

Alles Reibung oder was?

- Welchen Effekt oft genannte Einflüsse tatsächlich auf Messergebnisse haben -

Christoph Holz*, **Susanne Heinicke***

*WWU Münster, Institut für Didaktik der Physik, Wilhelm-Klemm-Straße 10, 48149 Münster
christoph.holz@uni-muenster.de, susanne.heinicke@uni-muenster.de

Kurzfassung

Eine typische Situation aus dem Physikunterricht: Die Ergebnisse einer Experimentierphase werden an die Tafel geschrieben und weichen deutlich von den theoretischen Erwartungen ab. Die Erklärung der Diskrepanz erfolgt mit den üblichen Verdächtigen: Reibung, Reaktionszeit, Luftfeuchtigkeit, ungenaue Messgeräte ... aber wie groß ist der Einfluss solcher Größen auf das Ergebnis tatsächlich? Wir halten ein adäquates Verständnis von physikalischen Messwerten, deren (Un-)Genauigkeit und Aussagekraft für die Ausbildung werdender Physiklehrkräfte für grundlegend wichtig. Insbesondere wird dies für ein sinnvolles Einbringen und Vorleben von Fehlerkultur und dem fachlich angemessenen Umgang mit Messunsicherheiten im Unterricht benötigt.

Die dazu konzipierte Lernumgebung beinhaltet daher im Wesentlichen drei Aspekte:

- Die Berücksichtigung der Vorstellungen der Studierenden
- Eine moderne fachliche Modellierung von Messwerten durch Wahrscheinlichkeitsdichtefunktionen, wie durch den ISO-Standard GUM vorgeschlagen
- Zwei hierzu erstellte Apps, die es ermöglichen
 - a) Messdaten nach GUM zu verrechnen und darzustellen und
 - b) die auf ein konkretes Experiment einflussnehmenden Faktoren zu variieren und deren Auswirkungen zu beobachten.

Im Rahmen des Vortrages wurden insbesondere die letzteren Applikationen vorgestellt.

1. Einleitung

„Bestimmen Sie den Ortsfaktor g experimentell.“ Eine Arbeitsanweisung, wie sie beispielsweise in einer Einführungsveranstaltung zu einem Grundpraktikum gegeben werden kann. Diese Aufgabe allein genügt in der Regel um – nach Zusammenführung der Ergebnisse verschiedener Teams – eine hitzige Debatte über die Aussagekraft und Qualität von empirisch gewonnenen Daten auszulösen. Sei es auf Grund von Verwunderung über stark abweichende Ergebnisse oder auf Grund von Unglaubwürdigkeit über die Ehrlichkeit der Gruppe, die ein Ergebnis minimaler Abweichung vom Literaturwert zu berichten hat. In unserer Studie verwenden wir hierzu ein einfaches Fallexperiment. Dieses kann mit einfachen Mitteln (Stoppuhr, Maßband und Tennisball) durchgeführt werden, birgt aber auch viele Unsicherheitsquellen in sich. Auch auf schulische Physikexperimente ist eine solche Situation leicht übertragbar: Die Lehrkraft benötigt hier das notwendige Wissen über und Verständnis vom Umgang mit Messdaten, um die Situation fachgerecht, aber auch für die Schülerinnen und Schüler zufriedenstellend anzuleiten.

Wo haben die Studierenden die Möglichkeit den Umgang mit empirischen Daten zu lernen? Meist findet dies in Formaten wie einem Grundpraktikum statt.

Ein näherer Blick in die Forschung rund um das Verständnis von Messdaten – häufig verbunden mit der Beforschung solcher Praktika – zeigt jedoch ein recht unzufriedenstellendes Bild: Sêrê et al. forderten bereits 1993: „*We must avoid what would be only calculating routines in favour of what promotes understanding*“ ([1], S. 438). Eine Erkenntnis die in vielen weiteren Studien erneut aufzufinden: Die mathematischen Routinen werden auswendig gelernt und abgespult, aber das Verstehen scheint oftmals auszubleiben [2], [3], [4]. Dies macht sich in verschiedenen Facetten wie bspw. Begriffsverwechslungen [1], [5], [3], beim Vergleich von Argumentationen über experimentelle Handlungen mit Argumentationen während der tatsächlichen Handlungen [6] oder geführten Diskussionen an Praktikumstagen sowie in Protokollen bemerkbar.

Für eine Vermittlung des Umgangs mit Messdaten sollten Methodik und Rechenroutinen selbstverständlich nicht außer Acht gelassen werden, vielmehr fordern wir sowohl die Vermittlung der Methodik, als auch die Vermittlung tieferen Verständnisses über letztere. (Warum werden spezielle Methoden eingesetzt? Wann ist der Einsatz sinnvoll? ...) Bei der Gestaltung einer Lernumgebung für Studierende, die

auch den Aspekt des tieferen Verständnisses zu fördern sucht, stößt man unter anderem auf zwei große Probleme. Der Vortrag stellt für beide Problemstellen jeweils eine mögliche Lösung, sowie eine Applikation, die die Umsetzung erleichtert, vor:

1. Die klassische Fehlerrechnung erscheint vielen Studierenden oft wenig zugänglich.

Die klassische Fehlerrechnung beinhaltet viele Faustregeln und Blackboxes, die teilweise sogar mathematisch nicht adäquat sind. Die Darstellung der Unsicherheit von Einzelmessungen verbirgt beispielsweise deren zufallsbehafteten Charakter und findet nicht mathematisch parallel zur Darstellung statistischer Größen statt.

Eine Lösung bietet die Unsicherheitsbetrachtung nach dem internationalen *Guide to the expression of Uncertainty in Measurement* (GUM). Mit Hilfe der vorgestellten Applikation können Messunsicherheiten nach diesem ISO-Standard (ISO 98-3:2008, [7]) berechnet werden. Des Weiteren kann die Wahrscheinlichkeitsdichtefunktion (im Folgenden WDF) der Ergebnisgröße betrachtet werden, was tieferes Verständnis fördern kann.

2. Die Verbindung von Ursache und Wirkung verschiedener Einflussgrößen wird nicht (konkret) thematisiert.

Woher kommt eigentlich das, was wir „Messfehler“ nennen? Bei unserer Befragung von Studierenden zu Studienbeginn und bei Durchsicht von Protokollen werden vor allem die folgenden genannt: Reibung, Luftzug im Labor, Reaktionszeit, Luftwiderstand, Luftfeuchtigkeit. Welche der Größen beeinflusst denn nun mein Ergebnis und vor allem inwiefern tut sie das? Diese Verbindungen werden häufig nicht explizit betrachtet, sondern von theoretischer Perspektive aus als gegeben vorausgesetzt. Um ein tiefes Verständnis zu vermitteln scheint dies aber nicht auszureichen. Eine Lösung kann die systematische Betrachtung möglicher Einflussgrößen eines konkreten Experiments bieten.

Die vorgestellte Applikation ermöglicht eben diese systematische Auseinandersetzung: Ein konkretes Experiment (g -Bestimmung beim freien Fall) wird realitätsnah simuliert und die einzelnen Einflussfaktoren können systematisch variiert werden.

Beide Applikationen und deren Anwendung beziehen sich auf die Anwendung des internationalen Leitfadens GUM, daher wird dieser im folgenden Abschnitt näher beschrieben und erläutert, bevor auf die spezifischen Applikationen eingegangen wird.

2. Messunsicherheitsbetrachtung nach GUM

Bei dem *Guide to the Expression of Uncertainty in Measurement* (GUM) handelt sich um einen festgelegten Standard bezüglich des Umgangs mit Messdaten und deren Unsicherheiten, der eine fachlich moderne Modellierung von Messunsicherheit ermöglicht. In Reaktion auf zunehmende Kritik an der herkömmlichen Fehlerrechnung und deren Schwächen in Wissenschaft und Wirtschaft wurde ein internationales Komitee, das *Joint Committee for Guides in Metrology* (JCGM) seitens des *internationalen Büros für Maß und Gewicht* (BIPM) gegründet, um diesen Leitfaden zu entwickeln. Unter anderem trug dazu von deutscher Seite auch die *Physikalisch-Technische Bundesanstalt* (PTB) bei [7], [8]. Trotz internationalem Standard ist die Unsicherheitsbetrachtung nach GUM in vielen Lehrwerken und damit auch Universitäten noch nicht angekommen [4] [9].

Verbesserungen, die von GUM vorgeschlagen werden betreffen hauptsächlich, aber nicht ausschließlich...

- die Schaffung klarer Begrifflichkeiten (bspw. Abweichung und Unsicherheit),
- eine höhere Anschaulichkeit verwendeter Modellierungen und
- das Ausbessern von Schwachstellen des traditionellen Ansatzes der Fehlerrechnung.

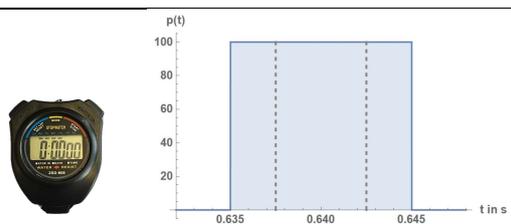
Bezüglich des letzten Punktes ist insbesondere die Verrechnung von statistischen mit nicht-statistischen Einflussgrößen zu nennen. Diese Unterteilung definiert bereits die Aufteilung, die GUM für Einflussgrößen vorschlägt:

Typ A: Solche Größen, die statistischer Natur sind und hinreichend durch die Normalverteilung modelliert werden können.

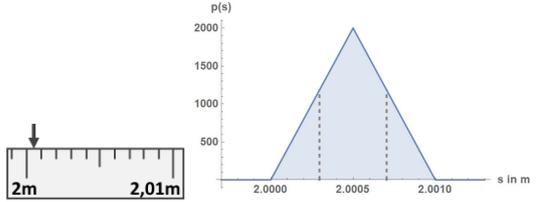
Typ B: Größen, bei denen dies nicht der Fall ist (beispielsweise Einzelmessungen, Geräteungenauigkeiten, usw.).

In der herkömmlichen Fehlerrechnung werden diese beiden Typen nicht analog modelliert: Im Falle von Typ A wird eine WDF in Form einer Normalverteilung mit einem festgelegten Vertrauensintervall betrachtet, das einer Überdeckungswahrscheinlichkeit von z.B. 68% oder 95% beschreibt. Im Fall von Typ B werden („herkömmlich“) schlichte Intervalle betrachtet ohne weitere Informationen über die Wahrscheinlichkeitsverteilung der Messwerte innerhalb des Intervalls einfließen zu lassen. Es kann daher bspw. kein Vertrauensintervall für Typ B gebildet werden.

Die Neuerung von GUM ist nun folgende: Auch Typ B wird in Form von WDFs, wie zum Beispiel Rechteck- oder Dreieckfunktionen modelliert (s. Bsp. 1 und 2 für beispielhafte Darstellung der Anzeigeunsicherheiten – die Unsicherheit durch die begrenzte Auflösung der Skala der Anzeige - des Eingangsbeispiels). Es wird also nicht nur das Intervall der Einflussgröße, sondern auch (auf der y-Achse aufgetragen) die jeweilige Wahrscheinlichkeit der Werte innerhalb des Intervalls betrachtet. Somit werden mehr Informationen der Einflussgröße in die Berechnung mit einbezogen und es resultiert eine mathematisch analoge Beschreibung zu Typ A. Es können nun ebenso Vertrauensintervalle gebildet werden und das Vertrauensniveau der Ergebnisgröße ist nach Verrechnung bekannt. Eine große Problemstelle dieses Ansatzes liegt darin, dass nach der Betrachtung aller Eingangsgrößen in Form von WDFs, die Berechnung und damit Angabe der Ergebnisgröße wiederum in Form eines Intervalls erfolgt. Dies ist zweckmäßig, da eine Berechnung der Ergebnis-WDF mathematisch sehr kompliziert wird (eine Faltung der jeweiligen Funktionen unter Berücksichtigung des jeweiligen physikalischen Zusammenhangs wäre erforderlich). Jedoch gehen viele der Informationen, die in die Eingangsgröße einbezogen wurden durch diese Reduktion verloren. Ein konkretes Beispiel hierzu wird weiter unten genannt.

Art der Anzeige: Digital
Intervall: 0,635 s bis (ausschließend) 0,645 s
 <p>The image shows a digital stopwatch on the left and a graph on the right. The graph plots probability density p(t) on the y-axis (ranging from 0 to 100) against time t in seconds on the x-axis (ranging from 0.635 to 0.645). A blue rectangular area represents a uniform distribution between 0.635 and 0.645 s.</p>
Erklärung: Die Stoppuhr hat eine Digitalanzeige, Sekunden können daher bis auf zwei Nachkommastellen angegeben werden. Ein angegebener Wert von bspw. 0,64 s beschreibt also ein Intervall von 0,635s bis ausschließend 0,645s. Die Wahrscheinlichkeit, der einzelnen Werte in diesem Intervall ist in jedem Fall gleich: Die Einflussgröße kann daher sinnvoll durch eine Rechteckverteilung beschrieben werden.
WDF: Rechteckfunktion mit Breite $a = 0,01s$
Vertrauensintervall: $u = \frac{a}{2\sqrt{3}}$

Bsp. 1: Anzeigeunsicherheit der Stoppuhr

Art der Anzeige: Analog
Intervall: 2,000 m bis 2,001 m
 <p>The image shows an analog scale on the left and a graph on the right. The graph plots probability density p(s) on the y-axis (ranging from 0 to 2000) against distance s in meters on the x-axis (ranging from 2.0000 to 2.0010). A blue triangular area represents a triangular distribution centered at 2.0005 m.</p>
Erklärung: Hier kann mehr über die Verteilung der Werte im Intervall ausgesagt werden. Die Randwerte sind unwahrscheinlicher als die mittig liegenden. Eine Modellierung kann dementsprechend durch eine Dreiecksfunktion erfolgen (auch andere Funktionen wie eine Trapezfunktion sind denkbar).
WDF: Dreieckfunktion mit Breite $a = 1mm$
Vertrauensintervall: $u = \frac{a}{2\sqrt{6}}$

Bsp. 2: Anzeigeunsicherheit des Maßbandes

3. Zur ersten Applikation

Die von uns entwickelte App hingegen bezieht alle Informationen der Eingangs-WDFs mit ein. Die Webapp erlaubt die Eingabe der Eingangsgrößen des Typs B durch WDFs, sowie Mehrfachmessungen nach Typ A, berechnet die Unsicherheitsanalyse nach GUM und bestimmt in einem letzten Schritt die WDF der Ergebnisgröße mithilfe einer Monte-Carlo-Simulation (s. Abbildung 1 für Beispiel der Benutzeroberfläche). Somit wird nicht nur die Berechnung der resultierenden Messunsicherheit nach GUM erleichtert, es werden zusätzlich die bei der Reduktion auf Intervalle verlorenen Informationen sichtbar gemacht. Die folgende Beispielanalyse des Eingangsbeispiels demonstriert diesen Sachverhalt.

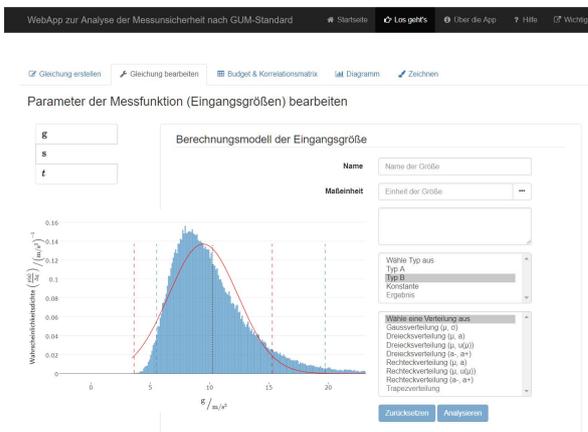


Abb. 1: Beispielhafte Darstellung der Benutzeroberfläche der ersten Applikation

4. Eine Beispielanalyse

Als Beispiel wird dieser Vorgang nun anhand des Eingangsexperimentes durchgeführt. Der Einfachheit halber werden dabei nur zwei Einflüsse bedacht: Die Anzeigeunsicherheit des Maßbandes (Dreieckverteilung wie in Bsp. 2) und die Reaktionszeit (approximiert durch eine Normalverteilung mit konservativ geschätzter Breite von $\sigma = 0,1s$). Werden beide WDFs in die Webapp eingetragen, so ergibt sich die in Abbildung 2 gezeigte Ergebnis-WDF.

Die Betrachtung dieses Beispiel zeigt bereits sehr deutlich, inwiefern die Betrachtung der Ergebnis-WDF mehr Informationen erschließt als die reine Betrachtung der berechneten Intervalle. Bei einer reinen Betrachtung der Intervalle wäre an dieser Stelle von einem symmetrischen, normalverteilten Ergebnis ausgegangen worden (in der Abbildung rot). Unter anderem wird hier aber sichtbar, dass trotz überwiegender normalverteilter Eingangsgröße (die Längenunsicherheit spielt auf Grund der geringen Größe so gut wie keine Rolle) eine deutlich nicht-normalverteilte Ergebnisverteilung entsteht. Auch die Wirksamkeit des Mittelwertes als Bestwert des Ergebnisses kann hier somit in Frage gestellt und diskutiert werden. Es wird ein kritisches aber sinnvolles Hinterfragen der herkömmlich für gegeben angenommenen Methoden ermöglicht, das Raum für die Förderung tieferen Verständnisses schafft.

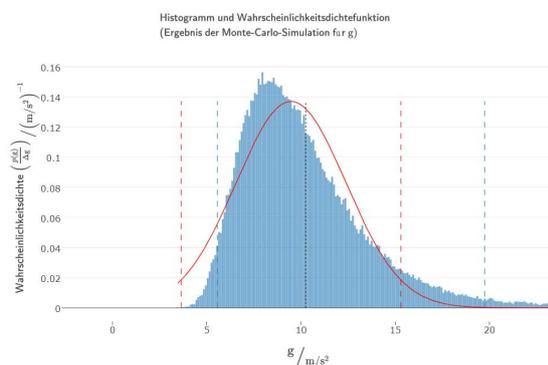


Abb. 2: Ergebnis-WDF des Ortsfaktors g (blau) mit nach Fehlerfortpflanzung bestimmter Normalverteilung (rot)

5. Vom Konkreten zum Allgemeinen

Während die Wirkung spezifischer Einflussfaktoren und gerade der statistische Charakter der Einflüsse von vielen Studierenden nicht ausreichend verstanden wird, scheint es vielen recht einfach zu fallen mögliche Einflussgrößen auf konkrete Experimente zu nennen. Ein tieferes Eingehen auf die konkrete Einflussnahme der einzelnen Faktoren bleibt aber in der Regel aus. Wir plädieren daher für eine tiefgehende, systematische Auseinandersetzung mit

exemplarischen Experimenten und dessen konkreten Einflussgrößen.

Gerade für die Lehramtsausbildung bietet sich dies an, da sich Störeinflüsse in typischen Schulexperimenten je nach Inhaltfeld oft wiederholen. Beispielsweise ziehen sich Störeinflüsse wie Reibung, Reaktionszeit, Luftwiderstand, Unsicherheit von Zeit-/ Längenmessungen und andere durch große Teile der Mechanik. Die tatsächliche Auswirkung dieser Einflüsse auf das Messergebnis abzuschätzen, fällt den Studierenden allerdings nicht leicht.

Gefordert ist also die systematische Variation zum Ziel der näheren Untersuchung einzelner Einflussfaktoren auf ein spezifisches Experiment. Dabei ergeben sich allerdings zwei große Probleme:

1. Die systematische Variation der Einflüsse kann zu sehr komplexen Aufbauten führen und sogar unmöglich sein.
2. Um die tatsächliche Auswirkung auf das Ergebnis zu finden müssen für statistische Einflüsse pro Variation sehr viele Messungen durchgeführt werden.

Dies ist somit als Realexperiment nicht denkbar, kann allerdings durch eine realitätsnahe Simulation ermöglicht werden. Hier knüpft die zweite erstellte Applikation an.

5.1. Zur zweiten Applikation

Die Applikation simuliert ein konkretes Experiment (das Eingangsexperiment) möglichst realitätsnah und ermöglicht eine direkte Ausgabe des Ergebnisses in Form einer WDF (s. Abbildung 1). Was bedeutet in diesem Fall „realitätsnah“? Der freie Fall wird reibungsbehaftet simuliert und sämtliche einstellbaren Eingangsgrößen (Anzeigeunsicherheiten, Reaktionszeit, Geräteunsicherheiten, Masse, Volumen, ...) werden durch WDFs im Programm simuliert, dass anhand dieser WDFs zu jeder erneuten Messung zufallsbasiert „misst“. Werden viele tausend Messungen (im Beispiel 10^5) auf einmal ausgeführt, so erhält man die Ergebnis-WDF (vgl. Monte-Carlo-Simulation). In Abbildung 1 ist der bereits beispielhaft diskutierte Fall eingestellt. Die einstellbaren Größen können durch den jeweiligen Slider (auf Abbildung 1 rechts) systematisch variiert werden. Ebenso können auch hier Standardunsicherheit und Normalverteilung über die Ergebnisfunktion gelegt werden.

- dents' perceptions of the quality of experimental measurements. In: *International Journal of Science Education* 20 (4), S. 447–459.
- [6] Heinicke, S.; Riess, F. (2012): Missing Links in the Experimental Work. Student's Actions and Reasoning on Measurement and Uncertainty. In: L. Maurines & A. Redfors (Hg.): *ESERA 2011 Proceedings. Nature of Science, History, Philosophy, Sociology of Science, Part 5*, S.52-56
- [7] Internationale Organisation für Normung, *ISO/IEC Guide 98-3:2008 (2008): Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement*. ISO, Genf.
- [8] Deutsches Institut für Normung (1995): *Leitfaden zur Angabe der Unsicherheit beim Messen*. Beuth Verlag, Berlin.
- [9] Heinicke, Susanne; Glomski, Julia; Priemer, Burkhard; Rieß, Falk (2010): Aus Fehlern wird man klug - Über die Relevanz eines adäquaten Verständnisses von "Messfehlern" im Physikunterricht. In: *Praxis der Naturwissenschaften - Physik in der Schule* 59 (5), S. 26–33.

Dieses Projekt wird im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern aus Mitteln des Bundesministeriums für Bildung und Forschung gefördert.