

„Wenn wir mal von der Reibung absehen...“



Im Physikunterricht über Validität sprechen

IRENE NEUMANN – STEFFEN WAGNER – BURKHARD PRIEMER

Das Aufzeichnen und Auswerten quantitativer Daten hat einen zentralen Stellenwert nicht nur in der Physik, sondern auch im Physikunterricht. Schüler/innen sollen dabei unter anderem beurteilen, welche Schlussfolgerungen aus Daten gezogen werden können und welche nicht. Die Frage nach der Validität der Schlussfolgerungen wird dabei jedoch oft vernachlässigt. Im folgenden Beitrag beleuchten wir den Validitätsbegriff aus Sicht der Physik und des Physikunterrichts und zeigen an einer Beispielaufgabe auf, wie er mit Schüler/innen thematisiert werden kann.

1 Einleitung

Stellen Sie sich vor: Ein Lehramtsstudent macht gerade in Ihrer Klasse ein Praktikum und soll eine Unterrichtsstunde zur Erdbeschleunigung g durchführen. Er hat dazu einen Demonstrationsversuch vorbereitet mit dem Ziel, die Größe der Erdbeschleunigung mit einer Unsicherheit von rund 10 % zu bestimmen. Ein Schüler hält einen Softball auf der Höhe der 30 cm-Marke an einem Meterstock fest und lässt ihn auf Kommando los. Eine Schülerin misst die Fallzeit mit einer Stoppuhr. Der Praktikant lässt die Klasse nun aus den Messwerten mithilfe der Gleichung $g = 2s/t^2$ die Erdbeschleunigung berechnen und muss gemeinsam mit den Schüler/innen feststellen, dass der berechnete Wert deutlich vom Literaturwert abweicht.

Schon bei der Durchführung des Versuchs sehen Sie dieses Ergebnis kommen. Sie haben den Versuch schon häufig durchgeführt und dabei festgestellt, dass Sie dazu am besten eine Stahlkugel, eine fixe Montage für die Starthöhe (mindestens 50 cm, besser 100 cm) und eine elektronische Stoppuhr mit automatischer Auslösung und Stoppfunktion benutzen. Außerdem führen Sie die Messungen nicht nur einmal durch, sondern bestimmen zehnmals die Fallzeit und nutzen dann den Mittelwert. Die Werte, die sich aus diesen Messungen für g ergeben, passen dann eigentlich immer ganz gut. Warum? Mit Ihrem Versuchsaufbau können Sie zum einen durch das mehrfache Messen der Zeit und die längeren Fallstrecken die Unsicherheit der Messung gut abschätzen und ggf. verringern. Sie haben zum anderen auch den Einfluss von Störvariablen wie Luftreibung und Reaktionszeit verkleinert und damit systematische Abweichungen minimiert. Was Sie automatisch machen – systematische Messabweichungen zu minimieren – hatte der Lehramtsstudent schlicht (noch) nicht im Blick.

Diese beispielhafte Situation zeigt, wie wichtig die „Richtigkeit“, die Validität bei physikalischen Messungen ist. Nur, wenn systematische Abweichungen ausgeschlossen oder zumindest minimiert werden, kann die physikalische Größe, die gemessen werden soll, auch wirklich erfasst werden. Das Beispiel zeigt auch, dass in vielen Situationen im Physikunterricht die Verringerung bzw. der Ausschluss von Störvariablen und damit systematischer Abweichungen meist „automatisch“ geschieht.

Oft werden Versuche gerade so ausgefeilt, dass es eben nicht zu (unerwarteten) Abweichungen in der gewünschten Zielgröße kommt. Damit wird den Schüler/innen jedoch die Möglichkeit genommen, sich mit dem Konzept der Validität und seiner Bedeutung für Messungen und die Interpretation von Daten auseinanderzusetzen. Mit dem vorliegenden Beitrag plädieren wir daher dafür, diese (zufälligen) Lerngelegenheiten nicht verpuffen zu lassen, sondern vielmehr weitere Lerngelegenheiten zu schaffen, die eine explizite Auseinandersetzung mit dem Begriff der Validität erlauben. Derartige Lerngelegenheiten würden auch der Forderung in den neu eingeführten Bildungsstandards (KMK, 2020) nachkommen, Messunsicherheiten stärker zu thematisieren. Dazu klären wir zunächst den Begriff der Validität aus einer allgemeinen Perspektive und diskutieren, warum und wie er im Physikunterricht thematisiert werden könnte. Am Eingangsbeispiel – der Bestimmung der Erdbeschleunigung im Fallversuch – zeigen wir dann auf, wie die Thematisierung von Validitätsaspekten in konkreten Aufgaben im Physikunterricht adressiert werden kann.

2 Validität – Begriffserläuterung und fachliche Klärung

In Schulbüchern zur Physik (z.B. ACKERMANN et al., 2020) und Hochschullehrbüchern (z.B. GIANCOLI, 2019; Tipler & Mosca, 2015) zur Experimentalphysik wird die Validität als eigenständiges Konzept kaum expliziert. In anderen empirischen (vor allem sozialwissenschaftlichen und psychologischen) Fachgebieten ist er dagegen sehr präsent. Die dort vorgenommene Konzeptualisierung lässt sich auf bestimmte Aspekte der physikalischen Erkenntnisgewinnung sinnvoll übertragen. Um Klarheit über den Bedeutungsrahmen in der Physik und im Physik- und NaWi-Unterricht zu schaffen, soll der Begriff „Validität“ (bzw. „valide“) daher zunächst etwas genauer unter die Lupe genommen werden. Woher kommt er, was bedeutet er im engeren Sinn und wie wird er im Kontext des Messens verwendet?

Der Begriff „Validität“ stammt vom lateinischen „validus“ ab, wo er soviel bedeutet wie kräftig oder stark, befestigt, wirksam, gesund, mächtig und gewaltig (LANGENSCHIEDT LATEIN-DEUTSCH, kein Datum). Auch im Englischen wird das Adjektiv

„valid“ verwendet und meint hier wohlbegründet, stichhaltig, rechtswirksam und gesund (LANGENSCHIEDT ENGLISCH-DEUTSCH, kein Datum). Mit diesen leicht unterschiedlichen Bedeutungen in den Quellsprachen ist die Bedeutung des bildungssprachlichen Lehnwortes „valide“ zwar nicht exakt festgelegt, lässt sich aber eingrenzen.

Im wissenschaftlichen Kontext bezieht sich Validität im Wesentlichen auf Aussagen. Das Cambridge Dictionary beschreibt als Validität *“the quality of being based on truth or reason, or of being able to be accepted”* (Cambridge Dictionary, kein Datum). Konkret ist aus Perspektive der Logik eine Aussage (z.B. ein Argument) genau dann valide, wenn sie eine Prämisse und eine Schlussfolgerung so miteinander verknüpft, dass es unmöglich ist, mit dem Argument aus einer richtigen Prämisse eine falsche Schlussfolgerung zu ziehen (Internet Encyclopedia of Philosophy, kein Datum). Diese Forderung lässt sich auf experimentelle Kontexte in der Physik übertragen.

In Bezug auf empirisch gewonnene Aussagen in den Sozialwissenschaften beschreibt die Validität die Güte von Aussagen über Zusammenhänge zwischen zwei Variablen. Dabei gibt es verschiedene Facetten oder Konzepte der Validität (vgl. DÖRING & BORTZ, 2016), wie z.B. die externe und interne Validität bezüglich Aussagen über beobachtete Zusammenhänge (SCHULZ et al, 2012). Die externe Validität beschreibt, ob sich eine Aussage über einen experimentell beobachteten Zusammenhang auf einen anderen (externen) Kontext übertragen oder sogar verallgemeinern lässt. Als intern valide wird eine Aussage über einen Variablenzusammenhang dann bezeichnet, wenn er sehr wahrscheinlich eine kausale Beziehung darstellt und keine andere Variable für diesen Zusammenhang verantwortlich ist (DÖRING & BORTZ, 2016).

Ein Beispiel für eine nicht valide Aussage soll die Sinnhaftigkeit dieser Festlegung und die Übertragbarkeit auf physikalische Messvorgänge illustrieren. Wenn man einen Erlenmeyerkolben zur Hälfte mit Alkohol füllt, mit einem Stopfen mit Bohrung verschließt und durch diese Bohrung ein dünnes Steigrohr einführt, dann wird der Alkoholpegel in dem Steigrohr bei steigenden Temperaturen ebenfalls steigen (selbst gebautes Thermometer). Nicht valide wäre nun die Formulierung eines kausalen Zusammenhangs zwischen Alkoholtemperatur und -Volumen. Diese Aussage ist deshalb nicht valide, weil für den Anstieg des Alkoholpegels die Ausdehnung der Luft in dem Gefäß bei steigenden Temperaturen in dieser Aussage nicht berücksichtigt wäre. Wenn das Ziel der Messung darin liegt, einen möglichen Zusammenhang zwischen Alkoholtemperatur und -volumen zu zeigen, dann muss die Luft aus dem Gefäß eliminiert werden, z.B. indem das gesamte Gefäß mit Alkohol befüllt wird und sich keine Luftbläschen mehr darin befinden. Andernfalls stimmen Messung (Änderung des Gas- und Alkoholvolumens) und Messziel (Änderung des Alkoholvolumens) nicht mehr überein und die Messung des Alkoholvolumens beinhaltet eine systematische Abweichung. Zudem ist die direkt beobachtete, abhängige Variable eine Längenänderung, die vorgeblich mit einer Volumenänderung assoziiert wird. Nur dann, wenn einzig die Volumenänderung tatsächlich auch eine Längenänderung

hervorrufen kann, kann man aufgrund der beobachteten Längenänderung eine valide Aussage über das thermische Ausdehnungsverhalten von Alkohol machen. Darüber hinaus muss die Messung der unabhängigen Variable (Temperatur) mit einem anderen Instrument als mit einem Alkoholthermometer erfolgen, da sonst ebenfalls eine nicht-valide Aussage folgt. Die Volumenänderung könnte dann nicht mehr auf die Temperaturänderung zurückgeführt werden.

Im Kontext des Messens oder des naturwissenschaftlichen Umgangs mit Daten spricht man manchmal davon, Daten, ein Verfahren, ein Instrument oder Messungen seien valide oder nicht valide. Nach den obigen Ausführungen sind solche Formulierungen eigentlich nicht ganz korrekt. Weder das Verfahren, noch das Instrument, die Messungen oder die Daten sind nicht valide, sondern die Aussagen, die nach einem solchen Verfahren, mithilfe eines Instruments, der Messungen oder auf Basis von Daten getroffen werden. Für das anvisierte Messziel, eine bestimmte Aussage über einen Zusammenhang zu rechtfertigen, ist dann das Verfahren oder das Instrument schlicht ungeeignet, weil es „nicht das misst, was es messen soll“. Die oben genannte fehlende Thematisierung des Validitätsbegriffs in Schul- und Hochschulbüchern zur Physik könnte auf diese sprachliche Ausdifferenzierung einerseits und den engen Zusammenhang zwischen Validität und systematischen Abweichungen andererseits zurückzuführen sein. Letztere werden nämlich durchaus thematisiert (z.B. GIANCOLI, 2019), die Tatsache, dass Validität aber keine Eigenschaft von Daten, Instrumenten oder Messungen, sondern von Aussagen ist, aber in der Regel nicht.

3 Validität im Physikunterricht

3.1 Argumente für eine Thematisierung des Validitätsbegriffs

Warum aber sollte der Begriff der Validität im Physikunterricht behandelt werden, wenn er doch scheinbar nur selten als eigenständiges Konzept in der Physik expliziert wird?

Erstens stellt das, was implizit unter Validität verstanden wird, eine wichtige Säule physikalischer Erkenntnisgewinnung dar: Ohne Validität sind Aussagen über Zusammenhänge zwischen zwei Variablen nicht möglich. Insofern ist es in der Tat recht verwunderlich, dass der Begriff der Validität nur so selten in physikalischen Lehrbüchern Thema ist. Umgekehrt könnte aber auch gerade das seine zentrale Relevanz unterstreichen: Indem Validität die Grundvoraussetzung aller – sinnvollen – Aussagen über Messvorgänge ist, wird wie selbstverständlich darauf geachtet, dass sie gewährleistet ist (nämlich zum Beispiel durch das Vermeiden systematischer Abweichungen, vgl. BEVINGTON & ROBINSON, 2003).

Zweitens wird diese „Selbstverständlichkeit“ im Physikunterricht besonders deutlich. Hier geht es oft gerade darum, die Schüler/innen „einfache“ Zusammenhänge erkunden zu lassen und Einflüsse, beispielsweise der (Luft-)Reibung oder Ausdehnung der Körper, in der mathematischen Beschreibung nicht zu

berücksichtigen. Entsprechend werden Versuche im Physikunterricht meist direkt so geplant, dass solche Effekte möglichst nicht ins Gewicht fallen. Wenn derartige Vereinfachungen dann aber nicht hinreichend thematisiert werden, kann es vorkommen, dass die Validität der Aussagen über die untersuchten Zusammenhänge unbemerkt verletzt und somit ein verzerrtes Bild naturwissenschaftlicher Erkenntnisgewinnung vermittelt wird. Gerade in solchen Situationen wird eine Lerngelegenheit nicht ausgenutzt, über die Bedeutung und die Notwendigkeit von Validität zu reflektieren. Und auch wenn Schüler/innen Abweichungen als solche bemerken, fällt es ihnen erfahrungsgemäß oft schwer, Ursachen dafür zu finden (jenseits der Reibung). Eine Reflexion über die Validität der Messung würde ihnen die Möglichkeit geben, diese umfassend zu ergründen.

Drittens wären gerade solche Lerngelegenheiten auch vor dem Hintergrund der Vermittlung von zukunftsrelevanten, überfachlichen Kompetenzen besonders erstrebenswert. So zählt die Bewertung der Validität von Daten beispielsweise zu den Teilaspekten des sogenannten kritischen Denkens (z.B. TIRUNEH et al., 2017). Schüler/innen müssen als zukünftige mündige Bürger/innen mit Daten umgehen – auch mit solchen, die nicht selbst generiert wurden und vielleicht auch komplexere Themen wie den Klimawandel o.ä. adressieren. Ein Verständnis der Wichtigkeit der Validität für die Interpretation von Daten, das Ziehen von Schlussfolgerungen und die daraus resultierenden Entscheidungen ist dafür unabdingbar. Dadurch, dass das Messen in der physikalischen Erkenntnisgewinnung einen so zentralen Stellenwert einnimmt, ist Physikunterricht damit besonders geeignet, das Reflektieren der Validität und damit einen Aspekt des kritischen Denkens mit Schüler/inne/n einzuüben.

3.2 Teilaspekte für den Physikunterricht

Um Schüler/innen an das Konzept der Validität von Aussagen über Messungen heranzuführen, schlagen wir vor, Teilaspekte dieses doch umfassenden Begriffs zu thematisieren. Dabei können unter anderem die folgenden Leitfragen und Teilschritte hilfreich sein:

- a) Was ist das Ziel der Messung? Die Schüler/innen
 1. ... beschreiben das Messziel – in Worten und ggf. mithilfe mathematischer Formelsprache.
 2. ... identifizieren unabhängige, abhängige und Kontrollvariablen des zu untersuchenden Zusammenhangs.
 3. ... ordnen einem selbst erzeugten oder vorgegebenen Datensatz eine oder mehrere physikalische Größen zu (wie z.B. Stromstärke, Spannung, Zeit, Kraft, etc.).
- b) Wie wurden die einzelnen Größen gemessen? Die Schüler/innen
 1. ... benennen, welche Größen von der Messapparatur wie erfasst werden (Funktion des experimentellen Aufbaus).
 2. ... beschreiben bzw. begründen die Auswahl der Stichprobe (z.B. Anzahl der Messwiederholungen, Auswahl verschiedener Probe-Objekte).
 3. ... nennen die experimentellen Voraussetzungen/ Randbedingungen, unter denen ein Datensatz entstanden ist, wie beispielsweise Reibung, Messbereich, genutzte Messgeräte (und deren Auflösungsvermögen) etc.

4. ... schätzen bei externen Datensätzen die Vertrauenswürdigkeit der Quelle ein (z.B. durch Überprüfung der Plausibilität/Qualität der Messung).
- c) Inwieweit passen Messziel, Messung und Theorie überein? Die Schüler/innen
 1. ... identifizieren Störvariablen, die die Messung beeinflussen (könnten) und schätzen deren Einfluss ab.
 2. ... nennen Unterschiede zwischen Messziel bzw. Messgleichung und realer Durchführung der Messung.
 3. ... schätzen ab, wie wichtig ggf. Unterschiede zwischen Messziel und Messung zu bewerten sind.
 4. ... vergleichen Messergebnisse mit Theorie/Literaturwert/anderen Messungen.
 - d) Wie lässt sich die Messung verbessern? Die Schüler/innen ... nennen Möglichkeiten, Unsicherheiten und Abweichungen zu verringern.
 - e) Was ist Validität im Allgemeinen? Die Schüler/innen ... geben eine allgemeine Definition bzw. Beschreibung des Begriffes der Validität.

Diese Liste verdeutlicht, wie umfassend das Konstrukt der Validität ist und wie viele Teilaspekte dazugehören. Um die dahinterstehenden Kompetenzen zu verdeutlichen, illustrieren wir die Einschätzung der Validität anhand eines Beispiels.

4 Die Validität bewerten am Beispiel „Bestimmung der Erdbeschleunigung im Fallversuch“

Eine mögliche unterrichtliche Umsetzung, bei der der Fokus der Schüler/innen auf die Bewertung der Validität von Aussagen gelenkt wird, ist für das eingangs erwähnte Beispiel in Kasten 1 dargestellt. Die Lernenden erhalten zunächst Hintergrundinformationen darüber, wie in einem Fallversuch die Erdbeschleunigung g gemessen werden soll. Neben dem Versuchsaufbau werden ihnen auch Messwerttabellen, inklusive der berechneten Werte von g , zur Verfügung gestellt. Entsprechend den oben dargestellten Teilaspekten leiten die eigentlichen Arbeitsaufträge die Schüler/innen nun durch einen Prozess, in dem sie die Messungen und die damit erhaltenen Daten mit Blick auf die Validität reflektieren. Dabei sollen sie zunächst das Ziel der Messung und die dafür nötigen physikalischen Größen identifizieren sowie ausgehend von einer genaueren Betrachtung der eigentlichen Durchführung mögliche Störvariablen benennen. Anschließend sollen die Schüler/innen abschätzen, wie sehr diese Störvariablen die berechneten Werte von g beeinflussen und wie sie ggf. reduziert oder ausgeschlossen werden können, um den Messvorgang valider zu gestalten. Eine abschließende Aufgabe soll dann anregen, auf der Metaebene die Idee der Validität und ihre Bedeutung für physikalische Messvorgänge zu reflektieren.

Diese Beispielaufgabe adressiert viele – z. T. recht anspruchsvolle – Kompetenzen, die zu einem Verständnis des Begriffs der Validität führen sollen. In der Praxis ist es durchaus sinnvoll, ggf. zunächst nur Teile dieser Aufgabe zu verwenden, um Schüler/innen nicht zu überfordern. Hier in diesem Beitrag ist es aber unser Ziel, das Konstrukt der Validität möglichst vollständig durch operationalisierte Arbeitsaufträge abzubilden.



Abb. 1. Aufbau eines Experiments zur Bestimmung der Erdbeschleunigung

In einem Experiment soll die Erdbeschleunigung g mit Hilfe der Messgleichung $g = 2s/t^2$ bestimmt werden. Dafür werden für zwei gleich große Kugeln (Durchmesser 30 mm) aus unterschiedlichem Material (Stahl und Styropor) die Fallzeiten für zwei verschiedene Höhen (50 cm und 150 cm, gemessen mit einem Gliedermaßstab) je sechsmal bestimmt. Die Zeitmessung erfolgt mit einer elektronischen Stoppuhr, die die Fallzeiten zwischen dem manuell ausgeführten Start und dem Stopp durch eine Lichtschranke misst (siehe Abb. 1 und Tab. 1).

	Styroporkugel ($m = 0,6 \text{ g} \pm 0,1 \text{ g}$)		Stahlkugel ($m = 110,2 \text{ g} \pm 0,1 \text{ g}$)	
	$s = 50,0 \text{ cm} \pm 0,5 \text{ cm}$	$s = 150,0 \text{ cm} \pm 0,5 \text{ cm}$	$s = 50,0 \text{ cm} \pm 0,5 \text{ cm}$	$s = 150,0 \text{ cm} \pm 0,5 \text{ cm}$
Höhe s				
1. Messung für t	0,329 s	0,604 s	0,319 s	0,561 s
2. Messung für t	0,320 s	0,619 s	0,313 s	0,547 s
3. Messung für t	0,316 s	0,600 s	0,316 s	0,553 s
4. Messung für t	0,332 s	0,594 s	0,317 s	0,551 s
5. Messung für t	0,357 s	0,621 s	0,348 s	0,548 s
6. Messung für t	0,319 s	0,597 s	0,313 s	0,556 s
MW und Unsicherheit (SD) für t	0,329 s \pm 0,015 s	0,606 s \pm 0,011 s	0,321 s \pm 0,013 s	0,553 s \pm 0,017 s
g und Unsicherheit für g nach $g = 2s/t^2$	9,25 m/s ² \pm 0,94 m/s ²	8,17 m/s ² \pm 0,33 m/s ²	9,71 m/s ² \pm 0,91 m/s ²	9,82 m/s ² \pm 0,63 m/s ²

Tab. 1. Fallzeiten t (in s) für die Styropor- und die Stahlkugel bei unterschiedlichen Höhen s , Mittelwerte (MW) und Standardabweichungen (SD) für t sowie die Messergebnisse für die Fallbeschleunigungen g mit deren Unsicherheiten.

Aufgaben

1. Formuliere das Ziel der Messungen mit eigenen Worten (vgl. a.1. in Abschnitt 3.2).

Skizze einer Musterlösung: Bestimmung der Fallbeschleunigung nach $g = 2s/t^2$.

2. Benenne alle Größen und Eigenschaften, die für die Messung festgehalten werden (vgl. a.3. in Abschnitt 3.2).

Skizze einer Musterlösung: Material, Oberflächenbeschaffenheit, Massen und Durchmesser der Kugeln, Fallhöhen und dafür benötigte Zeiten

3. Beschreibe kurz den Ablauf der Messungen (vgl. b.1. in Abschnitt 3.2).

Skizze einer Musterlösung: Die Kugeln befinden sich in einer bestimmten Höhe in Ruhe. Durch Loslassen beginnen sie frei zu fallen und lösen unmittelbar die erste Lichtschranke aus, sodass eine digitale Zeitmessung beginnt. Nach Durchlaufen einer vorgegebenen Strecke (gemessen mit einem Gliedermaßstab) beendet eine zweite Lichtschranke die Zeitmessung.

4. Beschreibe auf Basis der Beschreibung in 3., wie die Messungen durch weitere (nicht in 2. genannte) Größen beeinflusst werden könnten. Schätze das Ausmaß dieses Einflusses auf Basis der Messergebnisse des Experiments oder anhand der Betrachtung des Versuchsablaufs begründet ab. Unterscheide hierbei zwischen den Messungen mit den unterschiedlichen Kugeln (vgl. b.3. und c.1. in Abschnitt 3.2).

Skizze einer Musterlösung: Weitere physikalische Größen sind:

- Die Reibungskraft während des freien Falls: die unterschiedlichen Massen der Kugeln führen zu unterschiedlichen Fallbeschleunigungen (kleinere Masse → geringere Fallbeschleunigung).
- Die Oberflächenbeschaffenheit der Kugeln hat auch einen Einfluss auf die Reibungskraft (größerer c_w -Wert → geringere Fallbeschleunigung).
- Aus den Messungen lässt sich nicht entnehmen, ob die bedeutsamen Unterschiede zwischen Stahl- und Styroporkugeln bei $s = 150$ cm auf die Masse oder die Oberflächenbeschaffenheit zurückzuführen sind.
- Unterschiedliche Zeitverzögerungen bei Start und Stopp: dieser Effekt kann auftreten, ist aber aufgrund der sehr kurzen zu erwartenden Verzögerungen als eher klein (unbedeutsam) einzuschätzen.

5. Beurteile auf Basis deiner Überlegungen in 4., für welche Messungen (welche Kugel, welche Höhe) eine Abweichung zwischen Messgleichung und realer Durchführung der Messung anzunehmen ist und schätze diese bzgl. der Stärke der Abweichung ab (vgl. c.2. und c.4. in Abschnitt 3.2).

Skizze einer Musterlösung:

- Wenn keine Abweichungen zwischen realer Durchführung und Messgleichung vorliegen, dann sollten alle Messungen (zwei Kugeln, zwei Höhen) miteinander vereinbar sein, d.h. im Rahmen der Unsicherheiten vereinbare Ergebnisse erzeugen. Das ist aber nicht der Fall: bei der Styroporkugel und 150 cm Fallhöhe weichen die Ergebnisse bedeutsamen von den Ergebnissen der Stahlkugel ab. Das deutet auf eine fehlende Validität hin, die in der Masse oder der Oberflächenbeschaffenheit der Styroporkugel begründet sein kann.
- Wird deduktiv vorgegangen, so kann ein Referenzwert hinzugezogen werden: der Literaturwert sollte in das Unsicherheitsintervall der Messergebnisse fallen. Das ist bei allen Messungen bis auf den Fall der Styroporkugel mit 150 cm Fallhöhe der Fall. Das deutet auf eine fehlende Validität hin.
- Die Werte der Stahlkugel zeigen die größte Vereinbarkeit mit dem Literaturwert, d.h. hier ist die Abweichung am geringsten.
- Die vergleichsweise großen Unsicherheiten der Messungen bei 50 cm führen dazu, dass Verletzungen der Validität schwieriger zu erkennen sind, da bedeutsame Unterschiede zwischen den Messungen schlechter erkennbar sind.

6. Beurteile auf Basis deiner Überlegungen in 5. für jeden Messdurchgang (Spalte in der Tabelle), ob Messvorgang (realisierte Durchführung) und Messziel miteinander vereinbar sind. Fülle ein begründetes Urteil mit den vorangegangenen Überlegungen für jeden Messvorgang, ob die generierten Messdaten hinreichend valide sind. Valide („gültig“) ist ein Messvorgang dann, wenn tatsächlich das gemessen wird (realisierte Messung), was auch gemessen werden sollte (Ziel der Messung) und somit systematische Störvariablen ausgeschlossen wurden (vgl. c.3. und c.4. in Abschnitt 3.2).

Skizze einer Musterlösung:

- Styropor-Kugel mit 150 cm Fallhöhe: Vereinbarkeit ist nicht gegeben; Messung ist nicht valide; Grund (bei induktivem Vorgehen): bedeutsamer Unterschied zu Ergebnissen mit der Stahlkugel; Grund (bei deduktivem Vorgehen): Nicht-Vereinbarkeit mit Literaturwert.
- Stahlkugel mit 150 cm Fallhöhe: Vereinbarkeit ist gegeben; Messung ist valide.
- Styropor- und Stahlkugel mit 50 cm Fallhöhe: Vereinbarkeit ist gegeben; Messungen sind valide (allerdings ist die relative Unsicherheit mit rund 10 % recht groß).

7. Formuliere auf Basis deiner Überlegungen in 6. einen Vorschlag, wie der Versuchsaufbau oder die Durchführung der Messungen geändert werden könnten, sodass nicht ausreichend valide Messungen valide werden könnten (vgl. d. in Abschnitt 3.2).

Skizze einer Musterlösung:

- Verzicht auf Styropor-Kugeln bei Fallhöhen größer 0,5 m
- mehrere Messungen mit der Stahlkugel (Erhöhung auf 10 Messungen je Höhe, weitere Höhen bei der Durchführung mit aufnehmen)
- Verringerung der Unsicherheiten durch größere Fallhöhen
- theoretisch: Messungen im Vakuum

8. Reflektiere die Schritte 1 bis 7 und erläutere anhand deiner Überlegungen damit den Begriff der Validität. Begründe, ob man Aussagen über eine Messung in valide oder nicht valide einteilen sollte oder ob Validität unterschiedliche Ausprägungen haben kann – führt z.B. eine Messung zu valideren Aussagen als eine andere (vgl. e. in Abschnitt 3.2)?

Skizze einer Musterlösung:

- Treten starke Variationen innerhalb der Messdaten auf (unterschiedliche Werte für g bei unterschiedlichen Versuchsbedingungen), so deutet dies auf unkontrollierte (Stör-)Einflüsse hin; diese haben einen Einfluss auf die Validität der Messungen.
- Sind diese Einflüsse (Abweichungen) klein, können sie ggf. „in Kauf“ genommen werden; die Aussagen über die Messungen sind dann zwar nicht „vollständig“ valide, die Abweichungen sind aber tolerierbar. Bspw. lässt sich dieser Versuch zur Bestimmung von g in der Schule nicht im Vakuum durchführen (die Reibung wäre dann Null).
- Eine perfekte Validität ist nie zu erreichen, da selbst physikalische Größen – die durch Messgleichungen ausgedrückt werden – nie perfekt unter Angabe aller Bedingungen festgelegt werden können (Stichwort: „wahrer Wert“). Deshalb ist immer eine kritische Abschätzung notwendig, wie stark die Validität verletzt ist. Eine einfache Einteilung in valide/ unvalide ist deshalb nicht möglich und auch nicht sinnvoll.

Kasten 1. Skizze einer Aufgabe mit Musterlösung zur Beurteilung der Validität der Aussagen über am Beispiel der Bestimmung der Erdbeschleunigung mit einem Fallversuch

5 Diskussion

Mit dem vorliegenden Artikel soll ein Plädoyer für die explizite Thematisierung des Begriffs der Validität im Kontext von Messungen im Physikunterricht gegeben werden. Die Validität ist zentral, um aus Daten gültige Schlussfolgerungen ziehen zu können. Es ist anzunehmen, dass im Physikunterricht bei Experimenten mit Messungen die Validität berücksichtigt wird, jedoch oft für die Schüler/innen „versteckt“ bleibt und nicht explizit als wichtiges Qualitätsmerkmal der Physik thematisiert wird. Um Lehrkräfte für dieses Thema zu sensibilisieren, haben wir ausgehend von einer Begriffsklärung Leitfragen und Teilaspekte vorgeschlagen, mit deren Hilfe Schüler/innen schrittweise an das wichtige Konstrukt der Validität herangeführt werden können. Die dabei geförderten Kompetenzen reichen von niedrigen (z.B. Messziel beschreiben, Variablen identifizieren) bis hin zu höheren Anforderungen (z.B. Einfluss von Störvariablen abschätzen, Passung von Messziel und Messung beurteilen).

Wie diese Kompetenzen konkret in Arbeitsaufträgen umgesetzt werden können, haben wir im Beispiel zur Bestimmung der Erdbeschleunigung aufgezeigt. Dabei ist die Aufgabe als Prototyp zu verstehen, der in diesem Umfang in der Regel nur von sehr leistungsstarken Schüler/innen bewältigt werden kann. Denken Sie zurück an die vermeintlich gescheiterte Unterrichtsstunde des fiktiven Lehramtsstudierenden. Sobald dieser mit seinen Schüler/innen feststellt, dass der gemessene Wert von g erheblich vom Literaturwert abweicht, ist eine ideale Lerngelegenheit gegeben, über Validität zu reflektieren. In Anlehnung an die Fragen, wie sie in der Beispielaufgabe formuliert sind, könnte der Studierende beispielsweise zunächst die Frage aufwerfen, ob es kritisch ist, dass der gemessene Wert vom Literaturwert abweicht, und dann sukzessive mit den Schüler/innen ergründen, wo eventuell Probleme in der Messung liegen, wie diese behoben werden könnten, und warum es so wichtig ist, sich über derartige Messabweichungen bewusst zu werden.

Aber nicht nur in Situationen, in denen die Schüler/innen selbst Daten generieren, sollte an Validität gedacht werden. Die vorgeschlagenen Teilaspekte und Leitfragen können überall da

genutzt werden, wo aus Daten Schlüsse gezogen werden. Dies gilt beispielsweise auch für Datenmaterial, das im Rahmen von Bewertungsaufgaben benutzt wird (für eine beispielhafte Bewertungsaufgabe, die Datenmaterial einbezieht, siehe IQB, ohne Datum). In derartigen Situationen sind die konkreten Messvorgänge sicher nicht so zugänglich wie sie dies in Situationen eigenen Experimentierens sind. Umso wichtiger wird daher dann die Frage nach der Vertrauenswürdigkeit der Quellen.

Insgesamt hoffen wir, mit diesem Beitrag Physiklehrkräfte sensibilisiert zu haben, Validität als einen wichtigen Pfeiler der physikalischen Erkenntnisgewinnung zu thematisieren – gerade auch in Experimentiersituationen, die sie bewusst „vereinfacht“ haben, um den Zusammenhang zweier Variablen aufzuzeigen und den Einfluss von systematischen Störvariablen möglichst gering zu halten (oder sie gar nicht berücksichtigen).

Literatur

ACKERMANN, P., BECKER, P., BÖHLEMANN, R., BREUER, E., BURZIN, S., BUSCH, C., ... WINTER, R. (2020). *Fokus Physik S II. Gesamtband* (1. Auflage, 6. Druck). Berlin: Cornelsen.

BEVINGTON, P.R. & ROBINSON, D.K. (2003). *Data reduction and error analysis for the physical sciences*, 3rd edition. Boston: McGraw-Hill.

Cambridge Dictionary (ohne Datum). <https://dictionary.cambridge.org/dictionary/english/validity> (15.07.2022).

DÖRING, N. & BORTZ, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin: Springer.

GIANCOLI, D. C. (2019). *Physik* (4., aktualisierte Auflage). Hallbergmoos: Pearson.

Internet Encyclopedia of Philosophy (ohne Datum). <https://iep.utm.edu/val-snd/> (15.07.2022).

Institut zur Qualitätsentwicklung im Bildungswesen (IQB) (ohne Datum). <https://www.iqb.hu-berlin.de/appsrc/taskpool/data/taskpools/getTaskFile?id=p03^windenergie^f4534> (15.07.2022).

Kultusministerkonferenz (KMK) (2020). Bildungsstandards im Fach Physik für die Allgemeine Hochschulreife. Beschluss vom 18.06.2020. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2020/2020_06_18-BildungsstandardsAHR_Physik.pdf (02.09.2022).

Langenscheidt Englisch-Deutsch (ohne Datum). <https://de.langenscheidt.com/englisch-deutsch/valid> (15.07.2022).

Langenscheidt Latein-Deutsch (ohne Datum). <https://de.langenscheidt.com/latein-deutsch/validus> (15.07.2022).

SCHULZ, A., WIRTZ, M. & STARAUSCHEK, E. (2012). Das Experiment in den Naturwissenschaften. In W. RIESS, M. WIRTZ, B. BARZEL & A. SCHULZ (Hg.), *Experimentieren im mathematisch-naturwissenschaftlichen Unterricht. Schüler lernen wissenschaftlich denken und arbeiten* (S. 15-38). Münster, New York: Waxmann.

SHADISH, W. R., COOK, T. D. & CAMPBELL, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton, Mifflin and Company.

TIPLER, P. A. & MOSCA, G. (2015). *Physik: Für Wissenschaftler und Ingenieure* (7. dt. Aufl). Berlin: Springer Spektrum.

TIRUNEH, D.T., DE COCK, M., WELDESASSIE, A.G., ELEN, J. & JANSSEN, R. (2017). Measuring Critical Thinking in Physics: Development and Validation of a Critical Thinking Test in Electricity and Magnetism. *International Journal of Science and Mathematics Education* 15, 663–682. <https://doi.org/10.1007/s10763-016-9723-0>

IRENE NEUMANN ist Leiterin der Forschungsgruppe „Lehren und Lernen an der Schnittstelle zwischen Physik und Mathematik“ am Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN) in Kiel. Kontakt: ineumann@leibniz-ipn.de

BURKHARD PRIEMER ist Professor für Didaktik der Physik an der Humboldt-Universität zu Berlin. Er forscht zum Umgang mit Messdaten und zum naturwissenschaftlichen Argumentieren. Kontakt: priemer@physik.hu-berlin.de

STEFFEN WAGNER ist wissenschaftlicher Mitarbeiter der Didaktik der Physik an der Humboldt-Universität zu Berlin. Er forscht zu Erklärungen und Modellen, zum kritischen Denken und zu naturwissenschaftlichen Interessen.

Kontakt: steffen.wagner@physik.hu-berlin.de ■□