetalkraft erläutert wurde, ist der Lernzuwachs (g = 0,34) deutlich geringer als im zweiten Video (g = 0,64), in welchem die Zentripetalkraft nur anhand eines erklärenden Satzes erläutert wurde: „*Die Kraft, die dafür notwendig ist, ist die sogenannte Zentripetalkraft und zeigt immer in Richtung des Kreismittelpunkts*“. Aus diesem Grund handelte es sich beim Bearbeiten dieses Items für die Probanden mit Video 1 um eine Anwendungsaufgabe, wohingegen die Probanden mit Video 2 vor allem reproduzieren mussten. Obgleich des großen Unterschieds im Zahlenwert der hake-Faktoren, befinden sich die Lernzuwächse beider Videos im mittleren Bereich.

Desweiteren unterscheidet sich der Lernzuwachs bei Item A7 zwischen den Probanden, denen Video 1 vorgeführt wurde, signifikant von den Probanden, denen Video 2 gezeigt wurde. Auffällig hierbei ist, dass es sich bei dem Video guter Qualität um einen positiven, geringen Lernzuwachs handelte (g = 0,22), während im Video schwacher Qualität ein negativer Lernzuwachs zu verzeichnen ist (g = -0,05). Für Video 1

wurde die Erläuterung für das „Herausfliegen“ aus der Kurve nur nebenbei erwähnt. So musste beim Bearbeiten des Fragebogens eine Transferleistung von der ausgehenden Erklärung zur Kreisbewegung geleistet werden. Diese lautete: „*Je größer die Zentripetalkraft, desto geringer ist der Radius der Kreisbahn, die der Motorradfahrer fährt*“ und stand in Verbindung mit der Erklärung für die Reibung: „*Fährt der Motorradfahrer nun hier lang (auf die nasse Stelle zeigend), so verringert sich der Reibungskoeffizient kurzfristig zwischen Reifen und Asphalt. Dadurch wird nicht mehr die gleiche Kraft auf das Motorrad ausgeübt*“. Dieses Phänomen wurde im zweiten Video mithilfe von Formeln, zum einen für die Reibungskraft (vgl. Formel (3)), zum anderen für die Zentripetalkraft (vgl. Formel (1)), erläutert. Dabei wurden diese beiden Formeln gleichgesetzt und die Antiproportionalität zwischen dem Radius r und dem Haftreibungskoeffizienten μ herausgestellt. Somit mussten die Probanden des zweiten Videos zuerst das Gleichsetzen der Formeln nachvollziehen und anschließend zurück auf die Aufgabe transferieren. Erklären könnte man dies zusätzlich mit dem geringen Interesse von Schülerinnen und Schülern an mathematischen Themen. Dieses liegt üblicherweise unter dem durchschnittlichen Interesse an anderen Schulfächern (vgl. Holstermann/Bögeholz 2007). Die damit zusammenhängenden Schwierigkeiten im Umgang mit Formeln könnten die Probleme beim Verständnis der Erklärung noch verstärkt haben (vgl. Fritz/Ricken/Schmidt 2009).

In der achten und letzten Aufgabe des Fragebogens (Item A8) sollte das Phänomen der Kurvenfahrt als Ganzes in eigenen Worten beschrieben werden. Folglich mussten alle vorher erwähnten Aspekte von den Probanden verstanden, verknüpft und folgerichtig angewendet werden. Beide Stichgruppen haben bei der Bearbeitung dieser komplexen Aufgabe entsprechend große Probleme (Video 1: g = 0,15, Video 2: g = 0,04). Der signifikante Unterschied im Lernzuwachs, wenngleich es sich bei beiden Videos um einen geringen Lernzuwachs handelt, lässt sich möglicherweise doch mit der Qualität des verständlicheren Videos 1 erklären.

Abschließend soll sich mit einer möglichen Erklärung des negativen Punktezuwachses bei Item A2 und A7 seitens der Probanden, denen Video 2 vorgeführt wurde, beschäftigt werden. Die Diskussion zum Item A7 ist bereits erfolgt. Der negative Lernzuwachs wird als nicht vorhandener Lernzuwachs eingestuft und ergibt sich höchstwahrscheinlich aus dem unterschiedlichen „Raten“ der Probanden bei der Bearbei-

tung des Fragebogens zwischen Prä- und Posttest. Bezogen auf Item A2 ist zusätzlich zu erwähnen, dass weder in Video 1 noch in Video 2 explizit auf die zweite Aufgabe eingegangen wurde. Hierbei handelte es sich um einen anderen Kontext, sodass in beiden Fällen ein Transfer von der Erläuterung der Trägheit zur Aufgabe erbracht werden musste. Aufgrund dieser Tatsache ist wahrscheinlich bei beiden Videos kein oder nur ein äußerst geringer Lernzuwachs (Video 1: $g = 0{,}07$, Video 2: $g = -0{,}04$) aufgetreten, die sich statistisch nicht voneinander unterscheiden.

6 Fazit und Ausblick

Die vorliegende Arbeit wurde verfasst, um den Qualitätsmaßstab für „gutes Erklären“ am Beispiel der Kurvenfahrt mit Hilfe von entsprechenden Videos validieren zu können. Die gewonnenen zentralen Ergebnisse sollen an dieser Stelle resümiert und ein entsprechender Ausblick gegeben werden.

Den Ausgangspunkt bildete das im fachdidaktischen Teil erarbeitet Modell zum Prozess des Erklärens mit Vermittlungscharakter von Physik. Hier wurde formuliert, dass das Erklären sich stets durch Sach- und Adressatengerechtheit auszeichnet, zusätzlich aber die Faktoren Vorwissen und Interesse der Probanden einen besonderen Einfluss haben. Auf dieser Grundlage wurde im Projekt EWis der zu validierende Qualitätsmaßstab erstellt, welcher sich in 12 Punkten der Adressatengerechtheit widmet.

In der durchgeführten Empirischen Erhebung wurde lediglich der Faktor des Vorwissens mit Hilfe eines Prätests erhoben und mit dem durch den entsprechenden Posttest erhobenen Fachwissen, nach einem Input in Videoform, verglichen. Die verwendeten Videos zum Thema Kurvenfahrt wurden selbst erstellt, eines war gemäß Maßstab von guter Qualität (8 von 12 Punkten) und das andere von schwacher Qualität (0 von 12 Punkten).

Es konnte gezeigt werden, dass mit Hilfe des Videos guter Qualität ein höherer Lernzuwachs als durch das Video schwacher Qualität erreicht werden konnte. Der jeweils über den hake-Faktor klassifizierte Lernzuwachs war beim Video guter Qualität im mittleren, beim Video schwacher Qualität im geringen Bereich. Somit konnte der Qualitätsmaßstab im engeren Sinne validiert werden. Im weiteren Sinne gilt dies jedoch nicht, da die Unterschiede im Lernzuwachs zwischen beiden Videos zu gering und somit statistisch nicht signifikant waren. Zudem ergab sich auch für das Video guter Qualität nicht der erwartete hohe Lernzuwachs. Beim Video schwacher Qualität wurde ein nicht vorhandener Lernzuwachs erwartet, aber nicht bestätigt. Als mögliche Ursache wurde die in beiden Erklärvideos gleichermaßen vorhandene Sachgerechtheit diskutiert.

Die Untersuchung ergab zum einen eine Leistungshomogenität im Prä- und Posttest bezüglich des Alters. Zudem wurde festgestellt, dass die weiblichen Probanden im Prätest signifikant schlechtere Leistungen als die männlichen Probanden erbracht haben. Der Posttest ergab unabhängig vom Geschlecht und vorgeführtem Video gleiche Leistungen. Die Ergebnisse im Prätest konnten durch das in der Regel geringere Interesse und Vorwissen von Mädchen allgemein im Fach Physik erklärt werden. Die dem entgegenstehende Leistungshomogenität im Posttest wurde durch den Fleiß und die Zuverlässigkeit der Mädchen erklärt. Dennoch ist festzuhalten, dass unabhängig vom Geschlecht und vom vorgeführten Video die Leistungen im Prätest im mangelhaften, die Leistungen im Posttest auch nur im ausreichenden Notenbereich lagen.

Dass die Leistungen der Probanden im Durchschnitt nicht über den ausreichenden Bereich hinaus gesteigert werden konnten, ließ sich mit Hilfe der gewählten Methode, dem Medium der Erklärvideos, erläutern. So konnte aufgrund dieser Wahl auch beim Video guter Qualität keine optimale Adressatengerechtheit (12 von 12 Punkten) gewährleistet werden. Aus diesem Grund wurde eine Validierung des Qualitätsmaßstabs für „gutes Erklären“ mit Hilfe von Erklärvideos zwar angestrebt, aber nicht vollends erreicht.

Abschließend gilt es in Form eines Ausblicks alternative Möglichkeiten vorzuschlagen. Die Ergebnisse und deren Analyse haben gezeigt, dass zu einer optimalen Validierung des Qualitätsmaßstabs künftig reale Erklärsituationen geschaffen und aufgezeichnet werden sollten, anstatt fertig gedrehte Erklärvideos – ohne echte Möglichkeit vollkommener Adressatengerechtheit – zu verwenden. Allerdings sollte der Erklärer die 10 Kriterien „guten Erklärens“ verinnerlicht haben und diese in der entsprechenden Erklärsituation anwenden, damit wirklich annähernd alle der 12 Punkte des Qualitätsmaßstabs erreicht werden. Bezüglich des Beispielthemas Kurvenfahrt könnte zudem ein anderer, adressatenspezifischer Kontext gewählt werden. Statt eines Motorrads ist beispielsweise die Verwendung eines Fahrrads alltagsgemäßer für die angestrebte Zielgruppe. Speziell Mädchen könnten zudem durch die Verwendung von Inlineskates bei der Kurvenfahrt eher motiviert werden.

7 Literaturverzeichnis

Alonso, M. / Finn. E.J. (2000). Physik, Oldenbourg Verlag, München, Wien.

Baid, J. (1988). Teachers in science Education. In: Kulgemeyer, Christoph: Gelingensbedingungen physikalischer Erklärungen – Zu einer kostruktivistischen Auffassung des Erklärens.

Bartels, H. (2015). Zur Validität und Objektivität eines Maßes für Erklärungsfähigkeit, Universität Bremen.

Bickel-Sandkötter, S. (2003). Computer, Internet & Co. Im Biologieunterricht. Cornelson Skriptor, Berlin.

Bolte, F. / Meldau, H. (1925). Physik für Seefahrer unter besonderer Berücksichtigung der Funkentelegraphie, Springer Verlag, Wiesbaden.

Bortz, J. / Döring, N. (2006). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler, 4. Auflage, Springer Medizin Verlag, Heidelberg.

Braun, J. (2015). Über die Erstellung und Erprobung einer Bewertungsmethode zur Prüfung der Sachgerechtheit einer Erklärung, Universität Bremen.

Danner, D. (2015). Reliabilität – die Genauigkeit einer Messung, Leibniz Institut für Sozialwissenschaften Mannheim.

Demtröder, W. (2015). Experimentalphysik 1, 7. Auflage, Springer-Verlag, Berlin, Heidelberg.

Die Senatorin für Bildung und Wissenschaft (2006): Physik. Bildungsplan für die Gymnasiale Oberstufe - Einführungsphase, Bremen.

Die Senatorin für Bildung und Wissenschaft (2010): Naturwissenschaften – Biologie, Chemie, Physik. Bildungsplan für die Oberschule, Bremen.

Elster, D. (2007). Interessante und weniger interessante Kontexte für das Lernen von Naturwissenschaften. Erste Ergebnisse der deutschen ROSE- Erhebung. In: Der mathematische und naturwissenschaftliche Unterricht.

Fritz, A./Ricken, G./Schmidt, S. (2009). Handbuch Rechenschwäche, Beltz Verlag, Weinheim, Basel

Gage, N. (1968). The microcriterion of effectiveness in explaining. In: Kulgemeyer, Christoph: Gelingensbedingungen physikalischer Erklärungen – Zu einer kostruktivistischen Auffassung des Erklärens.

Grüß-Niehaus, T. (2010). Zum Verständnis der Löslichkeitskonzepts im Chemieunterricht – der Effekt von Methoden progressiver und kallaborativer Reflexion – Logos Verlag, Berlin.

Hammer, K. / Hammer, H. (1995). Grundkurs der Physik 1. Oldenbourg Verlag, München, Wien.

Heintze, J. / Bock, P. (2014). Lehrbuch zur Experimentalphysik Band 1: Mechanik. Springer-Verlag, Berlin, Heidelberg.

Holstermann, N. / Bögeholz, S. (2007). Interesse von Jungen und Mädchen an naturwissenschaftlichen Themen am Ende der Sekundarstufe I. In: Zeitschrift für Didaktik der Naturwissenschaften.

Jenßen, L. / Dunekacke, S. / Blömeke, S. (2015). Qualitätssicherung in der Kompezentforschung.

Kattmann, U. (2007). Didaktische Rekonstruktion - eine praktische Theorie. In: Krüger, D./ Vogt, H.: Theorien in der biologiedidaktischen Forschung. Ein Handbuch für Lehramtsstudenten und Doktoranden. Springer-Verlag, Berlin, Heidelberg.

Kempin, M. (2015). Entwicklung und Evaluation eines Lernarrangements zur Förderung der Auswertungskompetenz bei nicht-linearen Zusammenhängen in Diagrammen, Universität Bremen.

Killermann, W. / Hiering, P. / Starosta, B. (2005). Biologiedidaktik heute – eine moderne Fachdidaktik, Donauwört, Auer.

Krapp, A. (1992). Das Interessenkonstrukt – Bestimmungsmerkmale der Interessenhandlung und des individuellen Interesses aus der Sicht einer Person-Gegenstands-Konzeption, Münster, Aschendorff.

Krapp, A. / Schiefele, U. / Wild, K. / Winteler, A. (1993). Der Fragebogen zum Studiuminteresse. In: Diagnostika, Universität Potsdam.

Kröll, D. (2010). Gender und MINT – Schlussfolgerungen für Unterricht, Beruf und Studium, Universität Kassel.

Kulgemeyer, C. (2010). Physikalische Kommunikationskompetenzen überprüfen, Universität Bremen.

Kulgemeyer, C. (2013). Physikalische Erklärungen, Universität Bremen.

Kulgemeyer, C. / Tomczyszyn, E. (2015). Physik Erklären – Messung der Erklärenfähigkeit angehender Physiklehrkräfte in einer simulierten Unterrichtssituation. In: Gesellschaft für Didaktik der Physik und Chemie (GDCP).

Kulgemeyer, C. (2013). Gelingensbedingungen physikalischer Erklärungen, Universität Bremen.

Kulgemeyer, C. / Schecker, H. (2009). Kommunikationskompetenz in der Physik: Zur Entwicklung eines domänespezifischen Kommunikationsbegriffs. In: Zeitschrift für Didaktik der Naturwissenschaften.

Leisen, J. (2013). Trägst du noch vor oder erklärst du schon? Der Lehrer als Erzähler oder als Erklärer. In: Naturwissenschaften im Unterricht. Physik.

Leisen, J. (2008). Anderen Physik erklären. Adressatengerecht schreiben lernen. In: Naturwissenschaften im Unterricht. Physik.

Merzyn, G. (2008). Naturwissenschaften, Mathematik, Technik – immer unbeliebter? Hohengehren: Schneider-Verlag GmbH.

Muckenfuss, H. (2006). Lernen im sinnstiftenden Kontext. Cornelsen Verlag Berlin.

Neuhaus, B. / Braun, E. (2007). Testkonstruktion und Erhebungsstrategien – praktische Tipps für empirisch arbeitende Didaktiker. In: Bayrhuber, Horst et. Al.: Kompetenzentwicklung und Assessment. Studienverlag, Innsbruck.

Ramb, B.-T. / Wübbenhorst, K. (2013): Ojektivität. In: Springer Gabler Verlag (Hrsg.), Gabler Wirtschaftslexikon.
URL: http://wirtschaftslexikon.gabler.de/Archiv/8649/objektivitaet-v8.html

Rasch, B. / Hofmann, W. / Friese, M. / Naumann, E. (2010): Quantitative Methoden, Bd. 1. Springer-Verlag, Berlin, Heidelberg.

Riese, J. / Reinhold, M. (2014). Entwicklung eines Leistungstests für fachdidaktisches Wissen. In: Krüger, D. / Parchmann, I. / Schecker, H.: Methoden in der naturwissenschaftsdidaktischen Forschung. Springer-Verlag Berlin.

Schecker, H.; Gerdes, J. (1999). Force Concept Inventory von Hestenes, D./Wells, M./Swackhammer G. Universität Bremen.

Schiefele, U. (2008). Lernmotivation und Interesse. In: Schneider, Wolfgang/ Hasselhorn, Marcus: Handbuch der Pädagogischen Psychologie. Göttingen, Hogrefe.

Schmiemann, P. / Lücken, M. (2014). Validität – Misst mein Test, was er soll? In Krüger, D./Parchmann, I./Schecker, H.: Methoden in der naturwissenschaftsdidaktischen Forschung. Springer-Verlag Berlin.

Spreckels, J. (2009). Erklären im Kontext. Neue Perspektiven aus der Gesprächs- und Unterrichtsforschung. Schneider-Verlag, Freiburg.

Töpfer, A. (2012). Erfolgreich Forschen: Ein Leitfaden für Bachelor-, Master-Studierende und Doktoranden. Springer-Verlag, Berlin, Heidelberg.

Wirth, J. (2004). Selbstregulation von Lernprozessen. Waxmann Verlag, Münster.

Wübbenhorst, K. (2013): Validität. In: Springer Gabler Verlag (Hrsg.), Gabler Wirtschaftslexikon.
URL: http://wirtschaftslexikon.gabler.de/Archiv/978/validitaet-v7.html

Wübbenhorst, K. / Maier, G.W. (2013): Reliabilität. In: Springer Gabler Verlag (Hrsg.), Gabler Wirtschaftslexikon.
URL: http://wirtschaftslexikon.gabler.de/Archiv/57492/reliabilitaet-v6.html

Anhang

Universität Bremen

_____ _____

Name in Druckbuchstaben **Matrikelnummer**

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die Abschlussarbeit selbstständig verfasst und nur die angegebenen Quellen und Hilfsmittel benutzt habe.

Wörtlich oder dem Sinn nach aus anderen Werken entnommene Stellen habe ich unter Angabe der Quellen kenntlich gemacht.

_____ _____

Datum Unterschrift

Erklärung über die Einsichtnahme in die Abschlussarbeit

Hiermit erkläre ich, dass meine Abschlussarbeit in unveränderter Fassung der Öffentlichkeit zwecks Einsichtnahme zur Verfügung gestellt wird.

☐ ja ☐ nein

_____ _____

Datum Unterschrift