

Analyse studentischer Fehlvorstellungen mittels des Force Concept Inventory Item Response Curves im internationalen Vergleich

Silke Stanzel

TH Rosenheim, Hochschulstraße 1, 83024 Rosenheim
silke.stanzel@th-rosenheim.de

Kurzfassung

Das Force Concept Inventory ist ein international etabliertes Diagnoseinstrument für das Konzeptverständnis der Newtonschen Mechanik. Wir haben über einen Zeitraum von zehn Jahren von knapp 5000 Studierenden der Ingenieurwissenschaften an der TH Rosenheim zu Studienbeginn Testdaten erhoben. Die Auswertung mittels Item Response Curves erlaubt für jede Frage die Analyse sowohl der richtigen als auch aller falschen Antwortoptionen. Insbesondere lassen sich auf diese Weise Distraktoren identifizieren, die gängige Fehlvorstellungen widerspiegeln. Es zeigt sich, dass die am häufigsten gewählten Distraktoren auf der Annahme beruhen, jeder Bewegung müsse eine Kraft in Bewegungsrichtung zu Grunde liegen. Daten von Universitäten der USA zeigen ein vergleichbares Ergebnis. Die vorgestellte Analyse dient als Grundlage zur Verbesserung von Lehrsequenzen.

1. Motivation und Ausgangslage

Die physikalischen Vorkenntnisse Studierender zu Beginn des Hochschulstudiums sind Gegenstand diverser Untersuchungen. Dabei steht in vielen Fällen ein generelles Niveau der Physikkenntnisse im Fokus, z.B. um Unterschiede in den Hochschulzugangsberechtigungen aufzuzeigen [1] oder Korrelationen zum Studienerfolg herzustellen [2]. Aus Sicht der Lehrenden ist das Interesse an dem Vorwissen auch deshalb hoch, weil die Physiklehrveranstaltungen an den Hochschulen darauf aufbauen und das Erreichen der Lernziele je nach Stand der Vorkenntnisse mehr oder weniger Schwierigkeiten bereiten wird. Lehrende, die den Kenntnisstand der Studierenden kennen, können ihre Lehre entsprechend anpassen und damit das Erlernen physikalischer Konzepte erleichtern.

Dabei ist erwiesen, dass es nicht ausreicht, den Stand in Bezug auf korrekte physikalische Kenntnisse zu erheben, sondern dass insbesondere berücksichtigt werden muss, welche falschen Vorstellungen von physikalischen Zusammenhängen bei den Lernenden vorherrschen. Diese müssen zunächst in Frage gestellt und überwunden werden bevor das physikalisch richtige Konzeptverständnis aufgebaut werden kann [3].

Gegenstand der vorliegenden empirischen Untersuchung ist, das Vorwissen in der Newtonschen Mechanik von Studierenden des ersten Semesters in den Ingenieurwissenschaften an der TH Rosenheim auf

vorherrschende Fehlvorstellungen hin zu analysieren. Damit sollen Lehrende in die Lage versetzt werden, ihre Lehrveranstaltungen z.B. durch die Wahl geeigneten Studiermaterials so weiterzuentwickeln, dass Studierende ihre Kompetenzen in diesem Bereich der Physik effektiv verbessern können.

2. Methodik

2.1. Force Concept Inventory (FCI)

Für die Erhebung der Vorkenntnisse im Themenfeld der Newtonschen Mechanik hat sich das Force Concept Inventory (FCI) als Testinstrument weltweit etabliert. Im Folgenden wird auf Inhalt, Testqualität und ein häufiges Anwendungsfeld dieses Konzepttests eingegangen.

2.1.1. Inhalt des FCI

Das FCI wurde von Hestenes, Wells und Swackhammer in den USA entwickelt und 1992 veröffentlicht [4]. Die überarbeitete Version von 1995 ist mittlerweile in 30 Sprachen verfügbar und liegt unserer Studie zugrunde [5]. Dieser Konzeptverständnistest besteht aus 30 Single Choice Fragen zu Kinematik, Kräften und Dynamik. In den Testfragen werden Bewegungen unterschiedlicher Körper beschrieben und z.T. mit graphischen Darstellungen veranschaulicht (s. Abb. 1). Gefragt wird z.B. nach Bahnkurven, Geschwindigkeiten oder zugrundeliegenden Kräften.

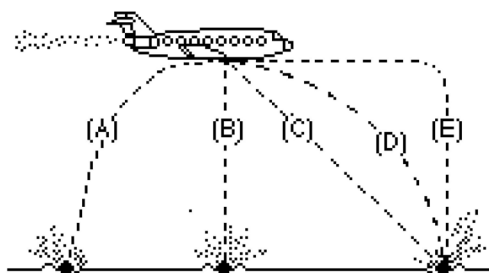


Abb.1: Graphik zur Frage 14 des FCI; gefragt ist nach der Bahnkurve einer Kugel, die aus einem Flugzeug fällt und von einem am Boden stehenden Beobachter gesehen wird [5].

Von den jeweils zur Wahl stehenden fünf Antwortoptionen beruhen die vier Distraktoren überwiegend auf wissenschaftlich erwiesenermaßen häufig vertretenen falschen Vorstellungen [4]. Diese intuitiven Konzepte gehen z.B. zurück auf Alltagsbeobachtungen oder auf ein unspezifisch breites Bedeutungsspektrum der Worte in der Alltagssprache, die gleichzeitig in der Physik für wohldefinierte Größen verwendet werden. „Kraft“ ist ein Beispiel dafür. In diesem Artikel wird der Begriff „Fehlvorstellungen“ verwendet. In der Literatur finden sich daneben auch Begriffe wie „Alltagsvorstellungen“ oder „Schülervorstellungen“ [3], [6].

2.1.2. Testqualität des FCI und Analysen des Antwortverhaltens

Der FCI wird seit 30 Jahren an unterschiedlichen Bildungseinrichtungen weltweit eingesetzt. Dementsprechend umfangreich sind die erhobenen Datensätze und die Erfahrungen mit dem Instrument. Die Überprüfung der Test-Qualitätsmerkmale wie Validität, Reliabilität, Leichtigkeitsindex und Trennschärfe ist Gegenstand etlicher Studien [7], [8]. Für die Analyse von FCI – Ergebnissen ist es wichtig zu wissen, dass die Fragen Nr. 6, 7 und 16 als nicht valide identifiziert wurden [9]. Das bedeutet, dass die richtigen Antworten auf diese Fragen zu einem hohen Anteil auf der Basis eines falschen physikalischen Verständnisses ausgewählt werden.

Ferner wurden Analysen durchgeführt, um Korrelationen der Antworten zu identifizieren und sowohl die richtigen als auch die falschen Antwortoptionen thematischen Gruppen zuordnen zu können [10], [11], [12]. Mit Hilfe sogenannter Item Response Curves können die Distraktoren weiter klassifiziert werden [13], [14] und Modelle für studentische Erklärungsmuster aufgestellt werden [15]. Diese Darstellungsform wird in Abschnitt 3.2 beschrieben.

2.1.3. Bewertung von Lehrformen mittels FCI

In vielen Fällen wird der FCI verwendet, um die Effektivität verschiedener Lehrformate bzw. Lehrinterventionen quantitativ zu erfassen. Beispielsweise wurden in der Meta-Studie von R. R. Hake [16] FCI Ergebnisse von interaktiven Lehrformaten mit denen konventioneller Formate verglichen und damit die Wirksamkeit aktivierender Lehre nachgewiesen.

Daran anknüpfend werden an der TH Rosenheim seit zehn Jahren begleitend zur Umstellung didaktischer Methoden ebenfalls FCI – Daten vor und nach den Physikkursen erhoben. Die Analyse dieser Daten bestätigt die höhere Effektivität aktivierender Lehre [17], [18].

2.2. Datenbasis

An der TH Rosenheim wird der FCI seit dem Wintersemester 2013/14 kontinuierlich in allen zwölf ingenieurwissenschaftlichen Studiengängen von den Studierenden durchgeführt. Für die sogenannten „Vortests“ zu Studienbeginn wird ihnen in den ersten Wochen des ersten Semesters im Rahmen der Physiklehrveranstaltung 30 min Bearbeitungszeit in Präsenz gegeben. Im Verlauf der zehn Jahre von WS2013/14 bis WS2022/23 sind so 4957 Testergebnisse entstanden, die dieser Studie zugrunde liegen. 46% der Teilnehmenden haben ein allgemeinbildendes Abitur erworben, 28% kommen vom Technikzweig einer Fach- bzw. Berufsoberschule, 15% von einem anderen Zweig dieser Schulform (z.B. Wirtschaft oder Soziales) und 4% haben sich beruflich für das Ingenieurstudium qualifiziert.

Die Verteilung der im Test erreichten Punkte ist in Abbildung 2 dargestellt. Jede richtig beantwortete Frage ergibt einen Punkt, falsche Antworten werden nicht eingerechnet. Im Mittel über alle 4957 Tests sind 12,1 Punkte von maximal 30 Punkten erreicht worden. Nicht alle Testpersonen wählen bei jeder Frage eine Antwort aus. Im Schnitt sind 4% der Fragen unbeantwortet gelassen. Der Anteil steigt vermutlich aus Zeitgründen zu Ende des Fragebogens an.

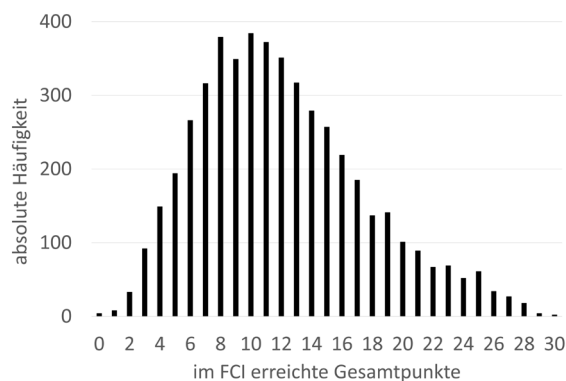


Abb.2: Verteilung der im FCI von maximal 30 möglichen Punkten erreichten Punktzahl für die 4957 zu Studienbeginn von WS2013/14 bis WS2022/23 durchgeführten Tests an der TH Rosenheim.

3. Studentische Fehlvorstellungen

3.1. Analyse der richtigen Antworten

Für ein Verständnis der Schwierigkeiten, die Studierende im Themenfeld der Newtonschen Mechanik mitbringen, wird zunächst für jede Frage im FCI die Anzahl richtiger Antworten ermittelt. Abbildung 3 zeigt in Form eines Paretos die relative Häufigkeit nach Größe sortiert.

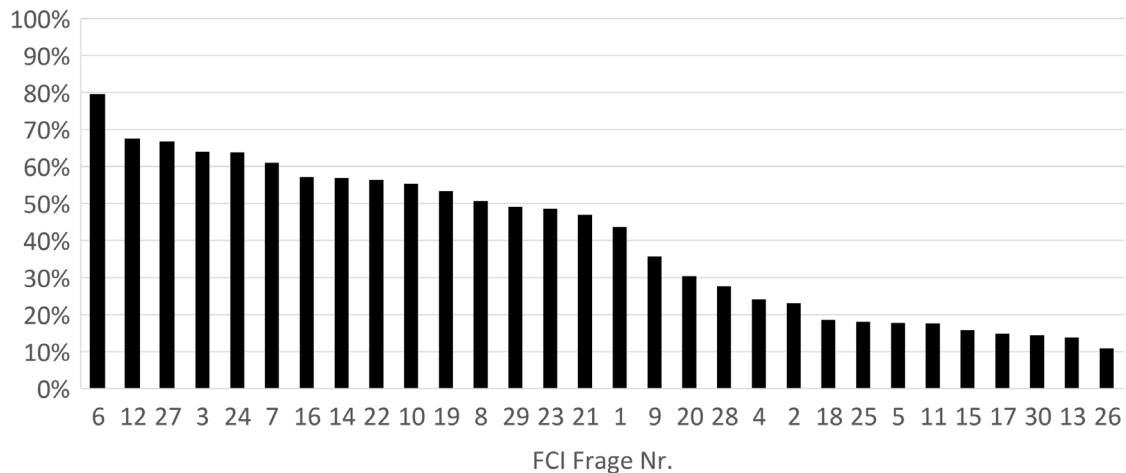


Abb.3: Pareto – Darstellung der Häufigkeit, mit der die richtigen Antworten im FCI von den 4957 Studierenden an der TH Rosenheim zu Studienbeginn gewählt wurden. Die Nummern entsprechen der Nummerierung der Fragen.

Die neun Fragen, die zu weniger als 20% richtig beantwortet wurden, lassen sich thematisch zusammenfassen. Dafür verwenden wir die Klassifizierung, die von P. Eaton und S. D. Willoughby anhand von Faktoranalysen und inhaltlicher Bewertung der Fragen aufgestellt wurde (Tabelle III in [11]). Demnach bilden fünf dieser Fragen die Gruppe der „force identification“ (Identifikation der wirkenden Kräfte, Fragen 5, 11, 13, 18, 30) und drei die der „mixed concepts“ (Fragen 17, 25, 26). Die verbleibende Frage Nr. 15 bezieht sich auf das dritte Newtonsche Axiom im Kontext beschleunigter Bewegung. Bei den Fragen aus der erstgenannten Gruppe geht es um die Identifikation von Kräften auf bewegte Körper. Zur richtigen Beantwortung der „mixed concepts“-Fragen müssen mehrere Konzepte wie Kräfteaddition, Reibungskräfte und Newtonsche Axiome zusammengenommen verstanden sein.

Aus der Darstellung der Häufigkeit der richtigen Antworten allein lässt sich noch nicht auf möglicherweise vorliegende Fehlvorstellungen schließen. Dafür wird zusätzlich ausgewertet, welche Distraktoren zu welchem Anteil gewählt werden.

3.2. Analyse der Distraktoren mittels Item Response Curves

Die Item Response Curves (IRC) sind eine besonders informative Darstellung der in einem Multiple Choice Test gewählten Antworten [13], [14]. Sie zeigen für jede Frage alle Antwortoptionen an. Damit erlauben sie Rückschlüsse nicht nur auf Basis der richtigen Antwort sondern auch in Bezug auf die gewählten Distraktoren.

Die IRC stellen in einem Diagramm pro Frage die Häufigkeit aller gewählten Antwortoptionen als Funktion eines allgemeinen Kompetenzniveaus dar. Für den FCI hat sich die im Test insgesamt erreichte Punktzahl als ein gutes Maß für die Verständnistiefe der Newtonschen Mechanik erwiesen [7], [14]. Daher wird für die IRC des FCI die Häufigkeit jeder der fünf möglichen Antwortoptionen als Funktion

der insgesamt über alle 30 Fragen erreichten richtigen Antworten aufgetragen. In Abbildung 4 sind exemplarisch die IRC für die Frage 14 dargestellt. Zum Inhalt der Frage siehe Abbildung 1. Zum Verständnis der Konstruktion der IRC ist in Abb. 4 bei einer FCI-Gesamtpunktzahl von 10 eine gestrichelte Linie gezogen. Auf dieser ist für jede Antwortoption aufgetragen, mit welchem relativen Anteil der 384 Testergebnisse, die in Summe über den ganzen Test genau zehn richtige Antworten enthalten, sie gewählt wurde: zu 53% ist es die richtige Antwort D, zu 25% der Distraktor A, zu 14% B, zu 8% C und zu 0% E. Die Daten mit einer FCI-Gesamtpunktzahl von 0, 1 oder 29 sind nicht in die IRC-Graphen aufgenommen, da sie nur in wenigen Fällen vorkommen und damit eine hohe statistische Streuung aufweisen (vgl. Abb. 2). Der Datenpunkt mit FCI-Gesamtpunktzahl 30 zeigt bei jeder Frage die richtige Antwort an.

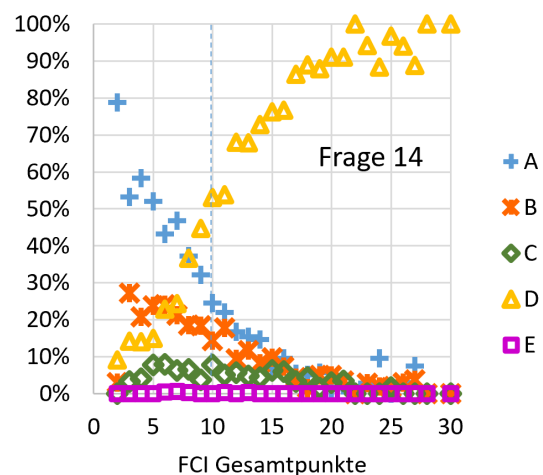


Abb.4: Item Response Curves für die Frage 14 des FCI. Aufgetragen ist die Häufigkeit der von den Studierenden jeweils gewählten Antwortoptionen A - E als Funktion der Anzahl richtiger Antworten im FCI. Die richtige Antwort auf diese Frage ist D.

Die Häufigkeit der in Frage 14 richtigen Antwort D nimmt wie zu erwarten als Funktion der Gesamtpunktzahl bis auf statistisch begründete Schwankungen monoton zu. Die Häufigkeit der vier Distraktoren nimmt entsprechend monoton ab. Dabei ist Antwort A der am häufigsten gewählte Distraktor. Antwortoption E hingegen wird kaum jemals gewählt und ist also als Distraktor nicht effektiv. Der monoton steigende Verlauf der richtigen Antwort ist in allen 30 Fragen zu sehen. Dies steht im Einklang damit, dass die FCI – Gesamtpunktzahl ein Maß für das Konzeptverständnis ist [7], [14]. Die Item Response Curves sind offensichtlich auch eine geeignete Darstellung, um Aussagen über die Qualität von Antwortoptionen in Testfragen zu treffen [13].

Der Verlauf der Distraktoren stellt sich in den vorliegenden Daten je nach Frage durchaus unterschiedlich dar. Neben dem oben beschriebenen monoton fallenden Verhalten von Distraktoren gibt es auch solche, deren Häufigkeit mit steigender FCI-Gesamtpunktzahl zunächst zunimmt, um dann nach einem Maximum oder Plateau wieder zu fallen, siehe Antwort A der Frage 17 in Abbildung 5. Sie bezieht sich auf die reibungsfreie Bewegung eines Fahrstuhls mit konstanter Geschwindigkeit nach oben. Gefragt wird nach der Größe der Seilkraft auf die Kabine im Vergleich zu anderen vorliegenden Kräften.

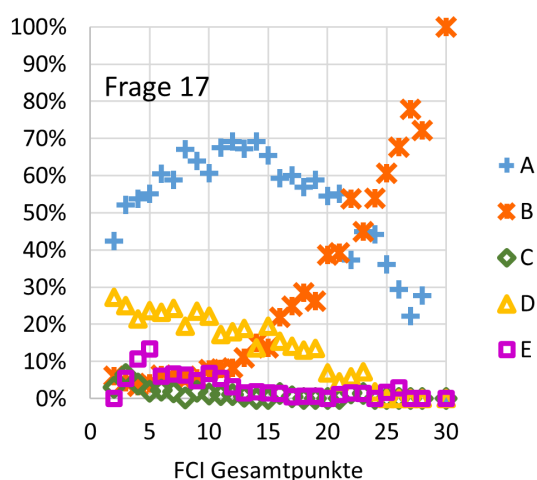


Abb.5: Item Response Curves für die Frage 17 des FCI. Aufgetragen ist die Häufigkeit der von den Studierenden jeweils gewählten Antwortoptionen A - E als Funktion der Anzahl richtiger Antworten im FCI. Die richtige Antwort auf diese Frage ist B.

Studierende mit geringem Niveau an Vorkenntnissen, also z.B. weniger als zehn Gesamtpunkten im FCI, wählen neben der Antwort A als zweithäufigstes die Antwort D. Beiden Antworten liegt das Fehlkonzepkt zugrunde, dass die Seilkraft in Bewegungsrichtung größer sein muss als die Summe aller ent-

gegengesetzt gerichteten Kräfte. Antwort D enthält im Unterschied zu A neben der Gewichtskraft noch eine physikalisch falsche Kraft des Luftdrucks. Mit zunehmender FCI – Gesamtpunktzahl nimmt die Antwort D ab, d.h. mit zunehmendem Verständnis wird diese Kraft des Luftdrucks als falsch erkannt. Dafür nimmt aber der Distraktor A an Häufigkeit zu. Die Fehlvorstellung der in Bewegungsrichtung wirkenden Kraft dominiert also weiterhin. Erst bei über 23 Gesamtpunkten setzt sich die richtige Antwort B in der Häufigkeit gegenüber dem Distraktor A durch. Generell spiegeln Antwortoptionen im FCI mit einem zunächst ansteigenden Verlauf vorhandene Fehlvorstellungen wider [14], [15].

Diese Frage Nr. 17 gehört in der Klassifizierung von P. Eaton und S.D. Willoughby [11] zur Gruppe der „mixed concepts“. In allen drei zu dieser Gruppe gehörenden Fragen (17, 25, 26) sowie den fünf Fragen (Nr. 5, 11, 13, 18, 30), welche die Gruppe der „force identification“ bilden, durchläuft der am häufigsten gewählte Distraktor in den Item Response Curves ein Maximum. Alle diese Falschantworten beruhen auf der Annahme, für jede Bewegung sei eine Kraft in Richtung der Bewegung notwendig. Dieses aus der Literatur (z.B. [3], [6]) bekannte Konzept liegt demnach bei den Studienanfängern an der TH Rosenheim als dominierende Fehlvorstellung vor.

4. Vergleich mit US-amerikanischen Daten

Die Darstellung der FCI Antworten mittels Item Response Curves (IRC) wurde 2006 erstmals von Morris et al für drei ausgewählte Fragen veröffentlicht [13]. In einer nachfolgenden Veröffentlichung sind die IRC für alle 30 Fragen des FCI dargestellt [14]. Die zugrundeliegende Datenbasis stammt von den US-amerikanischen Universitäten Harvard, Mississippi State und Rice. Aus den Veröffentlichungen wird nicht ersichtlich, zu welchem Zeitpunkt im Laufe des Studiums die Daten genommen wurden. Der Datensatz umfasst über 4500 studentische Antworten und ist damit von vergleichbarem Umfang wie die hier vorgestellten Daten der TH Rosenheim. Im Mittel werden bei den US-amerikanischen Fragebögen 12,6 von 30 möglichen richtigen Antworten erzielt. Bei den Rosenheimer Daten beträgt dieser Mittelwert 12,1 (s. Abschnitt 2.2).

In Abbildung 6 sind exemplarisch die IRC für die FCI-Fragen 13 – 18 der beiden Datensätze im Vergleich dargestellt. Die Symbole in Schwarz stammen aus Abbildung 1 der Publikation [14]. In Farbe sind mit der gleichen Symbolzuordnung die Häufigkeiten der Antworten A (+), B (*), C (◊), D (Δ) bzw. E (□) für die Daten der TH Rosenheim hinzugefügt. Die jeweils richtige Antwort lässt sich identifizieren als diejenige, welche bei der maximalen Gesamtpunktzahl 30 zu 100% gewählt wurde.

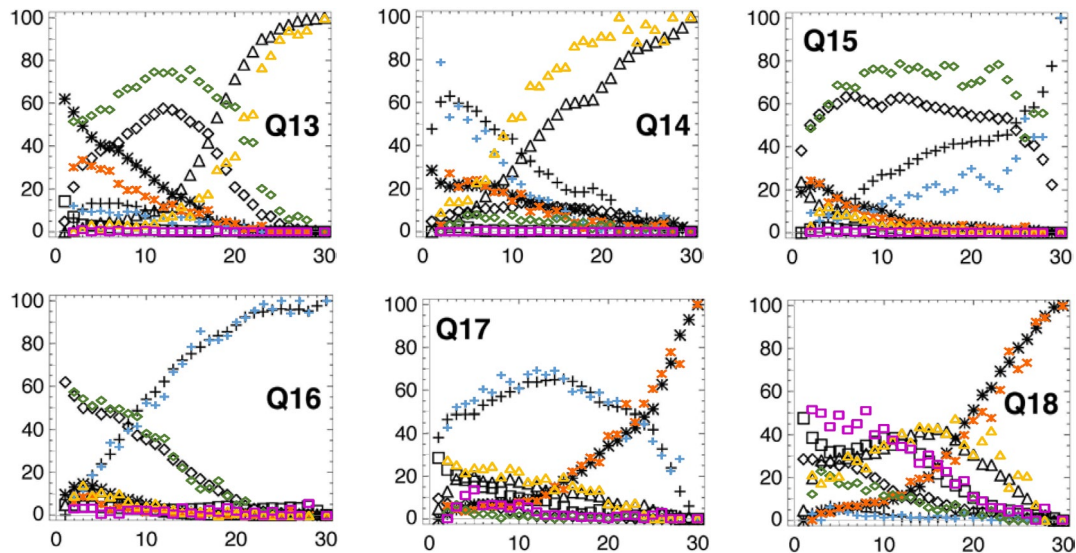


Abb.6: Item Response Curves für die Fragen (Q) 13 – 18 des FCI: in Schwarz die Daten der >4500 studentischen Antworten von US-amerikanischen Universitäten aus Abb. 1 in [14], in Farbe die Daten der Erstsemester-Studierenden an der TH Rosenheim; die relative Häufigkeit der jeweiligen Antwort A (+), B (*), C (◇), D (Δ) bzw. E (□) (von 0 bis 100%) ist aufgetragen als Funktion der Anzahl der insgesamt im FCI erreichten richtigen Antworten (von 0 bis 30).

Der qualitative Verlauf der IRC dieser sechs Fragen stimmt für beide Datensätze gut bis sehr gut überein. Insbesondere bei den Fragen 16 und 17 liegen die IRC der TH Rosenheim auf denen der US-amerikanischen Daten. Bei Frage 18 gibt es leichte Unterschiede in der Verteilung der Häufigkeit der Falschantworten im Bereich niedriger FCI-Gesamtpunktzahl (< 10). Bei den Fragen 13, 14 und 15 sind die quantitativen Unterschiede in den Antworthäufigkeiten größer, auch wenn der qualitative Verlauf sehr ähnlich ist. Auch die übrigen 24 Fragen des FCI zeigen im Vergleich der beiden Datensätze ähnliche bis keine Unterschiede. Wir können damit die in [14] gezeigten IRC zumindest qualitativ bestätigen.

Die große Ähnlichkeit des Verlaufs der IRC für die FCI-Daten von Erstsemester-Studierenden der TH Rosenheim im Vergleich zu den US-Studierenden bestätigt die Qualität der FCI-Fragen und Antwortoptionen. Die relative Häufigkeit der gewählten Antworten einer Frage ist offenbar nur abhängig von der erreichten Gesamtpunktzahl im FCI und unabhängig von dem spezifischen zu Grunde liegenden Datensatz. Damit besitzt bereits die insgesamt erreichte Anzahl richtiger Antworten im FCI eine hohe Aussagekraft in Bezug auf das Verständnis der hinter den einzelnen Fragen stehenden physikalischen Konzepte. Des Weiteren zeigt sich in diesem Vergleich, dass zumindest für die betrachteten Gruppen von Studierenden in USA und Deutschland Unterschiede z.B. im Bildungssystem keinen Einfluss auf die Ausbildung von Fehlvorstellungen und den prinzipiellen Verlauf des Übergangs zu physikalisch richtigen Konzepten haben.

Die Item Response Curves geben nur die relativen Häufigkeiten der Antwortoptionen als Funktion der im FCI erreichten Punkte wieder. Sie machen damit keine Aussage über die dem jeweiligen Datensatz zugrunde liegende Verteilung der insgesamt im Test erreichten richtigen Antworten. Im vorliegenden Fall weisen die beiden Datensätze zufällig einen nahezu gleichen Mittelwert im FCI-Gesamtergebnis auf (12,6 vs. 12,1). In Kombination mit dem ähnlichen Verlauf der Item Response Curves könnte vermutet werden, dass auch die relative Häufigkeit der richtigen Antworten jeder Frage vergleichbar ist. In Abbildung 7 ist die Häufigkeit der richtigen Antworten für die US-amerikanischen Daten in Form eines Pareto dargestellt. Dies lässt sich mit den Daten der TH Rosenheim in Abbildung 3 vergleichen. Je nach Frage ist der Anteil richtiger Antworten für beide Datensätze durchaus unterschiedlich. Für die Fragen 13 – 18, deren IRC in Abbildung 6 dargestellt ist, ist der Anteil richtiger Antworten für beide Datensätze zum besseren Vergleich in Tabelle 1 gelistet. Die Unterschiede sind insbesondere gering in den Fragen 16, 17 und 18, für die auch der Verlauf der richtigen Antwort im IRC keinen wesentlichen Unterschied aufzeigt.

Tab.1: Häufigkeit der richtigen Antworten auf die FCI – Fragen 13 – 18 für den Datensatz der TH Rosenheim (TH Ro) im Vergleich zu den Daten aus [14] (USA).

Frage	13	14	15	16	17	18
TH Ro	14%	57%	16%	57%	15%	19%
USA	26%	40%	29%	59%	18%	22%

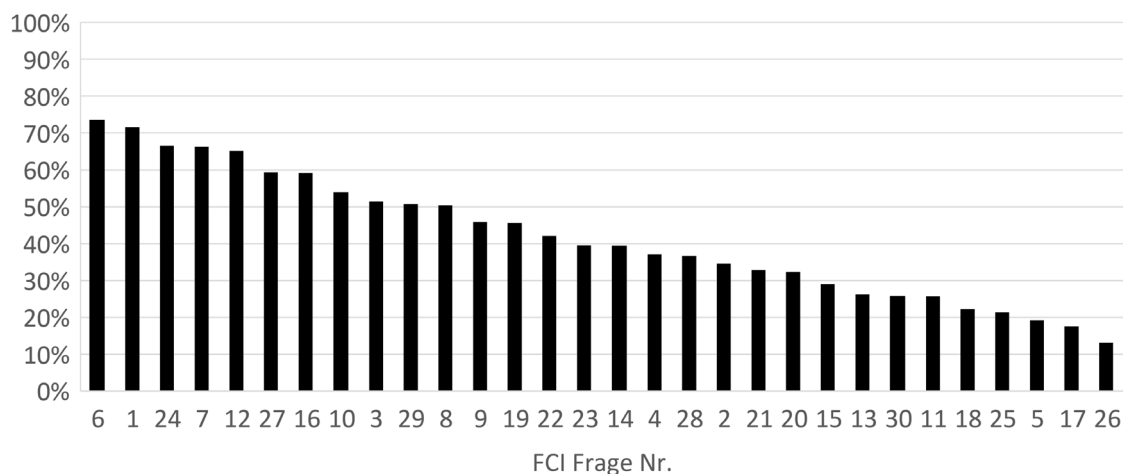


Abb.7: Pareto – Darstellung der Häufigkeit, mit der die richtigen Antworten im FCI von den über 4500 Studierenden der Universitäten Harvard, Mississippi State und Rice gewählt wurden. Die Nummern entsprechen der Nummerierung der Fragen. Die Darstellung beruht auf Tabelle 1 in [14].

Auffallend beim Vergleich der Pareto-Darstellungen (Abb. 3 und 7) ist, dass in beiden Datensätzen die neun Fragen, die am schlechtesten ausfallen, die gleichen sind, nämlich Fragen Nr. 5, 11, 13, 15, 17, 18, 25, 26, 30. In Abschnitt 3.1 ist die thematische Klassifizierung dieser Fragen angegeben. Damit besteht offenbar auch bei den Studierenden der untersuchten Gruppen in den USA im FCI das dominierende Fehlkonzent in der Annahme, für jede Bewegung sei eine Kraft in Richtung der Bewegung nötig.

5. Fazit und Ausblick

Die Analyse der FCI Daten von knapp 5000 Studierenden zu Beginn ihres ingenieurwissenschaftlichen Studiums an der TH Rosenheim ergibt, dass die Zuordnung von Kräften, die auf bewegte Körper wirken, die größte Schwierigkeit im Verständnis der Newtonschen Mechanik darstellt. Eine detaillierte Untersuchung der im Test gewählten Antworten zeigt, dass selbst Studierende mit einem insgesamt guten Grundverständnis davon ausgehen, dass zu jeder Bewegung eine Kraft in Richtung dieser Bewegung vorhanden sein müsse. Damit werden Daten, die von Studierenden in den USA erhoben wurden, bestätigt.

In der Konsequenz bedeutet das für Lehrende, dass geeignete Aufgabenstellungen und Materialien eingesetzt werden müssen, um dieses Fehlkonzent effektiv zu überwinden und durch das richtige physikalische Verständnis zu ersetzen. In den USA sind im Rahmen der Physics Education Research eine Fülle von forschungsbasierten Arbeitsblättern, Simulationen und Problemstellungen entwickelt worden, die zielgerichtet die Schwierigkeiten im Aufbau der richtigen physikalischen Konzepte adressieren. Exemplarisch seien die Tutorials von L.C. McDermott [19], die Materialien von R.D. Knight [20] und die PhET-Simulationen der University of Colorado [21] genannt.

In einigen Physiklehrveranstaltungen der TH Rosenheim werden in den laufenden Semestern auf Basis dieser Studie Lehrmaterial und Lehrinterventionen gezielt modifiziert. Der FCI wird am Semesterende nochmal durchgeführt. Daraus sollte in einer nachfolgenden Analyse die Wirksamkeit der gewählten Aktivitäten zu erkennen sein.

6. Literatur

- [1] Kurz, G.; Käß, H. (2019): Physikkenntnisse von Studienanfängern des Maschinenbaus – eine Fallstudie an der Hochschule Esslingen. In: PhyDid B, Didaktik der Physik, Beiträge zur DPG-Frühjahrstagung Aachen (2019), S. 189 - 195
- [2] Binder, T.; Sandmann, A.; Sures, B.; Friege, G.; Theyssen, H.; Schmiemann, P. (2019): Assessing prior knowledge types as predictors of academic achievement in the introductory phase of biology and physics study programmes using logistic regression. International Journal of STEM Education (2019) 6:33.
- [3] Kautz, C. (2014): Verständnisschwierigkeiten und Fehlvorstellungen in Grundlagenfächern des ingenieurwissenschaftlichen Studiums. In: Rentschler, M.; Metzger, G. (Hrsg.) Perspektiven angewandter Hochschuldidaktik. Studien und Erfahrungsberichte. Shaker Verlag 2014, S. 81 - 112
- [4] Hestenes, D.; Wells, M.; Swackhamer, G. (1992): Force concept inventory. The physics teacher 30, 141 (1992)
- [5] Das Force Concept Inventory in der Version von 1995 kann nach Registrierung heruntergeladen werden von: <https://www.physport.org> (Stand 05/2023)
- [6] Schecker, H.; Wilhelm, T.; Hopf, M.; Duit, R. (2018): Schülervorstellungen und Physikunterricht. Springer Verlag (2018)

- [7] Wang, J.; Bao L. (2010): Analyzing force concept inventory with item response theory. *Am. J. Phys.* 78 (10), 1064 (2010)
- [8] Shoji, Y.; Munejiri, S.; Kaga, E. (2021): Validity of Force Concept Inventory evaluated by students' explanations and confirmation using modified item response curve. *Physical review physics education research* 17, 020120 (2021)
- [9] Yasuda, Jun-ichiro (2018): Analyzing false positives of four questions in the Force Concept Inventory. *Physical review physics education research* 14, 010112 (2018)
- [10] Stewart, J.; Zabriskie, C.; DeVore, S.; Stewart, G. (2018): Multidimensional item response theory and the Force Concept Inventory. *Physical review physics education research* 14, 010137 (2018)
- [11] Eaton, P.; Willoughby, S.D. (2018): Confirmatory factor analysis applied to the Force Concept Inventory. *Physical review physics education research* 14, 010124 (2018)
- [12] Wells, J.; Henderson, R.; Stewart, J.; Stewart, G.; Yang, J.; Traxler, A. (2019): Exploring the structure of misconceptions in the Force Concept Inventory with modified module analysis. *Physical review physics education research* 15, 020122 (2019)
- [13] Morris, G. A.; Branum-Martin, L.; Harshman, N.; Baker, S. D.; Mazur, E.; Dutta, S.; Mzoughi, T.; McCaule, V. (2006): Testing the test: Item response curves and test quality, *Am. J. Phys.* 74, 449 (2006)
- [14] Morris, G. A.; Harshman, N.; Branum-Martin, L.; Mazur, E.; Mzoughi, T.; Baker, S. D. (2012): An item response curves analysis of the Force Concept Inventory, *Am. J. Phys.* 80 (9) 825-831 (2012).
- [15] Stewart, J.; Drury, B.; Wells, J.; Adair, A.; Henderson, R.; Ma, Y.; P'erez-Lemonche, Á.; Pritchard, D. (2021): Examining the relation of correct knowledge and misconceptions using the nominal response model. *Physical review physics education research* 17, 010122 (2021)
- [16] Hake, R. R. (1998): Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* 66, 64 (1998)
- [17] Stanzel, S.; Schäfle, C.; Junker, E. (2019): Impact of interactive teaching methods on heterogeneity. In: *Proceedings of the 10th international conference on Physics Teaching in Engineering Education, Delft* (2019); auf <https://www.pro-aktjv.de/publikationen> (Stand 05/2023)
- [18] Stanzel, S.; Junker, E.; Graupner, F. (2021): Der Hörsaal als Labor: aktivierende Lehre auf dem Prüfstand. *Die Neue Hochschule*, 2/2021, S. 20 - 23
- [19] McDermott, L. C.; Shaffer, P. S. (1998): *Tutorials in introductory physics*. Prentice Hall, Upper Saddle River (1998); auf Deutsch: *Tutorien zur Physik*, Pearson Studium (2009)
- [20] Knight, R.D. (2023): *Physics for Scientists and Engineers: a Strategic Approach*. 5th edition Pearson Education 2023
- [21] PhET interaktive Simulationen <https://phet.colorado.edu/de/> (Stand 05/2023)

Danksagung

Besonderer Dank gilt den Kolleginnen und Kollegen der TH Rosenheim, die seit zehn Jahren in ihren Physik-Lehrveranstaltungen den FCI durchführen, sowie den 4957 Studierenden für die Bearbeitung des FCI. Für die Auswertung und Zusammenführung der Daten sowie die Erstellung der Graphen danke ich J. Lackovic.

Diese Arbeit wurde im Rahmen des Projektes „Hybride, individuelle und greifbare Hochschullehre in Rosenheimer Qualität (HigHRoQ)“ erstellt, das unter der Projektnummer FMM2020-EA-150 von der Stiftung Innovation in der Hochschullehre gefördert wird.