

## Entwicklung eines Bewertungsmodells zur handlungsorientierten Messung experimenteller Kompetenz (Physik)Studierender

Anna B. Bauer\*, Peter Reinhold\*, Marc D. Sacher<sup>+</sup>

\*Didaktik der Physik, Universität Paderborn, <sup>+</sup> Physikalisches Grundpraktikum, Universität Paderborn  
Anna.Bauer@uni-paderborn.de, Peter.Reinhold@uni-paderborn.de, Marc.Sacher@uni-paderborn.de

### Kurzfassung

Der Erwerb experimenteller Kompetenz stellt in naturwissenschaftlichen und technischen Studiengängen eines der elementarsten Lernziele dar. Typischerweise lernen die Studierenden die Erkenntnismethodik des Experimentierens im Rahmen von Laborpraktika kennen. Aufgrund vermuteter geringer Lernwirksamkeit, die auf die Organisations- und Aufgabenstruktur zurückgeführt wird, sind vermehrt Neukonzeptionierungen entstanden. Für die Überprüfung der Lernwirksamkeit besteht nun der Bedarf an validen Textinstrumenten für die experimentelle Kompetenz (Physik)Studierender. In dem hier vorgestellten Projekt ist ein handlungsorientiertes Bewertungsmodell experimenteller Kompetenz entwickelt worden, das mit Hilfe eines Expertenratings abgesichert wird. Basierend auf dem Modell wird ein Beobachtungsbogen entwickelt, mit dem die experimentellen Fähigkeiten und Fertigkeiten erhoben werden können. Dieser soll neben einer Wirkungsanalyse unterschiedlicher Praktikumskonzepte Hinweise für eine Diagnose individueller Fähigkeiten ermöglichen und für die Gestaltung von Laborpraktika genutzt werden können.

### 1. Ausgangslage

Laborpraktika sind ein fester Bestandteil naturwissenschaftlicher und technischer Studiengänge an Universitäten und Fachhochschulen. In ihnen sollen Studierende grundlegende experimentelle Fähigkeiten und Fertigkeiten erwerben. Dies schließt auch das Diskutieren über Fachinhalte oder das Verfassen von fachwissenschaftlichen Texten ein (Empfehlung der Konferenz der Fachbereiche Physik, 2010).

Bezogen auf diese Zielsetzung wird die Lernwirksamkeit traditioneller Laborpraktika allerdings als gering eingeschätzt (Welzel & Haller, 1998). Ihre Organisations- und Aufgabenstruktur - die Studierenden erhalten eine kochrezeptartige Anleitung, die schrittweise abgearbeitet wird - fördert den Erwerb elaborierter experimenteller Fähigkeiten nur gering, da kaum Entscheidungssituationen zu bewältigen sind (Holmes, Wieman & Bonn, 2015; Tesch, 2004). Als Konsequenz sind zum einen umfangreiche Zielkataloge für die zu erlernenden Fähigkeiten mit Hilfe von Expertenbefragungen (Lehrende, Studierende) entstanden (Empfehlung der Konferenz der Fachbereiche Physik, 2010; Nagel, Scholz & Weber, 2019; Ruickoldt, 1996; Zwickl & Finkelstein, 2013). Zum anderen sind Neukonzeptionierungen von Laborpraktika auf Basis unterschiedlicher didaktischer Ansätze (didaktische Rekonstruktion (Nagel & Wolny, 2013; Neumann, 2004; Theyßen, 1999), forschendes Lernen, Cognitive Apprenticeship (Bauer, A. B. & Sacher, 2018)) entwickelt worden.

Um zu überprüfen, inwiefern Studierende in den Laborpraktika und ihren konzeptionellen Überarbeitungen die gewünschten elaborierten experimentellen

Fähigkeiten erwerben, fehlen bisher valide Testinstrumente, die eine Messung experimenteller Kompetenz auf universitärem Niveau erlauben (Nagel et al., 2019). An dieser Stelle setzt das hier vorgestellte Projekt an. In einem ersten Schritt wurde anhand von videografierten Realexperimenten unterschiedlich fähiger Probanden ein Bewertungsmodell entwickelt. Das Modell erlaubt die Analyse der Tiefenstruktur experimenteller Performanz vor dem Hintergrund des Gesamtprozesses (Bauer, A. B., Reinhold, P. & Sacher, 2020). Ausgehend von diesem Modell soll ein Beobachtungsbogen entwickelt werden, der variabel an die Anforderungen der Lernumgebung angepasst werden kann. Bewertungsmodell und Beobachtungsbogen zusammen sollen eine standardisierte Beurteilung studentischer Leistungen beim Experimentieren oder in einer handlungsorientierten Prüfungssituation ermöglichen. Darüber hinaus kann das Bewertungsmodell Hinweise für die Gestaltung von Aufgabenstellungen zur Adressierung verschiedener Fähigkeiten in Versuchsanleitungen liefern. Mit dem Modell könnten auch unterschiedliche Praktikumskonzeptionierungen verglichen werden, um Ansätze zu identifizieren, die die Ausbildung einzelner, experimenteller Fähigkeiten möglichst effizient unterstützen. Daraus könnten Hinweise zur Weiterentwicklung des Veranstaltungsformats Laborpraktikum abgeleitet werden.

### 2. Forschungsstand Diagnostik experimentelle Fähigkeiten

Modellierungen und Instrumente zur Messung experimenteller Kompetenz in den naturwissenschaftlichen Fächern existieren gegenwärtig in zahlreicher Form nur für den Bereich Schule (Emden, 2011; Gut

et al., 2014; Gut-Glanzmann, 2012; Maiseyenko, 2014; Mayer, 2007; Meier, 2016; Schecker et al., 2016; Schreiber, 2012; Schwichow & Nehring, 2018). Vereinzelt wurden diese Modelle zur Diagnose experimenteller Kompetenz auf Hochschulebene, allerdings dann für Lehramtsstudierende und dies auch nur in Form von Paper-Pencil-Tests (Straube, 2016) herangezogen. Für die Erhebung experimenteller Fähigkeiten, so wie sie Physikstudierende in Laborpraktika erwerben sollen, entstehen daraus zwei Herausforderungen.

Eine Herausforderung besteht darin, dass Handlungskompetenzen oft mit Hilfe von Paper-Pencil Tests erhoben werden. Neuweg (2011) führt an, dass sich vorliegende kognitiven Dispositionen (Wissen 2), das auf Basis von Ausbildungswissen (Wissen 1) entsteht, im kompetenten Handeln in einer Situation, dem Können (Wissen 3) manifestieren. Eine qualitative Einschätzung des Könnens kann durch die Rekonstruktion der Logik des Handelns von außen erfolgen (Neuweg, 2011; Winther, 2007). Folgt man dieser Definition ermöglichen Paper-Pencil Tests die Messung von Wissen 1 und je nach Aufgabenkonstruktion auch Wissen 2, das eigentliche Können, also die gezeigte Performanz (Vogelsang, 2014) beim Experimentieren in realen Situationen, kann mit solchen Testverfahren nicht erhoben werden. Eine Lösung stellen prozessorientierte Instrumente für die Messung experimenteller Kompetenz dar. Auf universitärem Niveau sind prozessorientierte Messungen bisher vorrangig mit Fokus auf einzelne Fähigkeiten, wie das Durchführen von Experimenten (Joußen, Fraß & Heinke, 2017) oder mit Fokus auf die Identifikation globaler Handlungsstrategien (Arndt, 2016), realisiert worden. Heidrich (2017) hat ein Modell für die prozessorientierte Messung experimenteller Kompetenz von Physikstudierenden mit Fokus auf den Gesamtprozess des Experimentierens im Inhaltsbereich Optik basierend auf Realexperimenten entwickelt. Er hat dazu drei globale Gütekriterien (Richtigkeit, Zielorientierung und Strukturiertheit) definiert. Die ersten beiden Gütemaße konnten mit Hilfe seines Instrumentes valide erhoben werden. Das Gütemaß der Strukturierung, das die Verknüpfung der einzelnen Handlungen vor dem Hintergrund des Gesamtprozesses erfasst, konnte nicht valide erhoben werden. Für die Modellierung fehlten hinreichend genaue Beschreibungen der Zielsetzungen und der zu erlernenden Kompetenzen, um Indikatoren ableiten zu können. Der Autor führt dies auf eine Diskrepanz zwischen den theoretisch modellierten experimentellen Fähigkeiten und der erhobenen Performanz zurück.

Dies stellt zusammen mit dem Vorgehen, Kompetenzmodelle für Schülerinnen und Schüler theoretisch für Universitäten zu adaptieren, die zweite Herausforderung dar. Kompetenzstrukturmodelle (Hartig & Klieme, 2007) werden auf Basis der Lernziele, die die Lernenden zu einem bestimmten Zeitpunkt erreichen sollen, entwickelt, strukturieren die Facetten der

Kompetenz und ermöglichen es somit Testinstrumente zu entwickeln, die die Passung zwischen den Kompetenzen der Lernenden und den intendierten Outcomes messbar machen. Kompetenzniveaumodelle (Hartig & Klieme, 2007) schlüsseln die einzelnen Teilfacetten nach unterschiedlichen zu erreichenden Qualitätsstufen auf und ermöglichen somit Aussagen zum Grad der Ausprägung der einzelnen Fähigkeiten und Fertigkeiten. Eine Adaption schulischer Modelle muss folglich unter der Anpassung der Modelle hinsichtlich der zu erreichenden Kompetenzen auf Basis der universitären Zielsetzungen erfolgen, um dem höheren Anspruch des Experimentierens an Universitäten gerecht zu werden (Galison, 1987; Höttecke & Rieß, 2015). Für das Studienfach Physik existiert bisher kein einheitliches, differenziertes Curriculum. Auch fehlt eine passgenaue Definition universitären Experimentierens und der zu erwerbenden Kompetenzen. Die für Laborpraktika erhobenen Zielsetzungen stellen zwar eine umfangreiche Sammlung an experimentellen Fähigkeiten und Fertigkeiten dar, sind jedoch zumeist nicht differenziert und handlungsorientiert genug beschrieben. Die Zielsetzungen umfassen oftmals nur eine Teilfacette der experimentellen Kompetenz, wie z.B. „Experimentalaufbau effektiv planen“ (Nagel et. al, 2019, S.106). Hierbei bleibt z. B. offen, welche konkreten Fähigkeiten für das Planen von Aufbauten notwendig sind und was das Wort „effektiv“ in diesem Zusammenhang bedeutet.

### 3. Fragestellungen des Projektes

Zur Entwicklung eines handlungsorientierten Bewertungsmodells experimenteller Kompetenz Physikstudierender auf universitärem Niveau werden drei Fragestellungen untersucht.

- Welche Strukturmerkmale weist die Erkenntnismethodik des Experimentierens auf universitärem Niveau in Abgrenzung zum schulischen Experimentieren auf?
- Wie können qualitative Unterschiede in der Ausprägung der experimentellen Kompetenz auf Hochschulniveau strukturiert und beschreibbar gemacht werden?
- Welches Anforderungsniveau, d.h. welchen Ausprägungsgrad der einzelnen Facetten experimenteller Kompetenz, können am Ende des Anfängerlaborpraktikums des Studiengangs Physik erwartet werden?

### 4. Methodisches Vorgehen

In diesem Kapitel wird zunächst das Forschungsdesign für die Entwicklung des handlungsorientierten Bewertungsmodells vorgestellt. Im Weiteren liegt dann der Fokus auf der Aufgabenkonstruktion und der Validierung des Modells.

#### 4.1 Forschungsdesign

In einem ersten Schritt ist als Basis für die Modellierung der experimentellen Kompetenz eine Definition für das universitäre Experimentieren entwickelt wor-

den, die der Komplexität der Erkenntnismethodik gerecht wird: Ziel des universitären Experimentierens ist das stabilisierende (Asmussen & Heering, 2014; Galison, 1987) Herauspräparieren der Eigenschaften physikalischer Phänomene (Tetens, 1987), wobei die Eigenschaften „in ein kohärentes Verhältnis zu Theorie und Praxis gebracht werden können“ (Höttecke & Rieß, 2015, S. 133).

Dafür werden vertieft vernetzte experimentelle Fähigkeiten und Fertigkeiten zusammen mit höchst flexibel einsetzbaren Wissensbeständen benötigt. Neben dem Planen, Aufbauen, Durchführen und Auswerten experimentell gewonnener Daten (Emden, 2011) besitzen auf universitärem Niveau experimentelle Fähigkeiten, wie z.B. das Testen und Optimieren experimenteller Aufbauten sowie das differenzierte Beurteilen der Handlungen und Ergebnisse vor dem Hintergrund des Gesamtprozesses, eine hohe Relevanz.

Die Entwicklung des Modells auf Basis dieser Definition erfolgt methodisch in einem mehrschrittigen Verfahren durch komparative Fallstudien ( Abb.1).

Da keine einheitlichen Zielsetzungen und Beschreibungen experimenteller Fähigkeiten auf struktureller oder Gesamtprozess-Ebene auf universitärem Niveau existieren, wird in einem ersten Schritt aus Forschungsdesiderata (Übersicht: Arndt, 2016; Emden, 2011) normativ ein Kompetenzstrukturmodell abgeleitet, das schrittweise deskriptiv mit beobachtbaren Indikatoren angereichert wird. Ziel ist, eine differenzierte Strukturierung der Handlungsschritte beim Experimentieren zu erzeugen, die den Fokus auf die Handlungsübergänge vor dem Hintergrund des Gesamtprozesses und der damit verbundenen Differenziertheit des experimentellen Vorgehens legt.

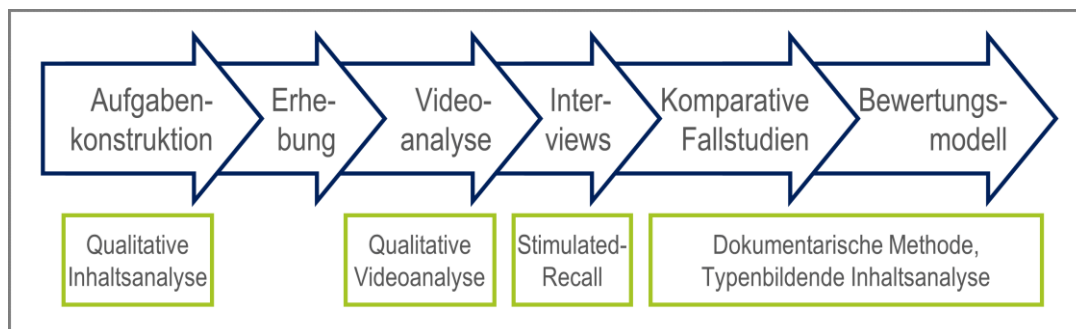
Für die deskriptive Anreicherung sind standardisierte Experimentieraufgaben entwickelt und Probanden unterschiedlicher Fähigkeitsniveaus zur Bearbeitung vorgelegt worden. Die Bearbeitung wurde videografiert und als zusätzliche Daten sind die Mitschriften der Probanden (in Form eines Laborbucheintrages), ihre Auswertungsdateien, die Browserverläufe des bereitgestellten Rechners sowie personenbezogene Informationen erhoben worden. Die Videos wurden mit Hilfe der qualitativen Videoanalyse (Bohnsack et al., 2018) analysiert und interpretiert. Im Anschluss

sind zu Validierungszwecken mit den Probanden Stimulated-Recall-Interviews (Messmer, 2015) geführt worden.

In einem vierschriftigen Verfahren sind die Daten unter Nutzung der dokumentarischen Methode (Bohnsack, 1999; Rosenberg, 2012) zu komparativen Fallstudien zusammengeführt worden. Ziel war die Rekonstruktion impliziter Wissensbestände unter Berücksichtigung der Sequentialität des Handelns (Bohnsack et al., 2013). Dieses Vorgehen erlaubt es, sowohl auf individueller Ebene als auch vergleichend zwischen den Fällen die Tiefenstruktur der gezeigten Performanz zu analysieren. Im zweiten Schritt wurden möglichst kontrastive Fälle ausgewählt (theoretical sampling (Robinson, 2014)), um das Kompetenzstrukturmodell deskriptiv anzureichern. Im Fokus der Analyse standen vor allem die Handlungssequenzen, in denen die Probanden Handlungsentscheidungen treffen, da dort die für die Entscheidungsprozesse notwendigen Fähigkeiten und Wissensbestände hinsichtlich ihrer Differenziertheit vor dem Hintergrund des Gesamtprozesses analysiert werden können. Im dritten Schritt wurden die Handlungsbeschreibungen der einzelnen Fallbeschreibungen den Facetten des Modells zugeordnet. Innerhalb der Facetten wurden die Handlungsbeschreibungen mit der typenbildenden Inhaltsanalyse (Kuckartz, 2012) analysiert, um typische Handlungsmuster zu identifizieren. Unter Anwendung des Außenkriteriums Komplexität (Commons, 2008) sowie typischer Handlungsmuster wurden dann Qualitätsstufen gebildet, die pro Facette eine Unterscheidung von Fähigkeitsniveaus ermöglichen. Der Fokus lag dabei auf Qualitätsunterschieden in der Differenziertheit der Argumentation innerhalb des Gesamtprozesses.

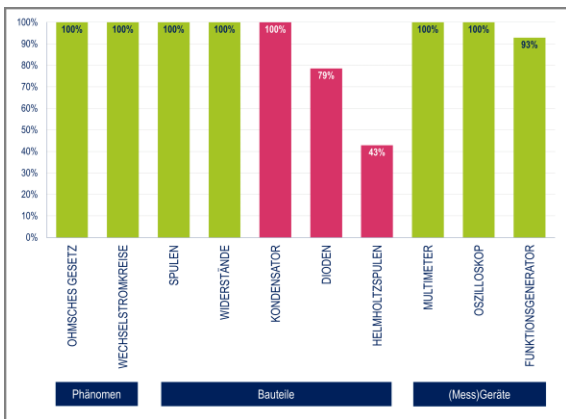
#### 4.2 Aufgabenkonstruktion

Für die Entwicklung des handlungsorientierten Bewertungsmodells wurden zwei unterschiedlich komplexe, experimentelle Aufgabenstellungen aus dem Inhaltsbereich Elektrodynamik erstellt. Die Elektrodynamik wird an allen Universitäten im Verlauf des Anfängerpraktikums mit Hilfe von Experimenten untersucht, sodass die Wahl des Themengebiets eine möglichst hohe Passung zu den Laborpraktika Deutschlands aufweist.



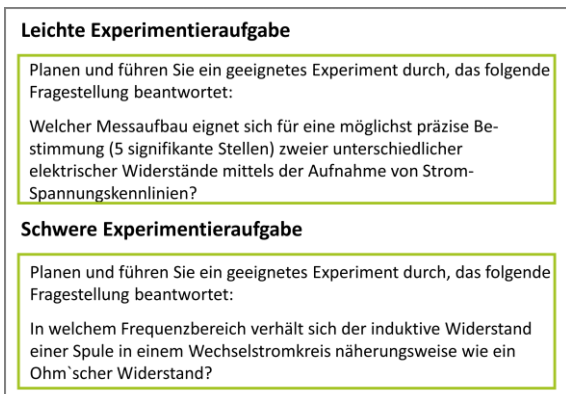
**Abb.1:** Schematische Darstellung der Bestandteile des Forschungsdesigns und des Ablaufs der Erstellung des Bewertungsmodells.

Da für Laborpraktika differenzierte Curricula fehlen, aus denen zu untersuchende experimentelle Fragestellungen und Experimentiermaterialien hätten abgeleitet werden können, wurden in einem ersten Schritt die Versuchsanleitungen aus zehn deutschen Universitäten mittels einer qualitative Inhaltsanalyse (Mayring, 2016) untersucht. Ziel war neben der curricularen Passung der Aufgabenstellungen außerdem die Sicherstellung der Vergleichbarkeit unterschiedlicher Praktikumskonzeptionen. Für die Inhaltsanalyse wurden die Aspekte Physikalisches Phänomen, Messmethode, Messgeräte, Bauteile und Materialien und Auswertungsmethode und Aufgaben zur Analyse des Vorgehens bzw. der Messergebnisse als Kategorien genutzt. Basierend auf den Ergebnissen (siehe Abb. 2) sind für die Konstruktion der Fragestellungen und Gestaltung der Experimentiermaterialien solche Aspekte gewählt worden (grün), die an (fast) allen Universitäten im Rahmen des Anfängerlaborpraktikums enthalten sind.



**Abb. 2:** Ergebnisse (Auszug) der Qualitativen Inhaltsanalyse von Versuchsanleitungen zehn deutscher Universitäten, die als Grundlage für die Aufgabenkonstruktion verwendet werden (grün).

Die unterschiedliche Komplexität der beiden Aufgabenstellungen (Abb. 3) wird dadurch erreicht, dass die leichtere Aufgabe nur Aspekte umfasst, die standardmäßig in Laborpraktika enthalten sind und somit

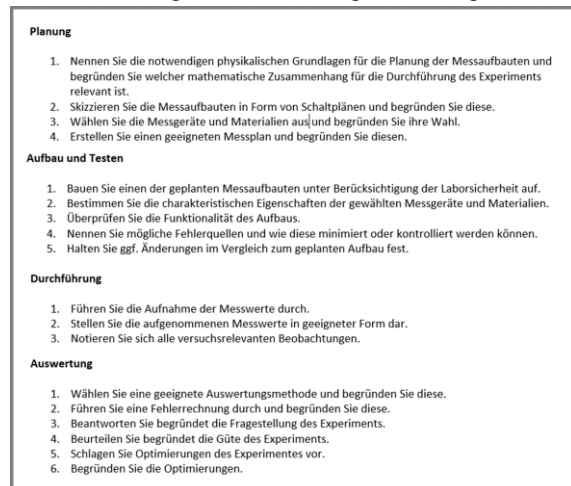


**Abb. 3:** Zu untersuchende experimentelle Fragestellungen der leichten (oben) und schwereren (unten) Aufgabe.

von den Probanden schon einmal experimentell bearbeitet wurden. Die zweite Aufgabenstellung ist schwerer gestaltet, da die Probanden zwar schon unterschiedliche Experimente mit dem Physikalischen Phänomen, mit den Bauteilen und Messgeräten oder dem Auswertungsverfahren durchgeführt haben, allerdings noch keine Fragestellung bestehend aus der Kombination aller Aspekte untersucht haben.

Bei der Konstruktion der Aufgabenstellung (Handlungsanweisungen) wurde eine einheitliche Formulierung der Aufgaben über alle Inhaltsbereiche realisiert, indem die Aufgaben entlang der Facetten des entwickelten Kompetenzstrukturmodells formuliert (Abb. 4) und lediglich die Fragestellungen der Experimentieraufgaben variiert wurden.

Um Unterschiede in der Differenziertheit der Argumentation bzw. hinsichtlich der gezeigten Performanz besser erheben zu können, wird eine Unterscheidung bei der Operationalisierung der experimentellen Fähigkeiten und Fertigkeiten aufgrund ver-



**Abb. 4:** Standardisierte Aufgabenstellung der beiden experimentellen Fragestellungen.

muteter unterschiedlicher Struktur (Gut et al., 2014) vorgenommen. Performanzaspekte, d.h. experimentelle Fähigkeiten, die sich vor allem in Handlungen, wie z.B. dem Aufnehmen von Messwerten, zeigen, sollen durch konkrete Handlungsaufforderungen in der Aufgabe ausgelöst werden. Die Fähigkeiten, die vor allem kognitive Vorgänge darstellen, wie z.B. das Planen des Experimentes, werden durch eine Aufforderung zur schriftlichen Begründung (Sander, 2000) adressiert. Weiterhin erhalten die Probanden pro Aufgabe mehrere unterschiedliche Messgeräte und Materialien. Ziel ist, dass die Probanden möglichst viele Handlungsentscheidungen während der Bearbeitung treffen können, um eine möglichst hohe Varianz in den aufgenommenen Daten zu generieren. Gerade die Stellen, in denen die Probanden selbstständig die nächsten Handlungsschritte festlegen, bieten das größtmögliche Potential, um Qualitätsunterschiede bei der Ausprägung experimenteller Fähigkeiten auf Tiefenebene zu identifizieren.

Die Aufgaben sind als open-ended investigation (Glasser, J., Gott, R., Roberts, R. & Cooper, B., 2009) konzipiert worden. Die Probanden erhielten aufgrund des standardisierten Vorgehens keine zusätzlichen Hilfestellungen durch die Testleitung und konnten also bei der Bearbeitung scheitern. Weiterhin wurde das für die Aufgabenstellung notwendige Fachwissen in Form eines Fachwissenstextes vorgegeben und die Probanden konnten während der Bearbeitung der Aufgabe einen Laptop zum Recherchieren im Internet verwenden. Dies sollte zum einen mögliche Bodeneffekte bei Probanden mit mangelndem Fachwissen verhindern und zum anderen sollte damit versucht werden, das Fachwissen als Einflussfaktor auf die Experimentierqualität konstant zu halten.

#### 4.3 Qualitätssicherung

Das Bewertungsmodell und das zu entwickelnde Testinstrument werden auf mehreren Ebenen hinsichtlich der Gütekriterien für Testverfahren überprüft.

##### 4.3.1 Stimulated-Recall Interviews

Die Rekonstruktionen (Abb. 1) der in den Videos gezeigten Handlungen werden mit Hilfe von Stimulated-Recall Interviews hinsichtlich ihrer Passung zur Wahrnehmung der Probanden abgesichert. Weiterhin konnte so an Stellen, an denen keine eindeutige Interpretation der Performanz möglich war, gezielt nachgefragt werden.

Dazu wurde theoriegeleitet ein standardisierter, vierteiliger Interviewleitfaden entwickelt. Der erste Teil umfasst Fragen zu den Einstellungen und Vorstellungen der Probanden zum Experimentieren. Die Antworten werden für die Charakterisierung der einzelnen Probanden verwendet. Im zweiten Teil sind Fragen ausgehend von der Antwort auf die jeweils untersuchte experimentelle Fragestellung enthalten. Sie folgen schrittweise den einzelnen Phasen des Experimentes von der Planung bis zur Auswertung des Experimentes. Die Struktur ist dabei für jeden Probanden gleich. Es werden allerdings individuell für jeden Probanden unterschiedliche Stimuli aus den erhobenen Produkten (Laborbucheintrag, Auswertungsdatei, Browserverlauf) verwendet. Im dritten Teil sehen sich die Probanden Ausschnitte aus ihrem videografieren Realexperiment an und kommentieren dieses. Im letzten Teil des Interviews wird auf den Fragebogen, den die Probanden vor Beginn des Experimentes ausgefüllt haben Bezug genommen. Die Probanden sollten dort ihre eigenen Fähigkeiten hinsichtlich der einzelnen in der Experimentieraufgabe enthaltenen Aspekte einschätzen. Diese Einschätzung wird den Probanden nun noch einmal vorgelegt, damit sie ggf. diese Einschätzung korrigieren können.

##### 4.3.2 Expertenrating

Das Bewertungsmodell wurde im Rahmen eines Expertenratings mit Praktikumsleitern und Lehrenden (n=30) aus dem deutschsprachigen Raum validiert (argumented-based-validity (Kane, 1992)). Dazu wurde eine Onlineumfrage erstellt, die aus drei Teilen

(Personenbezogene Daten, Einschätzung der Qualität von Handlungsbeschreibungen und Häufigkeit des Auftretens) besteht. Die Einschätzungen zur Qualität der einzelnen Handlungen werden aktuell hinsichtlich ihrer Passung zum Bewertungsmodell überprüft. Die Einschätzungen zur Auftretenshäufigkeit liefern Hinweise auf erwartbare Fähigkeiten der Studierenden am Ende des Anfängerlaborpraktikums und können hinsichtlich der Passung zu den Qualitätsstufen des Bewertungsmodells überprüft. Weiterhin sollen aus diesen Ergebnissen differenzierte, handlungsorientierte Zielsetzungen für Anfängerlaborpraktika abgeleitet werden, welche für die Gestaltung dieser Lehr-Lernumgebungen und für den Einsatz des entwickelten Instrumentes genutzt werden können.

#### 5. Stichprobe

Die Stichprobe zur Entwicklung des Bewertungsmodells wurde so gestaltet, dass eine möglichst hohe Varianz in der Performanz erreicht wird, um differenzierte Handlungsbeschreibungen generieren zu können. Dies ist vor allem vor dem Hintergrund, dass zunächst die für das Experimentieren an der Universität benötigten Fähigkeiten und Fertigkeiten sowie die strukturierenden Aspekte der Erkenntnismethodik identifiziert werden mussten, von hoher Relevanz.

Die Stichprobe (n=15) setzt sich aus 12 Studierenden nach dem dritten Semester sowie vier Experten (ein Masterand, zwei Doktoranden, ein Post-doc) zusammen. Die Studierenden erhielten Probandengelder, sodass es zu keiner Positivauswahl basierend auf dem Selbstkonzept der Probanden gekommen ist.

#### 6. Erste Ergebnisse

##### 6.1 Analyse der Strukturelemente universitären Experimentierens

Die Analyse der Strukturelemente des universitären Experimentierens, deren Ergebnisse in die Konstruktion des Bewertungsmodells eingeflossen sind, ist mit Hilfe eines Experten-Novizen-Vergleichs auf Basis der komparativen Fallstudien realisiert worden. Aufgrund des Experten-Novizen-Vergleichs anhand möglichst kontrastiver Fälle konnten drei Typen, Novizen, Fortgeschrittene und Experten identifiziert werden, die sich in drei Aspekten, die im Folgenden beschrieben werden, hinsichtlich ihrer Charakteristika bzw. Handlungsmuster deutlich unterscheiden und erste qualitative Unterscheidungen von Probanden hinsichtlich ihrer experimentellen Fähigkeiten zulassen.

##### 6.1.1. Zielorientierung

Unter dem Aspekt *Zielorientierung* wird in diesem Projekt verstanden, welchen handlungsleitenden Aspekt die Probanden beim Experimentieren besitzen. Experten arbeiten prozessorientiert und streben deswegen das möglichst stabile Herauspräparieren des physikalischen Phänomens an. Dies zeigt sich an mehreren Optimierungsschleifen während des Experimentierens, um eine möglichst präzise Messwertaufnahme sowie eine differenzierte Analyse der Ergebnisse und des Experimentes zu erreichen. Novizen



hingegen arbeiten ergebnisorientiert und zeigen sich meist schon mit dem Erhalt eines Zahlenwertes als Messergebnis, der meist unreflektiert hingenommen wird, zufrieden. Fortgeschrittene zeigen während des Experimentierens, dass sie versuchen ein möglichst präzises Messergebnis zu erreichen und durchlaufen meistens auch Optimierungsschleifen. Dies gelingt ihnen oftmals aufgrund fehlender Handlungsmuster für die Analyse und Beurteilung der einzelnen Schritte des Experimentierprozesses oder der erhaltenen Messwerte nicht effizient und präzise.

### 6.1.2. Methodischer Vernetzungsgrad

Der Aspekt *Methodischer Vernetzungsgrad* umfasst in diesem Projekt die methodische und damit argumentative Vernetzung der einzelnen Handlungen miteinander. Anhand der schriftlichen Begründungen der einzelnen Prozessschritte und Entscheidungen durch die Probanden können Aussagen zu qualitativen Unterschieden getätigt werden. Experten nutzen für die Einschätzung der Qualität ihrer Ergebnisse die Erkenntnisse aus allen Phasen des durchlaufenen Experimentierprozesses und können diese miteinander verknüpfen. Sie können gezielt Aspekte des Experimentes identifizieren, die noch Optimierungspotential besitzen und argumentieren vorrangig auf methodischer Ebene vor dem Hintergrund des Gesamtprozesses. Novizen sind meist mit den erzielten Ergebnissen zufrieden, wenn sie einen Zahlenwert erhalten, der „grob in die richtige Richtung“ geht. Sie können keine methodischen Argumente aus den einzelnen Phasen des Experimentierprozesses für die Einschätzung der Plausibilität ihres Ergebnisses generieren. Fortgeschrittene können einzelnen methodische Handlungen bei der Beurteilung der Plausibilität ihres Ergebnisses nutzen, können diese aber nicht miteinander in Beziehung setzen.

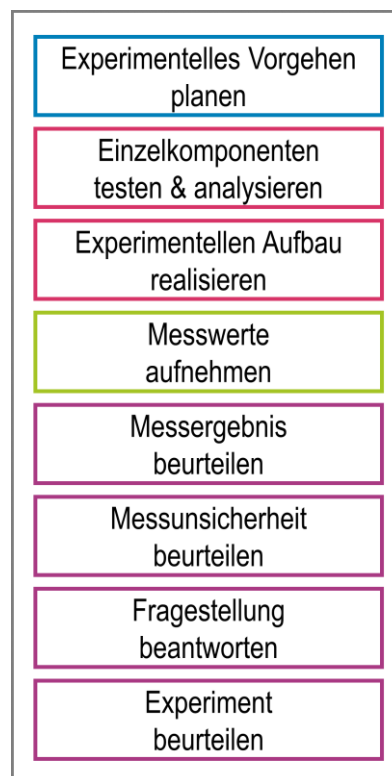
### 6.1.3. Kontrolliertes Vorgehen

Der Aspekt *Kontrolliertes Vorgehen* umfasst den Grad der Kontrolle der einzelnen Handlungsschritte beim Bearbeiten eines Experimentes. Experten überprüfen beim Experimentieren jeden Handlungsschritt hinsichtlich seiner Plausibilität vor dem Hintergrund der Fragestellung. Dies zeigt sich vor allem bei den Entscheidungssituationen hinsichtlich der Verwendung der unterschiedlichen Experimentiermaterialien. Es werden systematisch alle gegebenen Materialien getestet und hinsichtlich ihrer Eignung beurteilt sowie hinsichtlich unplausibler Messwerte oder Messergebnisse beurteilt. Novizen zeigen in Entscheidungssituationen kein kontrolliertes Vorgehen vor dem Hintergrund der Fragestellung. Dies zeigt sich daran, dass Experimentiermaterialien ohne Analyse gewählt werden, weil sie teuer aussehen oder ihnen bekannt sind oder daran, dass unplausible Messwerte ohne Kontrollmessung oder Diskussion hingenommen werden. Fortgeschrittene zeigen in Entscheidungssituationen, dass sie wissen, dass z. B. die charakteristischen Parameter von Experimentiermaterialien

Einfluss auf das Ergebnis des Experimentes besitzen, indem sie alle Materialien ansehen oder Datenblätter analysieren. Ihnen fehlen jedoch häufig die Fähigkeiten dieses Wissen handlungswirksam für z. B. Testmessungen einzusetzen.

## 6.2 Entwicklung qualitativer Abstufungen experimenteller Kompetenz

Durch die deskriptive Anreicherung des normativ abgeleiteten Kompetenzstrukturmodells konnten die Facetten der experimentellen Kompetenz weiter ausdifferenziert werden. Die Struktur und Benennung basiert auf beobachtbaren Fähigkeiten und Fertigkeiten. Das Bewertungsmodell besteht aus acht Facetten (Abb. 5), die sich jeweils noch einmal in zwei bis vier Unterfacetten auffächern. Jede Unterfacette weist vier qualitative Abstufungen auf, die mit handlungsorientierten Indikatoren aus den Fallstudien versehen sind und eine Beobachtung der jeweiligen Fähigkeit beim Experimentieren ermöglichen. Die qualitativen Abstufungen sind unter Nutzung des Außenkriteriums Komplexität sowie der aus der Analyse des universitären Experimentierens gewonnenen Erkenntnisse zu den strukturierenden Elementen konstruiert worden. Die Unterfacetten sowie deren qualitative Ausprägungen werden nach Abschluss der Auswertung des Expertenratings berichtet.



**Abb. 5:** Die beobachtbaren Kompetenzfacetten der experimentellen Kompetenz auf universitärem Niveau sind aus den komparativen Fallstudien abgeleitet worden.

Nach der Auswertung von acht der 15 Fälle zeichnet sich eine Sättigung bei der deskriptiven Anreicherung der Facetten ab. Die ausgewerteten Probanden konnten mit Hilfe des Bewertungsmodells charakterisiert

werden: Drei Probanden erreichen in den meisten Facetten die höchsten Qualitätsstufen und können als Experten deklariert werden, zwei Probanden erreichen vorrangig die niedrigsten Qualitätsstufen und können als Novizen eingestuft werden und drei Probanden erreichen die mittleren Stufen und können als Fortgeschrittene identifiziert werden.

Die Fälle, die nicht in die Anreicherung des Bewertungsmodell eingeflossen sind, zeigen keine Handlungen, die über das Bewertungsmodell hinausgehen. Sie können daher in einem weiteren Schritt für die Überprüfung des Gütekriteriums Objektivität im Rahmen eines Interratings verwendet werden.

## 7. Fazit & Ausblick

In dem vorgestellten Projekt wurde ein Bewertungsmodell für die qualitative Einschätzung der experimentellen Kompetenz (Physik)Studierender entwickelt. Das Bewertungsmodell stellt die Grundlage für das noch abzuleitende Testinstrument in Form eines Beobachtungsbogens dar, mit dem anhand der gezeigten Performanz von Studierenden beim Experimentieren Rückschlüsse auf den Grad der Ausprägung ihrer experimentellen Fähigkeiten und Fertigkeiten zu ziehen wären.

Es ist gelungen auf Basis komparativer Fallstudien ausgehend von videografierten Realexperimenten strukturierende Aspekte des universitären Experimentierens abzuleiten und diese Erkenntnisse für die standardisierte Konstruktion qualitativer Abstufungen innerhalb einzelner Fähigkeiten zu nutzen. Mit Hilfe des entwickelten Bewertungsmodells konnten die Fälle unterschiedlichen Fähigkeitsniveaus zugeordnet werden.

Das Modell wird mit Hilfe eines Expertenratings mit Praktikumsleitern aus dem deutschsprachigen Raum, das aktuell noch ausgewertet wird, validiert. Nach der Validierung soll das Modell in ein Testinstrument (Beobachtungsbogen) überführt werden und mit Hilfe eines Interratings hinsichtlich seiner Objektivität überprüft werden.

## 8. Literatur

Arndt, K. (2016). Experimentierkompetenz erfassen: Analyse von Prozessen und Mustern am Beispiel von Lehramtsstudierenden der Chemie: Dissertation. Studien zum Physik- und Chemielernen: Bd. 202. Berlin: Logos.

Asmussen, S. & Heering, P. (2014). Ein neuartiger Zugang zur Implementierung historischer Geräte und Experimente im Unterricht: Überlegungen zu Konzeption, Umsetzung und Evaluation. Physik und Didaktik in Schule und Hochschule, 13(1), 1–10.

Bauer, A. B., Reinhold, P., Sacher M. D. (2020). Bewertungsmodell zur experimentellen Performanz (Physik)Studierender. GDGP Tagungsband - Jahrestagung Wien 2019, 106–110.

Bauer, A. B., Sacher M. D. (2018). Kompetenzorientierte, universitäre Laborpraktika: Das Paderborner Physik Praktikum (3P). PhyDid B, Beiträge zur DPG-Frühjahrstagung 2018 in Würzburg, 65–72.

Bohnsack, R. (1999). Rekonstruktive Sozialforschung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Bohnsack, R., Geimer, A. & Meuser, M. (Hg.). (2018). UTB: Bd. 8226. Hauptbegriffe Qualitativer Sozialforschung (4. Aufl.). Opladen, Toronto: utb GmbH.

Bohnsack, R., Nentwig-Gesemann, I. & Nohl, A.-M. (Hg.). (2013). Die dokumentarische Methode und ihre Forschungspraxis: Grundlagen qualitativer Sozialforschung (3., aktualisierte Aufl.). Wiesbaden: Springer VS.

Commons, M. L. (2008). Introduction to the Model of Hierarchical Complexity and its Relationship to Postformal Action. World Futures, 64(5), 305–320.

Emden, M. (2011). Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens: Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I. Studien zum Physik- und Chemielernen: Bd. 118. Berlin: Logos.

Empfehlung der Konferenz der Fachbereiche Physik. (2010 in Berlin). Zur Konzeption von Bachelor- und Masterstudiengängen in der Physik. [Url: https://www.kfp-physik.de/dokument/KFP\\_Handreichung\\_Konzeption-Studiengaenge-Physik-101108.pdf](https://www.kfp-physik.de/dokument/KFP_Handreichung_Konzeption-Studiengaenge-Physik-101108.pdf) (Stand: 31.05.2020).

Galison, P. L. (1987). How experiments end. Chicago: University of Chicago Press.

Gut, C., Metzger, S., Hild, P. & Tardent, J. (Frankfurt 2014). Problemtypenbasierte Modellierung und Messung experimenteller Kompetenzen von 12- bis 15-jährigen Jugendlichen PhyDid B, Beiträge zur DPG-Frühjahrstagung 2014 in Frankfurt, 1-9.

Gut-Glanzmann, C. (2012). Modellierung und Messung experimenteller Kompetenz: Analyse eines large-scale Experimentiertests: Univ., Diss.-Basel, 2012. Studien zum Physik- und Chemielernen: Bd. 134. Berlin: Logos.

Hartig, J. & Klieme, E. (2007).: Empirische Erfassung von Kompetenzen und psychometrische Kompetenzmodelle. In J. Hartig & E. Klieme (Hg.), Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik.: Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung, 17–36.

Holmes, N. G., Wieman, C. & Bonn, D. A. (2015). Teaching critical thinking. Proceedings of the National Academy of Sciences of the United States of America, 112(36), 11199–11204.

Höttecke, D. & Rieß, F. (2015). Naturwissenschaftliches Experimentieren im Lichte der jüngeren Wissenschaftsforschung - Auf der Suche nach einem authentischen Experimentbegriff der Fachdidaktik: Zeitschrift für Didaktik der Naturwissenschaften, 21(1), 127–139.

Joußen, N., Fraß, S., Heinke, H. (2017). Diagnostik experimenteller Strategien: Validierung eines prozessorientierten Instruments. PhyDid B, Beiträge zur DPG-Frühjahrstagung 2017 in Dresden, 131-137.

- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112(3), 527–535.
- Kuckartz, U. (2012). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung*. Weinheim u.a.: Beltz Juventa.
- Maiseyenko, V. (2014). Modellbasiertes Experimentieren im Unterricht: Praxistauglichkeit und Lernwirkungen. *Studien zum Physik- und Chemielernen*: Bd. 166. Berlin: Logos.
- Mayer J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger; H. Vogt (Hg.), *Theorien der biologiedidaktischen Forschung* (S. 177–186). Berlin: Springer.
- Mayring, P. (2016). *Einführung in die qualitative Sozialforschung: Eine Anleitung zu qualitativem Denken* (6. Aufl.). Weinheim, Basel: Pädagogik. Beltz.
- Meier, M. (2016). Entwicklung und Prüfung eines Instrumentes zur Diagnose der Experimentierkompetenz von Schülerinnen und Schülern. *BIOLOGIE lernen und lehren*: Bd. 13. Berlin: Logos.
- Messmer, R. (2015). Stimulated Recall als fokussierter Zugang zu Handlungs- und Denkprozessen von Lehrpersonen. *FQS*, 16(1), Art. 3.
- Nagel, C., Scholz, R. & Weber, K. (2019). Umfrage zu Lehr/Lernzielen in physikalischen Praktika. *Phy-Did B - Didaktik der Physik- DPG-Schule Physikalische Praktika*, Bad Honnef, 97–109.
- Nagel, C. & Wolny, B. (2013). Ein adressatenspezifisches Physikpraktikum für Ernährungswissenschaften: Didaktische Rekonstruktion und Evaluation. *Physik und Didaktik in Schule und Hochschule*, 12(1), 48–61.
- Neumann, K. (2004). Didaktische Rekonstruktion eines physikalischen Praktikums für Physiker. *Studien zum Physik- und Chemielernen*: Bd. 38. Berlin: Logos.
- Neuweg, G. H. (2011). Das Wissen der Wissensvermittler. In Terhart, E., Bennewitz, H., Rothland, M. (Hg.), *Handbuch der Forschung zum Lehrerberuf* (S. 451–477). Münster: Waxmann.
- Robinson, O. C. (2014). Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide. *Qualitative Research in Psychology*. *Qualitative Research in Psychology*, 11, 25–41.
- Rosenberg, F. von. (2012). Rekonstruktion biographischer (Bildungs-)Prozesse: Überlegungen zu einer prozessanalytischen Typenbildung. In I. Miethe & H.-R. Müller (Hg.), *Qualitative Bildungsforschung und Bildungstheorie* (S. 193–207). Opladen: Budrich.
- Ruickoldt, G. (1996). Ergebnisse einer Umfrage zum Physikalischen Praktikum. *Physikalische Blätter* 52(10), 1022–1024.
- Sander, F. (2000). Verbindung von Theorie und Experiment im physikalischen Praktikum: Eine empirische Untersuchung zum handlungsbezogenen Vorverständnis und dem Einsatz grafikorientierter Modellbildung im Praktikum. *Studien zum Physiklernen*: Bd. 13. Berlin: Logos.
- Schecker, H., Neumann, K., Theyßen, H., Eickhorst, B. & Dickmann, M. (2016). Stufen experimenteller Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 22(1), 197–213.
- Schreiber, N. (2012). Diagnostik experimenteller Kompetenz: Validierung technologiegestützter Testverfahren im Rahmen eines Kompetenzstrukturmodells. *Studien zum Physik- und Chemielernen*: Bd. 139. Berlin: Logos.
- Schwichow, M. & Nehring, A. (2018). Variablenkontrolle beim Experimentieren in Biologie, Chemie und Physik: Höhere Kompetenzausprägungen bei der Anwendung der Variablenkontrollstrategie durch höheres Fachwissen? *Empirische Belege aus zwei Studien*. *Zeitschrift für Didaktik der Naturwissenschaften*, 24(1), 217–233.
- Straube, P. (2016). Modellierung und Erfassung von Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung bei (Lehramts-) Studierenden im Fach Physik. *Studien zum Physik- und Chemielernen*: Bd. 209. Berlin: Logos.
- Tesch, M. (2004). Experimentieren im Physikunterricht: Ergebnisse einer Videostudie. *Zeitschrift für Didaktik der Naturwissenschaften* (10), 51–69.
- Tetens, H. (1987). Experimentelle Erfahrung: Eine wissenschaftstheoretische Studie über die Rolle des Experiments in der Begriffs- und Theoriebildung der Physik. *Paradeigmata*: Bd. 8. Hamburg: Meiner.
- Theyßen, H. (1999). Ein Physikpraktikum für Studierende der Medizin: Darstellung der Entwicklung und Evaluation eines adressatenspezifischen Praktikums nach dem Modell der Didaktischen Rekonstruktion. *Studien zum Physik- und Chemielernen*: Bd. 9. Berlin: Logos.
- Vogelsang, C. (2014). Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften: Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz. *Studien zum Physik- und Chemielernen*: Bd. 174. Berlin: Logos.
- Welzel, M. & Haller, K. e. a. (1998). Ziele, die Lehrende mit dem Experimentieren in der naturwissenschaftlichen Ausbildung verbinden: Ergebnisse einer europäischen Umfrage. *Zeitschrift für Didaktik der Naturwissenschaft*, 4(1), 29–44.
- Winther, E. (2007). Performanz messen - Kompetenz diagnostizieren. In D. Lemmermöhle, M. Rothgangel, S. Bögeholz, M. Hasselhorn & R. Watermann (Hg.), *Professionell lehren, erfolgreich lernen* (S. 303–316). Münster: Waxmann.
- Zwickl, B. M. & Finkelstein, Noah, H.J. (2013). The process of transforming an advanced lab course: Goals, curriculum and assessment. *American Journal of Physics* (81), 63–70.