

## Entwicklung und Validierung eines Testinstruments zu einfachen Stromkreisen

Benjamin Groß\*, Lana Ivanjek<sup>†</sup>, Salome Flegr<sup>°</sup>, Judith Glaesser<sup>x</sup>, Augustin Kelava<sup>x</sup>,  
Jan-Philipp Burde\*

\*Eberhard Karls Universität Tübingen, AG Didaktik der Physik, Auf der Morgenstelle 14, 72076 Tübingen

<sup>†</sup>Johannes Kepler Universität Linz, Abteilung für MINT-Didaktik, Altenberger Straße 68, 4040 Linz

<sup>°</sup>Technische Universität Dresden, Professur für Didaktik der Physik, Haeckelstraße 3, 01069 Dresden

<sup>x</sup>Eberhard Karls Universität Tübingen, Methodenzentrum, Haußerstraße 11, 72076 Tübingen

[benjamin.gross@uni-tuebingen.de](mailto:benjamin.gross@uni-tuebingen.de)

### Kurzfassung

Zur Erfassung typischer Lernendenvorstellungen und des konzeptionellen Verständnisses zu einfachen elektrischen Stromkreisen wurde ein dreistufiges Multiple-Choice-Testinstrument entwickelt, d. h. die Items erfassen neben einer Antwort auf eine Frage auch die zugehörige Begründung und die subjektive Sicherheit der Teilnehmenden bei der Beantwortung der Frage. Die in diesem Artikel beschriebene Validierung basiert zunächst auf der zweistufigen Auswertung der Kombinationen aus Antwort und Begründung. Die 18 Testitems wurden auf Basis einer Literaturrecherche konstruiert und decken sechs verbreitete Vorstellungen zu den Konzepten Strom, Spannung, Widerstand, Energiefluss in Stromkreisen sowie deren Systemcharakter ab. Durch eine Expertenbefragung wurde die inhaltliche Validität der Testitems überprüft, die quantitative Evaluation erfolgte in einer Online-Befragung mit  $N = 160$  Studierenden naturwissenschaftlicher Studiengänge im ersten Studienjahr. Die Items zeigten angemessene Schwierigkeiten, hohe Trennschärfen und eine sehr gute interne Konsistenz ( $\alpha = 0.90$ ). Konfirmatorische Faktorenanalysen mit Bifaktormodellen ergaben gute Modellfit-Indizes und ausreichend hohe Subskalen-Reliabilitätswerte mit Blick auf Lernendenvorstellungen, was die Diagnosefähigkeit des Testinstruments unterstützt.

### 1. Einleitung

Einfache elektrische Stromkreise sind ein zentraler Bestandteil des Physikunterrichts in der Sekundarstufe I und stellen zugleich eine wichtige Grundlage für das Verständnis moderner technischer Anwendungen dar. Trotz der intensiven schulischen Behandlung dieses Themenfelds zeigen zahlreiche Studien, dass selbst nach dem Unterricht oftmals Verständnisprobleme bezüglich grundlegender Konzepte wie Strom, Spannung, Widerstand sowie der Funktionsweise elektrischer Stromkreise bestehen – und dies nicht nur bei Schülerinnen und Schülern, sondern auch bei Studierenden der Physik und angehenden Lehrkräften [1–5]. Diese konzeptionellen Schwierigkeiten äußern sich in stabilen, weit verbreiteten und gut dokumentierten Lernendenvorstellungen, die mit wissenschaftlichen Konzepten nicht vereinbar sind [6].

In den letzten Jahrzehnten wurden verschiedene Testinstrumente entwickelt, um das konzeptuelle Verständnis von Lernenden zu elektrischen Stromkreisen zu erfassen [7–10]. Diese Instrumente decken jedoch nur einen begrenzten Bereich von Konzepten ab, weisen eine unzureichende Reliabilität auf oder ermöglichen keine Erfassung von Lernendenvorstellungen. Einige von ihnen weisen zudem grundlegende Schwächen in der Itemkonstruktion auf. Um die Lernendenvorstellungen, das Verständnis der zugehörigen spezifischen physikalischen Konzepte sowie das

allgemeine Konzeptverständnis bezüglich einfacher elektrischer Stromkreise zu erfassen, wurde ein Multiple-Choice-Testinstrument entwickelt. Dieses wurde im Rahmen der hier beschriebenen Studie mit Studierenden im ersten Jahr eines naturwissenschaftlichen oder lehramtsbezogenen Studiums evaluiert, ließe sich aber (ggf. unter Einschränkung auf bestimmte Subskalen) auch im schulischen Kontext einsetzen.

### 2. Theoretischer Hintergrund

#### 2.1. Typische Lernendenvorstellungen zu Stromkreisen

Auf Basis einer umfassenden Literaturrecherche wurden zunächst sieben Lernendenvorstellungen ausgewählt, die als empirisch gut belegt gelten und einen direkten Bezug zu fundamentalen Konzepten bezüglich elektrischer Stromkreise aufweisen:

- Stromverbrauch (SV, Konzept „elektrischer Strom“): Strom wird als „verbrauchbare“ Substanz verstanden, die durch Bauteile wie Lampen oder Widerstände verbraucht wird – insbesondere in Form eines Abschwächungsmodells (engl. „attenuation model“) [6].
- Batterie als konstante Stromquelle (BKS, Konzept „Batterie“): Die Batterie wird als Stromquelle mit konstanter Stromstärke interpretiert,

obwohl sie physikalisch gesehen (idealisiert) eine konstante Spannung liefert ([3], [4], [11]).

- Vermischung von Spannungs- und Strombegriff (UI, Konzept „Spannung“): Die Spannung wird nicht als eigenständige physikalische Größe erkannt, sondern fälschlicherweise als eine Eigenschaft des elektrischen Stroms verstanden ([7], [12–15]).
- Sequenzielle Argumentation (SEQ, Konzept „Systemcharakter von Stromkreisen“): Der Stromkreis wird als linear-sequenzieller Prozess interpretiert, bei dem der „Strom“ (als Substanz) nacheinander die Schaltungselemente passiert. Daraus wird gefolgert, dass die Reihenfolge der Bauteile entscheidend sei ([3], [6], [16]).
- Lokales Denken (LOK, Konzept „Systemcharakter von Stromkreisen“): Statt den Stromkreis als zusammenhängendes System zu betrachten, wird er auf Basis lokaler Gegebenheiten (z.B. Verzweigungen) analysiert. So wird z.B. davon ausgegangen, dass sich der elektrische Strom an Knotenpunkten gleichmäßig aufteilt ([7], [17]).
- Vernachlässigung der Schaltungsart bei Widerständen (EW, Konzept „Widerstand“): Der Unterschied zwischen Reihen- und Parallelschaltung wird ignoriert; die Anzahl der Widerstände wird als allein entscheidend für den Ersatzwiderstand betrachtet [4].
- Vermischung von Strom- und Energiebegriff (EN, Konzept „Energie“): Energie wird sprachlich wie konzeptionell mit dem Strombegriff gleichgesetzt, was häufig zu falschen Schlüssen über Energieflüsse im Stromkreis führt [6].

## 2.2. Bisherige Testinstrumente zu elektrischen Gleichstromkreisen

Im Rahmen der Testentwicklung wurde zudem eine Analyse bestehender Testinstrumente zu Gleichstromkreisen durchgeführt. Ziel war es, zentrale Eigenschaften hinsichtlich Zielgruppe, Konzeptabdeckung, Item-Design und psychometrischer Qualität zu untersuchen. Berücksichtigt wurden nur Tests, die sich ausschließlich mit einfachen Gleichstromkreisen befassen. Instrumente mit breiterem thematischen Fokus – etwa auf Elektrostatik, Wechselstrom oder Elektrodynamik – wurden nicht einbezogen, selbst wenn sie teilweise auch Items zu Gleichstromkreisen enthalten (z. B. [18–20]).

Der Test von Rhöneck [7] war eines der ersten Instrumente im deutschsprachigen Raum, das typische Lernendenvorstellungen zu Strom, Spannung und Widerstand thematisierte. Der Test umfasst 12 Items verschiedener Formate (u. a. Multiple Choice, offene Aufgaben), wobei pro Vorstellung meist nur ein oder zwei Items eingesetzt wurden. Auch wenn eine systematische Validierung im Sinne aktueller psychometrischer Standards nicht erfolgte, erlangte der Rhöneck-Test erhebliche Bedeutung.

Ein ausgereifteres Instrument stellt der 2004 entwickelte, englischsprachige DIRECT-Test dar [8]. Dieser Test umfasst 29 einstufige Items zu Strom, Spannung, Widerstand und Energieübertragung. Er wurde mit einer großen Stichprobe von Schülerinnen und Schülern und Studierenden validiert und weist akzeptable Reliabilitätswerte auf. Die Items wurden jedoch nicht direkt auf Basis bestimmter Lernendenvorstellungen konstruiert. Vielmehr wurden diese post-hoc durch Interviews identifiziert. Auch aufgrund seiner Einstufigkeit eignet sich der DIRECT-Test eher zur Erhebung des allgemeinen Konzeptverständnisses als zur gezielten Diagnostik einzelner Vorstellungen.

Der SECDT-Test [21], ursprünglich in türkischer Sprache entwickelt, verwendet ein dreistufiges Item-Format. Neben Antwort und Begründung geben die Teilnehmenden auch ihre Antwortsicherheit an. Die Autoren sehen darin wie Hasan et al. [22] ein Mittel, zwischen fehlendem Wissen und stabilen Vorstellungen zu unterscheiden, was in Anbetracht neuerer empirischer Befunde aus anderen Inhaltsgebieten der Physik jedoch kritisch gesehen werden kann [23]. Inhaltlich fokussiert der Test auf den Strombegriff, während andere Konzepte wie Spannung oder Energie kaum berücksichtigt werden. Zudem wurde die Reliabilität von den Autoren als niedrig eingeschätzt, eine weiterführende Validierung (z. B. mittels Faktorenanalyse) wurde nicht durchgeführt.

Mit dem AECCP-Test [9] liegt ein deutschsprachiges Instrument mit 23 Items vor, das teilweise auch zweistufige Fragen verwendet. Der Schwerpunkt liegt auf den Konzepten Strom und Widerstand, der Spannungs- und Energiebegriff werden hingegen nicht abgedeckt. Jede Lernendenvorstellung wird durch mindestens drei Items operationalisiert. Die Ergebnisse einer Faktorenanalyse zeigen, dass einzelne Vorstellungen in diesem Format zuverlässig identifiziert werden können. Der Test weist eine gute Gesamtreliabilität auf [9].

Der 2T-SEC-Test [10] besteht aus 25 durchgängig zweistufigen Items mit je einer Antwort- und Begründungsstufe und deckt dabei auch den Spannungsbegriff ab. Die Auswertung mittels Rasch-Modell ergab gute Item-Fit-Werte sowie eine hohe Item-Reliabilität. Die Personenreliabilität war hingegen nur marginal akzeptabel, und die Schwierigkeit des Tests erwies sich für die Zielgruppe (Sekundarstufe I) als zu hoch. Zudem lag der Fokus der Entwicklung auf physikalischen Konzepten, nicht explizit auf Lernendenvorstellungen. Eine Folge davon ist, dass die durch den Test abgedeckten Lernendenvorstellungen mit unterschiedlich vielen Items erfasst werden und keine Faktorenanalysen zur Validierung der Lernendenvorstellungen durchgeführt wurden.

## 3. Ziele und Testentwicklungsprozess

Das hier beschriebene Testinstrument wurde im Rahmen eines gemeinsamen Forschungsprojekts dreier Universitäten in Deutschland und Österreich entwickelt.

### 3.1. Zielsetzung

Die Entwicklung des Testinstruments erfolgte auf der Grundlage der zuvor durchgeführten Analyse existierender Testinstrumente (vgl. Abschnitt 2.2). Das Instrument sollte dabei zentrale Konzepte wie Strom, Spannung, Widerstand, Batterie, Systemcharakter und Energie sowie zugehörige, verbreitete Lernendenvorstellungen ausgewogen abdecken. Im Rahmen der hier vorgestellten statistischen Evaluation sollte die Validität des Testinstruments untersucht werden – speziell inwieweit die Erfassung

- des übergreifenden Konzeptverständnisses,
- einzelner typischer Lernendenvorstellungen und
- des Verständnisses einzelner physikalischer Konzepte (z.B. Strom, Spannung und Widerstand)

möglich ist. Ziel war zunächst die Konstruktion zweistufiger Items, d. h. insbesondere von passenden Fragestellungen inklusive geeigneter Antwort- und Begründungsdistraktoren. Um weitere Fragestellungen bezüglich der Stabilität von Lernendenvorstellungen untersuchen zu können, wurden diese Items zudem um eine dritte Stufe erweitert, welche in einer sechststufigen Likert-Skala auch die von den Teilnehmenden subjektiv wahrgenommene Sicherheit bei der Beantwortung der Frage erfasst. Im Folgenden wird hierauf jedoch nicht weiter eingegangen, da dieser Artikel sich mit der Erfassung des konzeptionellen Verständnisses mittels der ersten beiden Stufen des Testinstruments befasst.

### 3.2. Entwicklungsschritte

Die initiale Testentwicklung verlief folgendermaßen:

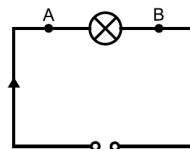
- Identifikation relevanter Vorstellungen: Basierend auf einer Literaturrecherche wurden sieben weit verbreitete Lernendenvorstellungen zu den Kernkonzepten bezüglich einfacher Stromkreise bestimmt (siehe Abschnitt 2.2).
- Sichtung bestehender Instrumente und Item-adaption: Aus bestehenden Testinstrumenten wurden Items ausgewählt, die zur Messung der definierten Vorstellungen, ggf. in adaptierter Form, geeignet erschienen. Im Falle der Adaption einstufiger Items wurde eine zweite Stufe in Form geeigneter Begründungsdistraktoren ergänzt. Wo dies nicht möglich war, wurden Items neu entwickelt.
- Itementwicklung: Für jede der sieben Vorstellungen wurden vier zweistufige Single-Choice-Items konstruiert, was zu einer vorläufigen Testversion mit 28 Items führte. Jedes Item enthält eine Schaltskizze, einen Beschreibungstext und eine fachliche Frage mit mehreren Antworten (Stufe 1) sowie passenden Begründungen (Stufe 2). Bei jedem Item ist aus physikalischer Sicht genau eine Kombination von Antwort und Begründung korrekt. Zusätzlich wurden jeweils meist eine, teils auch mehrere spezifische Distraktoren-Kombinationen (d.h. Antwort und Begründung)

konstruiert, die als Indikator für eine bestimmte Lernendenvorstellung dienen.

Besonderes Augenmerk wurde allgemein auf die Vergleichbarkeit innerhalb von Item-Gruppen zu einer bestimmten Vorstellung gelegt, indem z.B. nicht nur ähnliche Stromkreise betrachtet werden, sondern wo möglich auch auf eine Vergleichbarkeit der Frage- und Distraktoren-Formulierungen geachtet wurde. Ferner wurde auch die Anzahl der Distraktoren auf beiden Stufen innerhalb jeder Item-Gruppe konstant gehalten, um statistische Verzerrungen zu vermeiden. Alle Items des Testinstruments wurden so konstruiert, dass eine Beantwortung der Frage inklusive zugehöriger Begründung auf rein konzeptueller Ebene erfolgen kann und keinerlei Berechnungen nötig sind. Zum Test gehört ebenfalls eine Instruktion, in der den Teilnehmenden die verwendeten Schaltsymbole erläutert sowie allgemeine Hinweise und Anweisungen zur Testdurchführung gegeben werden. Ein Beispielimitem ist in Abb. 1 dargestellt.

Auf Basis der statistischen Evaluation und dem Experten-Feedback wurden je Vorstellungskategorie aus den ursprünglich vier Items jeweils drei in die finale Testversion übernommen. Da die Items zur Vorstellung „Lokales Denken“ entfernt werden mussten (siehe Abschnitt 4), umfasst die finale Testversion 18 Items, mit denen sechs Lernendenvorstellungen erfasst werden können.

**Im abgebildeten Stromkreis ist eine Glühlampe mit einer Batterie verbunden und leuchtet.**



Was lässt sich über die Stromstärke am Punkt B im Vergleich zur Stromstärke am Punkt A aussagen?

- ☐ Die Stromstärke am Punkt B ist Null.
- ☐ Die Stromstärke am Punkt B ist kleiner als am Punkt A, aber nicht Null.
- ☐ Die Stromstärke am Punkt B ist gleich groß wie am Punkt A.

**Wie haben Sie Ihre obige Antwort begründet?**

Wichtig: Wählen Sie diejenige Begründung aus, die am ehesten Ihrem eigenen Gedankengang bei der Beantwortung der Frage entsprach.

- ☐ Die leuchtende Glühlampe verbraucht den kompletten Strom.
- ☐ Die leuchtende Glühlampe verbraucht einen Teil des Stroms.
- ☐ Durch das Leuchten der Glühlampe wird kein Strom verbraucht.

**Abb. 1:** Beispielimitem aus dem Testinstrument zur Vorstellung „Stromverbrauch“ (adaptiert nach einem Item aus dem 2T-SEC-Test [10])

### 4. Evaluation

In diesem Abschnitt wird die grundsätzliche Vorgehensweise bei der Evaluation des Testinstruments sowie eine Auswahl der wesentlichen Ergebnisse vorgestellt. Für eine ausführlichere Darstellung sei auf die englischsprachige Publikation zum Testinstrument verwiesen, welche sich aktuell in Vorbereitung befindet.

#### 4.1. Expertenbefragung

Im Anschluss an die initiale Item-Entwicklung wurde eine Expertenbefragung durchgeführt, um die inhaltliche Validität der entwickelten Items zu überprüfen. Fünf Experten der Physikdidaktik mit aktiver Forschungstätigkeit sowie Lehrerfahrung auf Schul- oder Hochschulebene nahmen an der Befragung teil. Dabei war keiner der Experten an der Entwicklung der Items direkt beteiligt.

Die Befragung erfolgte anhand eines Dokuments mit sämtlichen 28 zweistufigen Items sowie den Instruktionen für Testteilnehmende. Die Experten erhielten eine Liste, in der für jedes Item die als korrekt intendierte Kombination aus Antwort und Begründung angegeben war. Zudem wurden jene Kombinationen ausgewiesen, die aus Sicht der Entwickler spezifischen Lernendenvorstellungen zugeordnet werden sollten, zusammen mit Definitionen der zu erhebenden Vorstellungen.

Die Aufgabe der Experten bestand darin,

- die jeweils angegebene Kodierung zu bewerten (Skala: volle Zustimmung / Zustimmung mit Einschränkungen / Ablehnung);
- eine kurze Begründung anzugeben, sofern einer Kodierung nicht vollständig zugestimmt wurde;
- ggf. Kommentare zu Items, Begründungen oder Formulierungen abzugeben.

Die Rückmeldungen belegten eine hohe Zustimmung zu den angegebenen Kodierungen: Alle Kombinationen von Antwort und Begründung, die aus physikalischer Sicht korrekt sind, wurden zumeist einstimmig oder mit großer Mehrheit von den Experten akzeptiert. Auch die angegebenen Zuordnungen bestimmter Kombinationen von Antwort und Begründung zu Lernendenvorstellungen wurden mehrheitlich von den Experten bestätigt. Einigen dieser Antwortkombinationen wurden von einzelnen Experten jedoch nur mit Einschränkungen zugestimmt – insbesondere dann, wenn die sprachliche Formulierung eine alternative Interpretation zuließ oder wenn sich Distraktoren nicht eindeutig von anderen Vorstellungen abgrenzen ließen. In solchen Fällen wurden entweder die betreffenden Zuordnungen zu Lernendenvorstellungen entfernt oder die Formulierungen überarbeitet. Die Rückmeldungen führten darüber hinaus zur generellen Verbesserung der Distraktoren hinsichtlich deren allgemeiner Verständlichkeit sowie der Konsistenz von Formulierungen innerhalb der zu den Lernendenvorstellung gehörenden Item-Gruppen. Insgesamt stützen die Ergebnisse der Expertenbefragung nicht nur die inhaltliche Validität des Instruments, sondern dienen darüber hinaus als Grundlage für die abschließende Kodierung der Lernendenvorstellungen und (neben statistischen Kriterien) auch für die Itemauswahl in der finalen Testversion.

#### 4.2. Quantitative Evaluation

##### 4.2.1. Durchführung und Stichprobe

Das Testinstrument wurde in den Sommersemestern 2023 und 2024 an sechs Universitäten im deutschsprachigen Raum online pilotiert. Teilnehmende waren Studierende der Physik sowie naturwissenschaftlicher Fächer mit Physikbezug sowie Lehramtsstudierende im ersten Studienjahr. Die Testdurchführung fand dabei vor der ersten Behandlung der Elektrizitätslehre im Rahmen universitärer Lehrveranstaltungen statt.

Die Teilnahme erfolgte freiwillig, im Rahmen von Lehrveranstaltungen oder individuell von zu Hause aus per Link zum Fragebogen über hochschulinterne Verteiler. Insgesamt wurden 176 Datensätze erhoben, von denen 13 aufgrund unvollständiger Bearbeitung oder sehr kurzer Bearbeitungszeiten ausgeschlossen wurden [24]. Die finale Stichprobe umfasste  $N = 160$  Personen. Die mittlere Bearbeitungszeit betrug 24 Minuten (Median).

Die Testadministration erfolgte online über die Plattform SoSci Survey. Die Items wurden einzeln präsentiert, wobei die zweite Stufe (Begründung) erst nach Beantwortung der ersten Stufe (Antwort) sichtbar wurde. Eine Änderung der Antwort auf der ersten Stufe war dann nicht mehr möglich. Diese sequenzielle Darstellung diente dazu, unerwünschte Antwortstrategien auszuschließen, die durch parallele Sichtbarkeit beider Stufen bei papierbasierten Tests begünstigt werden können (vgl. [25]). Eine Rückkehr zu vorherigen Fragen war ebenfalls nicht möglich.

##### 4.2.2. Methoden

Zur quantitativen Evaluation des Testinstruments wurden verschiedene Verfahren der klassischen Testtheorie sowie explorative und konfirmatorische Faktorenanalysen eingesetzt. Die verwendeten Methoden und deren Bewertungskriterien orientieren sich an etablierten Empfehlungen zur Prüfung der Güte von Konzepttests, z. B. nach Jorion et al. [26]. Nachfolgend sind die eingesetzten Verfahren und die jeweils zugrunde gelegten Gütekriterien aufgeführt.

- Itemschwierigkeiten: Anteil der Teilnehmenden, die die korrekte Kombination aus Antwort und Begründung wählen
- Item-Trennschärfen: Korrelation zwischen Item-Scores und Gesamtttestscore
- Reliabilität (Cronbach's Alpha): interne Konsistenz über alle Items hinweg
- Distraktorenanalyse: Anteil der Teilnehmenden, die eine spezifisch kodierte Lernendenvorstellung wählen
- Explorative Faktorenanalyse (EFA): Parallelanalyse zur Ermittlung der Anzahl latenter Faktoren; Interpretation der Faktorenlösungen bzgl. erwarteter Zuordnung zu Items

- Konfirmatorische Faktorenanalyse (CFA): Bifaktormodelle; Modellfit, Faktorladungen und zugehörige Reliabilitätskennwerte (Gesamtreliabilität  $\omega$ , hierarchische Reliabilitätswerte  $\omega_h$  des Generalfaktors und der Subskalen) [27]

#### 4.2.3. Ergebnisse

Die im Folgenden berichteten Ergebnisse beziehen sich auf die finale Testversion, d. h. es wurden bei der statistischen Evaluation nur die Antworten auf die darin enthaltenen 18 Items berücksichtigt. Die Items zum Lokalen Denken wurden auf Basis der Strukturanalyse entfernt (siehe auch Abschnitt 4.2.3.2).

##### 4.2.3.1. Distraktorenanalyse und deskriptivstatistische Kennwerte

Die Distraktoren der zu den Lernendenvorstellungen gehörigen Antwortkombinationen wurden je nach Item von 16 % bis 51 % der Teilnehmenden gewählt und sind daher ausreichend attraktiv. Die Items weisen mit 13 % bis 63 % korrekten Antworten eine breite Streuung der Schwierigkeit auf, was eine gute Differenzierung zwischen einfachen und schwierigen Items darstellt und mit wenigen Ausnahmen im empfohlenen Bereich liegt [26]. Die Trennschärfen der Items liegen im Bereich von 0.38 bis 0.69 und sind damit durchgehend als sehr gut zu bewerten [26].

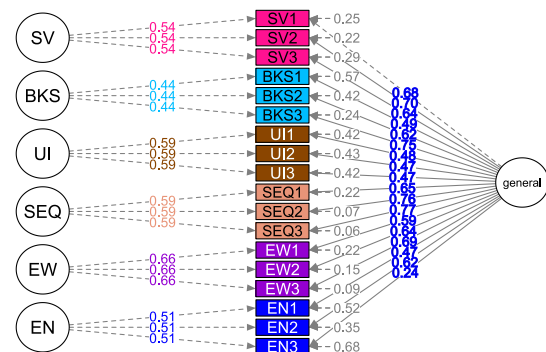
Die interne Konsistenz der finalen 18-Item-Version wurde mittels Cronbachs Alpha bestimmt ( $\alpha = 0.90$ ), was auf eine hohe Reliabilität des Instruments hinweist. Aufgrund der erwarteten Multidimensionalität des Tests muss dieser Wert jedoch mit Vorsicht interpretiert werden.

##### 4.2.3.2. Strukturanalysen

Bei der Strukturanalyse basierend auf Lernendenvorstellungen wurde eine dichotome Kodierung gewählt, abhängig davon, ob eine gewählte Kombination aus Antwort (Stufe 1) und Begründung (Stufe 2) die mit dem jeweiligen Item erfasste Lernendenvorstellung repräsentiert („1“) oder nicht („0“). Die Parallelanalyse mit anschließender explorativer Faktorenanalyse mit obliquen Rotation (korrelierte Faktoren) ergab eine fünffaktorielle Struktur, wobei in der resultierenden Faktorenlösung die latenten Faktoren größtenteils inhaltlich gut interpretierbar sind. Die Faktoren zeigen darüber hinaus hohe Korrelationen untereinander. Vor diesem Hintergrund wurde zudem ein Bifaktormodell mit einem auf allen Items ladenden Generalfaktor und zunächst sieben untereinander und zum Generalfaktor orthogonalen spezifischen Faktoren als Strukturgleichungsmodell gewählt. Der spezifische Faktor zur Vorstellung LOK (Lokales Denken) zeigte keine signifikant von Null verschiedenen Faktorladungen, weshalb diese Items aus dem finalen Testinstrument entfernt wurden und das Modell anschließend mit sechs spezifischen Faktoren für die übrigen Vorstellungen neu berechnet wurden. Die konfirmatorische Faktorenanalyse zeigt eine gute Passung des Modells zu den Daten und eine Reliabilität des

Gesamtmodells von  $\omega = 0.93$ , wobei der Generalfaktor mit  $\omega_h = 0.79$  dominant ist, d. h. gegenüber den spezifischen Faktoren ( $\omega_h$  zwischen 0.20 und 0.43) den Großteil der Gesamtvarianz der Items erklärt. Dennoch zeigen die spezifischen Faktoren hochsignifikante Faktorladungen. Die Faktorenlösung für das Bifaktormodell ist in Abb. 2 dargestellt.

Weiterhin wurden die Daten ebenfalls bezüglich des korrekten Konzeptverständnisses ausgewertet und analoge Strukturanalysen durchgeführt. Hierbei wurden nur solche Kombinationen als korrekt („1“) kodiert, die sowohl in der ersten (Antwort) als auch der zweiten Stufe (Begründung) korrekt waren, ansonsten als falsch („0“). Die explorative Analyse zeigte hier mit einem vierfaktoriellen Modell die beste Passung, jedoch keine klar interpretierbare Faktorenlösung, was sich auch in der wiederum hohen Korrelation zwischen den Faktoren zeigt. Die konfirmatorische Analyse mit einem analogen Bifaktormodell (ein Generalfaktor, sechs orthogonale spezifische Faktoren) ergab dagegen auch hier eine gute Passung zu den Daten und eine Reliabilität von  $\omega = 0.97$ . Der Generalfaktor dominierte mit  $\omega_h = 0.87$  wiederum deutlich im Vergleich zu den spezifischen Faktoren mit  $\omega_h$  zwischen 0.07 und 0.31, welche den zugrundeliegenden Konzepten (z.B. Spannung, Stromstärke, Widerstand) entsprechen.



**Abb. 2:** Faktorenlösung des Bifaktormodells bezüglich der Lernendenvorstellungen (eigene Abbildung). Die Faktoren sowie die Items wurden entsprechend der Kürzel der zugrunde liegenden Lernendenvorstellungen benannt (vgl. Abschnitt 2).

## 5. Zusammenfassung und Ausblick

Die im vorliegenden Beitrag berichteten qualitativen und quantitativen Untersuchungen stützen die Validität des Testinstruments, insbesondere in Bezug auf die in Abschnitt 3 genannten Entwicklungsziele. Die Ergebnisse der Expertenbefragung bestätigen qualitativ die inhaltliche Angemessenheit der Items - sowohl hinsichtlich ihrer Relevanz für die Erfassung des Konzeptverständnisses als auch zur diagnostischen Erfassung der beschriebenen Lernendenvorstellungen. Die quantitativen Evaluationen weisen auf gute psychometrische Eigenschaften des Testinstruments hin, z.B. mit Blick auf die Item-Schwierigkeiten und -trennschärfen sowie die interne Konsistenz des Testinstruments ( $\alpha = 0.90$ ). Die Strukturanalyse mittels

Bifaktormodellen weist – basierend auf signifikanten Faktorladungen und zufriedenstellenden Reliabilitätswerten der Subskalen – darauf hin, dass sowohl die definierten Lernendenvorstellungen bzw. das zugrundeliegende Verständnis einzelner physikalischer Konzepte als latente Konstrukte die Item-Antworten systematisch beeinflussen.

Vor dem Hintergrund, dass Lernendenvorstellungen oftmals als stabile Konstrukte beschrieben werden, wäre es mit Blick auf die weitere Forschung von Interesse zu untersuchen, inwiefern die Sicherheit, mit der Lernende eine entsprechende Vorstellung in der Elektrizitätslehre vertreten, mit der Konsistenz korreliert, mit der die jeweilige Vorstellung über verschiedene Items hinweg vertreten wird. In anderen Inhaltsbereichen wie z.B. dem radioaktiven Zerfall konnte eine solche Korrelation nicht nachgewiesen werden [23].

Das Testinstrument kann bei Interesse an einem Einsatz in Forschung oder Lehre auf Anfrage gerne bei den Autoren angefordert werden.

## 6. Literatur

- [1] ENGELHARDT, PAULA V. ; GRAY, KARA E. ; REBELLO, N. SANJAY: How Many Students Does It Take Before *We* See the Light? In: *The Physics Teacher* Bd. 42 (2004), Nr. 4, S. 216–221
- [2] MALONEY, DAVID P. ; O’KUMA, THOMAS L. ; HIEGGELKE, CURTIS J. ; VAN HEUVELEN, ALAN: Surveying students’ conceptual knowledge of electricity and magnetism. In: *American Journal of Physics* Bd. 69 (2001), Nr. S1, S. S12–S23
- [3] COHEN, R. ; EYLON, B. ; GANIEL, U.: Potential difference and current in simple electric circuits: A study of students’ concepts. In: *American Journal of Physics* Bd. 51 (1983), Nr. 5, S. 407–412
- [4] MCDERMOTT, LILLIAN C. ; SHAFFER, PETER S.: Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding. In: *American Journal of Physics* Bd. 60 (1992), Nr. 11, S. 994–1003
- [5] STETZER, MACKENZIE R. ; VAN KAMPEN, PAUL ; SHAFFER, PETER S. ; MCDERMOTT, LILLIAN C.: New insights into student understanding of complete circuits and the conservation of current. In: *American Journal of Physics* Bd. 81 (2013), Nr. 2, S. 134–143
- [6] WILHELM, THOMAS ; HOPF, MARTIN: Schüler-vorstellungen zum elektrischen Stromkreis. In: SCHECKER, H. ; WILHELM, T. ; HOPF, M. ; DUIT, R. (Hrsg.): *Schülervorstellungen und Physikunterricht: Ein Lehrbuch für Studium, Referendariat und Unterrichtspraxis*. Berlin, Heidelberg : Springer, 2018 — ISBN 978-3-662-57270-2, S. 115–138
- [7] RHÖNECK, C VON: Vorstellungen vom elektrischen Stromkreis und zu den Begriffen Strom, Spannung und Widerstand. In: *Naturwissenschaften im Unterricht-Physik* Bd. 34 (1986), Nr. 13, S. 10–14
- [8] ENGELHARDT, PAULA VETTER ; BEICHNER, ROBERT J.: Students’ understanding of direct current resistive electrical circuits. In: *American Journal of Physics* Bd. 72 (2004), Nr. 1, S. 98–115
- [9] URBAN-WOLDRON, HILDEGARD ; HOPF, MARTIN: Entwicklung eines Testinstruments zum Verständnis in der Elektrizitätslehre. In: *Zeitschrift für Didaktik der Naturwissenschaften* Bd. Jg. 18 (2012), S. 201–227
- [10] IVANJEK, LANA ; MORRIS, LOUISA ; SCHUBATZKY, THOMAS ; HOPF, MARTIN ; BURDE, JAN-PHILIPP ; HAAGEN-SCHÜTZENHÖFER, CLAUDIA ; DOPATKA, LIZA ; SPATZ, VERENA ; U. A.: Development of a two-tier instrument on simple electric circuits. In: *Physical Review Physics Education Research* Bd. 17 (2021), Nr. 2, S. 020123
- [11] MCDERMOTT, LC ; VAN ZEE, EH: Identifying and addressing student difficulties with electric circuits. In: *Aspects of understanding electricity*, IPN Kiel, Germany (1985), S. 39–48
- [12] RHOENECK, CHRISTOPH VON ; VOELKER, BRUNO: Vorstellungen vom Stromkreis und ihr Einfluss auf den Lernprozess. *Der Physikunterricht*, 18(2).
- [13] RHÖNECK, C V: Schüleräußerungen zum Begriff der elektrischen Spannung beim Erklären der Meßwerte am Schalter. In: *Naturwissenschaften im Unterricht. Physik/Chemie* Bd. 29 (1981), Nr. 6, S. 210–215
- [14] MAICHLE, U: Schülervorstellungen zu Stromstärke und Spannung. In: *Naturwissenschaften im Unterricht. Physik/Chemie* Bd. 30 (1982), Nr. 11, S. 383–387
- [15] SHIPSTONE, DAVID: Pupils’ understanding of simple electrical circuits. Some implications for instruction. In: *Physics Education* Bd. 23 (1988), Nr. 2, S. 92–96
- [16] SHIPSTONE, D. M.: A study of children’s understanding of electricity in simple DC circuits. In: *European Journal of Science Education* Bd. 6 (1984), Nr. 2, S. 185–198
- [17] SHIPSTONE, D. M. ; RHÖNECK, C. V. ; JUNG, W. ; KÄRRQVIST, C. ; DUPIN, J.-J. ; JOHSUA, S. ; LICHT, P.: A study of students’ understanding of electricity in five European countries. In: *International Journal of Science Education* Bd. 10 (1988), Nr. 3, S. 303–316
- [18] SOKOLOFF, DAVID R.: Teaching Electric Circuit Concepts Using Microcomputer-Based Current/Voltage Probes. In: TINKER, R. F. (Hrsg.): *Microcomputer-Based Labs: Educational Research and Standards*. Berlin, Heidelberg :

- Springer Berlin Heidelberg, 1996 — ISBN 978-3-642-61189-6, S. 129–146
- [19] DING, LIN ; CHABAY, RUTH ; SHERWOOD, BRUCE ; BEICHNER, ROBERT: Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. In: *Physical Review Special Topics - Physics Education Research* Bd. 2, American Physical Society (2006), Nr. 1, S. 010105
  - [20] MCCOLGAN, MICHELE W. ; FINN, ROSE A. ; BRODER, DARREN L. ; HASSEL, GEORGE E.: Assessing students' conceptual knowledge of electricity and magnetism. In: *Physical Review Physics Education Research* Bd. 13, American Physical Society (2017), Nr. 2, S. 020121
  - [21] PEŞMAN, HAKI ; ERYILMAZ, ALI: Development of a Three-Tier Test to Assess Misconceptions About Simple Electric Circuits. In: *The Journal of Educational Research* Bd. 103 (2010), Nr. 3, S. 208–222
  - [22] HASAN, SALEEM ; BAGAYOKO, DIOLA ; KELLEY, ELLA L: Misconceptions and the Certainty of Response Index (CRI). In: *Physics Education* Bd. 34 (1999), Nr. 5, S. 294–299
  - [23] HULL, MICHAEL M. ; JANSKY, ALEXANDRA ; HOPF, MARTIN: Does confidence in a wrong answer imply a misconception? In: *Physical Review Physics Education Research* Bd. 18 (2022), Nr. 2, S. 020108
  - [24] LEINER, DOMINIK J.: Too Fast, Too Straight, Too Weird: Post Hoc Identification of Meaningless Data in Internet Surveys. In: *SSRN Electronic Journal* (2013)
  - [25] GRIFFARD, PHYLLIS BAUDOIN ; WANDERSEE, JAMES H.: The two-tier instrument on photosynthesis: What does it diagnose? In: *International Journal of Science Education* Bd. 23 (2001), Nr. 10, S. 1039–1052
  - [26] JORION, NATALIE ; GANE, BRIAN D. ; JAMES, KATIE ; SCHROEDER, LIANNE ; DiBELLO, LOUIS V. ; PELLEGRINO, JAMES W.: An Analytic Framework for Evaluating the Validity of Concept Inventory Claims. In: *Journal of Engineering Education* Bd. 104 (2015), Nr. 4, S. 454–496
  - [27] SCHERMELLEH-ENGEL, KARIN ; GÄDE, JANA C.: Modellbasierte Methoden der Reliabilitäts-schätzung. In: MOOSBRUGGER, H. ; KELAVA, A. (Hrsg.): *Testtheorie und Fragebogenkonstruktion*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2020 — ISBN 978-3-662-61532-4, S. 335–368