

Lehre und Forschung

Achim Goerres, Caroline Kärger und Daniel Lambach

Aktives Lernen in der Massenveranstaltung: Flipped-Classroom-Lehre als Alternative zur klassischen Vorlesung in der Politikwissenschaft

1. Einleitung¹

Auf welche Weise erzielen Studierende in politikwissenschaftlichen Großveranstaltungen den besten Lernerfolg?² Anstatt die Studierenden in einer klassischen Vorlesung zu passiven Wissenskonsumenten zu machen, ist der „Flipped Classroom“ oder „Inverted Classroom“³ eine vielversprechende Alternative, die auf die Aktivierung der Studierenden setzt. Dabei wird die Vorlesungszeit zur Großübung in kleinen Arbeitsgruppen, während strukturierte Übersichtsinformationen und Aufgaben zur Vermittlung und Erarbeitung von grundlegendem Wissen in Form von Lektüre und Videos aus der gemeinsamen Vorlesungszeit in die Vorbereitungszeit der Studierenden verlagert werden. Dieser Artikel stellt das innovative Format des Flipped Classroom (FC) vor und bewertet es vorläufig anhand von

- 1 Alle drei Autoren haben das Papier entworfen und die Schlussredaktion durchgeführt. AG hat die Daten gesammelt, die empirische Analyse durchgeführt und den Erstentwurf des Papiers ohne Theorieteil geschrieben. CK und DL haben den Erstentwurf des Theorieteils geschrieben. Die Daten können aus Gründen des Datenschutzes bei Replikationsanfragen nur intern analysiert bzw. Syntax im Auftrag intern zur Anwendung gebracht werden. Wir danken dem Stifterverband für die deutsche Wissenschaft für die finanzielle Unterstützung an DL mit dem Projekttitle "Aktives Lernen in der Massenveranstaltung". Wir danken Michael Löttgen für Hinweise bezüglich des Texts und des Umgangs mit den Daten.
- 2 Unter Lernerfolg verstehen wir das Erreichen von Lernzielen. Den Ausführungen des Artikels und der beschriebenen Anwendung des Flipped Classroom liegen kompetenzorientierte Lernziele zugrunde. Kompetenz wird dabei als holistischer Begriff, der Kenntnisse, Fertigkeiten, Fähigkeiten, Motivationen, Emotionen, Handlungen und Metakognition erfasst, verstanden (European Union 2008; Weinert 2014: 14 f.).
- 3 Die Bezeichnungen Flipped Classroom und Inverted Classroom werden synonym verwendet (Sams 2012: 13). Während in den USA und Kanada verbreitet vom Flipped Classroom gesprochen wird, nutzt man in Deutschland und anderen europäischen Ländern häufiger den Begriff des Inverted Classroom.

explorativer Evidenz einer Methodenvorlesung im ersten Fachsemester des BA Politikwissenschaft an der Universität Duisburg-Essen.

Die Vorlesung ist ein traditionsreiches Veranstaltungsformat der universitären Lehre in nahezu allen Fachgebieten – so auch in der Politikwissenschaft. Im Zeitalter der Massenuniversität gibt es überdies kapazitäts Argumente für eine derartige Form der Lehrveranstaltung. Versteht man Lernen im Sinne eines „Input-Modells“, wonach Studierende durch den Dozenten mit Wissen „aufgefüllt“ werden, ist die Vorlesung ein ressourceneffizientes Lehrformat: für Hunderte von Studierenden wird nur eine einzige Lehrperson benötigt. Die Lehr-Lern-Forschung hat jedoch deutlich gezeigt, dass sich die Vorlesung nur für bestimmte Arten von Lernzielen eignet (Cooper/Robinson 2000: 9; Bligh 1998). Dazu gehört die Einordnung und Systematisierung der Literatur sowie die Erläuterung besonders komplexer Zusammenhänge. Die Lernfortschritte von Studierenden sind jedoch typischerweise auf relativ einfache kognitive Kompetenzen beschränkt und selbst diese Wissenszuwächse verfallen relativ schnell, wenn sie nicht eingeübt werden.

Es gibt viele Möglichkeiten diese klassische Vorlesung im Kleinen zu verändern, beispielsweise durch interaktive Abstimmungssysteme.⁴ Als grundsätzliche Frage bleibt jedoch bestehen, ob die Face-to-face-Zeit zwischen dem Dozenten und den Studierenden in der Vorlesung effizient für den Lernerfolg der Studierenden eingesetzt wird. Mögliche Effizienzverluste liegen auf der Hand: aufseiten der Studierenden ist die physische Teilnahme an der Vorlesung keine hinreichende Bedingung für das Verstehen des Stoffes und noch nicht einmal für dessen erste Aufnahme. Die Gefahren sind groß, die Vorlesung als soziales Ereignis zu sehen, in denen Studierende schweigend oder sprechend mit Kommilitonen kommunizieren, das Internet für soziales Netzwerken und Nachrichten nutzen (Junco 2012), und durch langes passives Sitzen unkonzentriert zu werden. Aufseiten der Dozierenden besteht die Gefahr der zu großen Vertrautheit mit dem Stoff. Nach einigen Durchläufen einer Vorlesungseinheit schwindet das Gespür für die Schwierigkeit des Stoffes und die eventuell notwendige Verlangsamung des Vortrags. Angesichts dieser beidseitigen Effizienzverluste sollte das grundsätzliche Vorlesungsschema kritisch hinterfragt werden. Das FC-Format setzt hier an und hat zum Ziel, die wertvolle Interaktionszeit zwischen den Studierenden untereinander und zwischen dem Dozenten und den Studierenden zu nutzen, die aktive Auseinandersetzung mit dem Thema zu fördern und höherwertige Kompetenzen

4 Ausführlicher zum Einsatz von Abstimmungssystemen siehe Wiemeyer 2013. Beispielhafte Abstimmungssysteme wären Pingo (Reinhardt u. a. 2012), OnlineTED oder ARSnova.

(Anwendung, Synthese, Beurteilung etc.; Bloom u. a. 1974: 29-33; Anderson/Krathwohl 2001: 67 f.) anzustreben.

Der Artikel ist wie folgt strukturiert: Zunächst geben wir eine Beschreibung der definierenden Eigenschaften des FC und fassen die Lehr-Lern-Forschung zu diesem didaktischen Instrument zusammen. Danach stellen wir die Umsetzung des FCs anhand einer Sitzung einer Methodenvorlesung dar. Wir präsentieren die empirische Evidenz einerseits über die Effektivität dieses Formats auf Prüfungsergebnisse und andererseits über die Wahrnehmung des Formats in der Studierendenschaft. Im letzten Teil fassen wir unsere Erfahrungen mit diesem Format themenartig zusammen.

2. Eigenschaften des Flipped-Classroom-Formats

Beim FC-Format wird die passive Rezeption von Lerninhalten aus dem Unterricht in die Vorbereitungsphase verlagert, beispielsweise durch die Lektüre von Texten, das Betrachten von aufgezeichneten Videovorlesungen, das Hören von Podcasts etc.⁵ Ergänzt wird dieses Vorgehen durch Instrumente, mit denen die Studierenden ihr Verständnis des Materials überprüfen können, etwa durch online bereitgestellte Übungsaufgaben. Folglich wird in dieser Phase vorrangig die Aneignung der Kompetenzen des Wissens und Verstehens von Inhalten verfolgt. Die Präsenzphase dient demgegenüber zur Anwendung des erworbenen Wissens sowie zur Vertiefung des Materials und zur Klärung von Fragen. Sie baut systematisch auf den Kompetenzen auf, die die Studierenden in der Vorbereitungsphase erworben haben.

Aus dieser Umkehrung des üblichen Formats können *theoretisch* verschiedene Vorteile entstehen. Erstens haben Studierende stärkere Anreize zur Vorbereitung auf Präsenzsitzungen, da das Verständnis der in der Präsenzzeit vermittelten Inhalte und der aktive Umgang damit das Verständnis der Inhalte aus der Vorbereitungsphase voraussetzen. Die Kosten der Prokrastination steigen, also in diesem Fall des systematischen Verschleppens von Vorbereitung. Zweitens kann die Lehrperson stärker im Rahmen des Just-in-Time-Teaching (Simkins/Maier 2010; Mazur/Watkins 2010) auf individuelle Lernprozesse eingehen und Studierenden bei Verständnisproblemen helfen, da sie den Lernstand und -fortschritt der Studierenden einerseits anhand der Ergebnisse der Vorbereitungsphase und anderer-

5 Zu den Grundlagen dieses Ansatzes vgl. Lage/Platt/Treglia 2000 und Schäfer 2012.

Schaubild 1: Traditionelles Vorlesungsformat und Inverted Classroom Vorlesungsformat im Vergleich

Traditionelles Vorlesungsformat		Inverted Classroom Vorlesungsformat	
Phase	Aktivität	Phase	Aktivität
1. Präsenzphase 	Wissensvermittlung und Verstehen	1. Individuelle Phase 	Wissensvermittlung und Verstehen
2. Individuelle Phase 	Festigung und Vertiefung des Wissens (Anwendung, Analyse, Diskussion, Transfer etc.)	2. Präsenzphase 	Festigung und Vertiefung des Wissens (Anwendung, Analyse, Diskussion, Transfer etc.)

Quelle: Eigene Darstellung.

seits durch die Interaktion in der Präsenzphase einschätzen kann. Im FC rücken die Studierenden in das Zentrum aller Aktivitäten (Bergmann/Sams 2012). Drittens greift die Methode die oft von Studierenden geäußerte Kritik am „Bulimie-Lernen“ – der kurzfristigen, schnellen Aufnahme von oberflächlichem Wissen – auf, indem mittels FC auch anspruchsvollere Kompetenzziele verfolgt werden, die über die reine Wissensreproduktion hinausgehen. Viertens hat der FC auch eine soziale Komponente, indem die Studierenden stärker mit der Lehrperson sowie untereinander interagieren. Durch dieses „peer teaching“ (Crouch and Mazur 2001; Crouch 2007; Mazur 1997) verbreiten sich Normen der angemessenen Vorbereitung, außerdem steigert das gegenseitige Erklären das Verstehen des Materials.

Die Wirksamkeit des FC und vergleichbarer Techniken gegenüber reinem Frontalunterricht ist bereits *empirisch* in einer Reihe von Studien nachgewiesen worden.⁶ Ein verbesserter Lerneffekt trat auch in Massenveranstaltungen mit mehr als 100 Teilnehmern auf (Bates/Galloway 2012; Moravec u. a. 2010). Dafür gibt es verschiedene Erklärungen. Erstens wird das klassische „Aufmerksamkeitsproblem“ der Vorlesung gelöst – Studien haben gezeigt, dass Studierende nach etwa 15 Minuten reinen Zuhörens immer weniger Informationen aufnehmen können (Bligh 1998; Cooper/Robinson 2000; Szpunar/Khan/Schacter 2013). Zweitens ermöglicht die Aktivierung der Studierenden die Verfolgung höherwertiger

6 Vergleiche Carlisle 2010; Crouch/Mazur 2001; Fagen/Crouch/Mazur 2002; Hake 1997; Talbert 2012.

Lern- und Kompetenzziele. Die Studierenden werden durch die Bearbeitung von Übungs- und Testfragen zur Abstraktion und zum Transfer des erworbenen Wissens aufgefordert und erhalten auf diese Aufgaben ein unmittelbares Feedback durch Peers und den Dozierenden (Cooper/Robinson 2000). Drittens bleibt angewandtes Wissen länger und besser haften als lediglich passiv rezipiertes Wissen. Viertens können Dozierende durch Just-in-Time Teaching, also zum Beispiel der spontanen Beantwortung von Fragen der Studierenden in der Präsenzphase oder im Rahmen der Vorbereitungsphase formulierten Fragen, auf die Fragen eingehen, bei denen Studierende die meisten Verständnisprobleme haben. Angesichts dieser Vorzüge überrascht es nicht, dass Studierende in Evaluationen den FC dem klassischen Vorlesungsformat vorziehen (Bates/Galloway 2012; Moravec u. a. 2010). Dies äußert sich auch in einer stärkeren Leistungsbereitschaft unter Studierenden: Carlisle (2010) berichtet, dass Studierende in FC-Kursen mehr Zeit für die Vorbereitung auf eine Präsenzsitzung verwenden als Teilnehmer einer klassischen Vorlesung (ähnlich Bates/Galloway 2012; Moravec u. a. 2010).

Bislang wird der FC hauptsächlich in MINT-Fächern und den Wirtschaftswissenschaften angewendet.⁷ In den Sozialwissenschaften wird dieser didaktische Ansatz bisher nur sporadisch eingesetzt (zu den Ausnahmen gehören Forsey/Low/Glance 2013) bzw. wird der Einsatz des Formats weniger in Publikationen vorgestellt und reflektiert. Touchton (2015) berichtet, dass der Einsatz des FC in einer Vorlesung zur Statistik für Politikwissenschaftler signifikante Lernzuwächse bei fortgeschrittenen Anwendungsproblemen gebracht hätte und dass Studierende das didaktische Format einer traditionellen Vorlesung vorziehen würden. Diesen Befund können wir mit unseren in diesem Artikel analysierten Anwendungsfall untermauern. Dies bestätigt zudem die Ergebnisse der Studien anderer Fächer und unterstreicht, dass das Format auch in der politikwissenschaftlichen Lehre Vorteile bringen kann.

Es gibt gute Gründe, warum die Methode in der politikwissenschaftlichen Ausbildung noch weiter eingesetzt werden sollte. Lernerfolge können an der Erreichung von Lernzielen festgemacht werden. Dabei kennzeichnet die Politikwissenschaft in vielen Fällen, dass neben der Kenntnis von Faktenwissen insbesondere die Entwicklung, Analyse und Diskussion von Argumenten im Fokus stehen. Das heißt, dass die Struktur, Originalität und Überzeugungskraft eines Argumentes genauso wichtig sind wie die Korrektheit von Faktenwissen. Gleichmaßen verlangt eine Vielzahl dieser Aspekte höherwertige Kompetenzen, die über reines

7 Vergleiche Bates/Galloway 2012; Foertsch u. a. 2002; Love u. a. 2013; Loviscach 2012; McGivney-Burelle/Xue 2013; Moravec u. a. 2010.

Faktenwissen und Verstehen hinausgehen. Folglich orientiert sich der Lernerfolg zahlreicher politikwissenschaftlicher Studiengänge an der Entwicklung und Förderung höherwertiger kognitiver Kompetenzen wie zum Beispiel Analyse, Synthese, Vergleich und Kritik (Bloom u. a. 1974). Dieser kompetenzorientierte Lernerfolg ist zudem in den entsprechenden Prüfungsordnungen fixiert und vorgegeben. Diese Kompetenzen können nicht durch die passive Rezeption von Wissen in einer klassischen Vorlesung ausgebildet werden. Der FC ermöglicht die Verfolgung dieser Kompetenzen. Durch den FC werden Studierende aktiviert und lernen, eigene Meinungen aufzustellen und diese im Austausch mit anderen Studierenden argumentativ zu verteidigen. Diese Form des Dialogs ist in Seminaren gang und gäbe. Durch den FC kann dies auch im Rahmen von Vorlesungen stattfinden.

3. Erfahrungen mit dem Flipped-Classroom-Format

a) Studierverhalten und Rezeption unter den Studierenden

Die Vorlesung „Methoden der Empirischen Sozialforschung“ des erstgenannten Autors ist im ersten Fachsemester des BA Politikwissenschaft der Universität Duisburg-Essen verortet und wird fast ausschließlich von Studierenden aus diesem Studiengang absolviert. Die Vorlesung ist dreistündig und wird mit einer Pause als eine 90minütige und eine 45minütige Einheit wöchentlich gehalten. Die typische Größe einer eingeschriebenen Kohorte lag in den Wintersemestern 2012/13, 2013/14 und 2014/15 bei etwa 220 bis 260 Studierenden.

Der tatsächliche Besuch der Vorlesung variierte im WS 2014/15 schätzungsweise zwischen 60 und 130 Studierenden pro Sitzung. Auf der Onlineplattform zur Vorlesung waren gegen Ende der Vorlesungszeit etwa 330 Studierende eingetragen, was als maximale Obergrenze für eine Zahl der eventuellen Prüflingskandidaten gesehen werden kann. Kurzum handelt es sich also um eine typische für die Studierenden verpflichtende Massenveranstaltung.

Im Wintersemester 2013/14 wurde das erste Mal eine der 14 substanziellen Vorlesungseinheiten durch eine FC-Sitzung ersetzt, nämlich die Einheit über Messen in der empirischen Forschung (Validität, Reliabilität, Objektivität, Skalen, Indizes etc.). Dieser Austausch im Format wurde im WS 2014/15 für die gleiche Sitzung wiederholt. Durch die Beibehaltung der Prüfungsfragen in der schriftlichen Klausur können die Prüfungsergebnisse zu den Fragen zu dieser Vorlesungseinheit aus dem Wintersemester 2012/13 (ohne FC-Einheit) mit den Ergebnissen der beiden folgenden Wintersemester verglichen werden. Anhand von Um-

frage- und Prüfungsergebnissen sowie beobachtetem Studierverhalten können wir im Folgenden einen Eindruck bezüglich der Umsetzung des FC gewinnen.

Das FC-Format wurde den Studierenden zu Beginn der Vorlesung ausführlich erklärt und dann noch einmal kurz vor der Sitzung (der vierten in der Abfolge der Vorlesungssitzungen). Die Informationen zur FC-Sitzung wurden schriftlich über die Onlineplattform sowie mündlich in der Vorlesung verteilt. Den Studierenden standen zur Sitzungsvorbereitung folgende Materialien zur Verfügung: ein Onlinevideo dieser Sitzung im klassischen Format aus dem WS 2012/13 von etwa 120 Minuten Laufzeit nebst Folien im PDF-Format, die Pflichtlektüre, ein Online-Multiple-Choice-Test mit acht Fragen zur Überprüfung des Grundverständnisses sowie die Möglichkeit, über ein Online-Forum oder per E-Mail Fragen an den Dozenten zu stellen.

Im WS 2014/15 wurde im Monat der Sitzung der erste Teil der Online-Vorlesung 882 Mal aufgerufen und der zweite Teil 344 Mal, wobei sich nicht erkennen lässt, ob Videos in toto angesehen wurden. Durch die unterschiedlichen Abrufraten der inhaltlich zusammenhängenden Videos, die nur aus technischen Gründen getrennt worden waren, lässt sich aber schlussfolgern, dass bei Weitem nicht alle Abrufversuche bis zum Ende vollzogen worden sind. Ein weiterer technischer Grund für die deutlich höhere Abrufrate beim ersten Video könnten technische Verbindungsprobleme mit langsamen Internetverbindungen sein (s. unten). Da die Videos nur digitale Kopien der Sitzung im klassischen Format waren, könnte hier das Aufmerksamkeitsproblem und eine mangelnde Instruktion des Zusammenhangs zwischen Onlinevideos und Inhalt der FC-Sitzung vorgelegen haben.

Für den Onlinetest liegen 117 beendete Testversuche vor, von denen 93 beendete Erstversuche sind. Für die Erstversuche liegt der Mittelwert der korrekt beantworteten Fragen bei 72 %, für die Letztversuche bei 78 %. Die Erfolgsquote spricht dafür, dass die Testteilnehmer im Großen und Ganzen den Stoff verstanden hatten, wobei unklar ist, inwieweit während des Tests Folien oder Lehrbücher zurate gezogen oder Fragen aus dem Gedächtnis beantwortet wurden.

In der FC-Sitzung wurden die Studierenden im Sinne des Peer Teaching aufgefordert sich in Kleingruppen à 3-4 Personen selbständig zusammenzufinden, um eigenständig und im eigenen Tempo 20 schriftliche Aufgaben in maximal 90 Minuten zu bearbeiten, beispielsweise: „Sie wollen die öffentliche Meinung in der Gesellschaft zu der Piratenpartei mithilfe von Twiternachrichten in einem bestimmten Zeitraum messen. In Bezug auf welches Gütekriterium der Messung sollten Sie sich Sorgen machen?“ Hier müssen die Studierenden die Kriterien Reliabilität, Objektivität und Validität in der konkreten Form anwenden. Nach einer Pause wurden die Aufgaben durch den Dozenten besprochen und weitere Fragen

geklärt. Es wurde hervorgehoben, dass die Antworten nicht eingesammelt werden um dem Missverständnis vorzubeugen, es handele sich um eine Prüfung. 85 Studierende waren anwesend, was etwas unter den Schätzungen der vorherigen Wochen lag.

Alle 85 Studierenden füllten einen kurzen schriftlichen Fragebogen zur Messung von Studierverhalten und Zufriedenheit aus. Nach diesen Angaben hatten nur 63 % der Studierenden die Onlinevideos in Gänze geschaut, weitere 27 % zum Teil. 76 % der Studierenden gaben an, die Folien genutzt zu haben. 67 % bestätigten eine Nutzung des Onlinetests. Wenn man berechnet, wie groß der Anteil der Studierenden ist, der die vollständige Vorbereitung gemäß der Empfehlung mit Onlinevideos, Folien und Selbsttest gemacht hat, ergibt sich nur ein Anteil von 41 % der Studierenden. Mit anderen Worten: Von den präsenten Teilnehmern waren weniger als die Hälfte optimal vorbereitet. Auf die Frage nach der Wahrnehmung der eigenen Vorbereitung gaben sich die Studierenden im Schnitt eine 2,5 auf einer Schulnotenskala von 1 bis 5. Die von ihnen berichtete sub-optimale Vorbereitung war den Befragten also bewusst.

Um die Zufriedenheit der Teilnehmer mit dem FC-Format zu messen, wurde ihnen folgende Frage gestellt: „Alles in allem, wie hat Ihnen die FC-Sitzung mit Ihrer Vorbereitung im Vergleich zu einer klassischen Vorlesung mit Ihrer persönlich typischen Vorbereitung gefallen?“ Hierauf gaben 38 % der Befragten an, dass ihnen die FC-Sitzung besser gefallen habe, 42 %, dass beide Modi gleich gut seien, und 20 %, dass FC schlechter sei. Summa summarum waren die Studierenden im Vergleich zu den vorangegangenen drei Vorlesungssitzungen im klassischen Format mit der FC-Sitzung genauso zufrieden oder zufriedener. Diese positive Bewertung spiegelt auch die Befunde aus der Literatur wider (Handke 2014: 26; Lage/Platt/Treglia 2000: 35 ff.).

Um zu verstehen, warum bestimmte Studierende nicht in der Sitzung anwesend waren, wurde eine Onlineumfrage programmiert, zu der explizit über die Onlineplattform alle diejenigen aufgerufen worden waren, die nicht an der Sitzung im FC-Format teilgenommen hatten. Es nahmen 49 Studierende teil, die aber nur einen Bruchteil aller potenziell Interessierten darstellten. Uns interessierte vor allem, warum diese Studierenden nicht an der Sitzung teilgenommen haben. Unter den möglichen Mehrfachnennungen ergab sich folgende Reihung: 39 % keine ausreichende Vorbereitung, 37 % krank, 31 % anderer Termin, 31 % Erwartung, dass in einer Vorlesung vor allem der Dozent reden soll, 27 % kein genaues Verständnis des Formats, 27 % FC-Format missfallen, 24 % niemals Besuch der Vorlesung, 10 % Unwillen, sich mit anderen auszutauschen, 2 % Onlinetest schon gut

bestanden. Eine weitere wiederholt auftretende offene Nennung war das schlechte Funktionieren des Onlinevideos.

Wie die Angaben der Nichtteilnehmer zeigen, hat die mangelnde Vorbereitung einige Studierende gewissermaßen präventiv davon abgehalten an der Präsenzsitzung teilzunehmen. Bemerkenswert sind die Erwartungshaltungen. Für zehn Prozent der Befragten verhinderte der zu erwartende Austausch unter Peers die Teilnahme an der Vorlesung, 31 Prozent führten die erwartete monologhafte Struktur an. Dass 27 % der Befragten das Format missfiel, das sie, sofern sie keine Wiederholer waren, nicht aus eigener Erfahrung kannten, zeigt ein weiteres Informationsdefizit.

Zusammengefasst: Bei der Durchführung der FC-Sitzung zeigte sich sowohl bei den anwesenden als auch bei den nicht anwesenden Studierenden das Problem der mangelnden Vorbereitung. Nur 41 % der anwesenden Studierenden waren optimal vorbereitet. Die Compliance der Studierenden mit den Vorgaben in Bezug auf die Vorbereitung war nicht ausreichend, wobei man die Frage stellen muss, ob dieses niedrige Niveau nicht schon die Vorbereitung bei einer normalen Vorlesung überschreitet und somit positiv zu sehen ist. Weiterhin zeigt die Befragung der nicht anwesenden Studierenden, dass es – selbst bei diesen Studierenden in ihrer vierten Vorlesungswoche – zumindest als Vorwand die vorgetragene Erwartung gibt, dass eine Vorlesung hauptsächlich durch Dozierende und frontal zu erfolgen habe. Schlussendlich zeigen unsere Befunde auch die grundsätzliche höhere Zufriedenheit von 38 % der Befragten gegenüber einer klassischen Vorlesung und weiteren 42 %, dass sie beide Modi gleich gut fänden.

b) Effekte auf das Klausurergebnis

Welchen Einfluss hat das FC-Format auf die Prüfungsergebnisse? Hier liegen Daten für vier Klausuren vor, jeweils der reguläre Termin und die Wiederholungsprüfung im WS 2012/13 und im WS 2013/14, insgesamt also 377 verwertbare Ergebnisse.

Vier Prüfungsfragen können über alle Prüfungen hinweg als äquivalent angesehen werden (siehe Tabelle 1). Frage 1 verlangte von den Studierenden, vier verschiedene Variablen anhand der Beschreibung bezüglich ihres Skalenniveaus (nominal, ordinal, metrisch) und der Unterscheidung latent/manifest einzuordnen. Frage 2 beschrieb eine Situation, in der Triangulation angewendet wurde, so dass die Studierenden den Begriff benennen mussten. Frage 3 beschrieb ein inhaltsanalytisches Kodierprojekt, bei dem Probleme des manuellen Kodierens mit Reliabi-

Tabelle 1: Übersicht über Prüffragen

#	Fragenstichworte	Lernziel nach Bloom u. a. 1974	Art der Frage	Max. Punktzahl	MW erreichbare Punkte (in % des Maximums)	MW FC-Gruppe	MW Nicht-FC-Gruppe
1	Variablen einschätzen	Anwendung	MC	4	3,1 (77,5 %)	3,1	3,0
2	Triangulation erkennen und benennen	Verstehen	Offen	4	0,5 (12,5 %)	0,5	0,4
3	Messqualitätskonzepte bei Inhaltsanalyse anwenden	Anwendung	Offen	2	0,9 (45,0 %)	1,1	0,9
4	Idealeigenschaften eines Indizes benennen	Verstehen	MC	3	1,8 (60,0 %)	1,6	1,9

Quelle: Eigene Darstellung. MC = Multiple Choice.

lilität und Objektivität benannt werden mussten. Frage 4 fragte nach den Idealeigenschaften eines Indizes. Die vier Fragen unterschieden sich im Aufwand (maximale Punktzahl zwischen 2 und 4), Schwierigkeitsgrad (% des Maximums am Mittelwert), im Format (Multiple Choice versus offene Fragen) und im Lernziel nach Bloom (Verstehen versus Anwenden).

Diese Varianz in der Fragenart können wir systematisch in einer Regressionsanalyse ausbeuten. Im einfachen Vergleich der Punktergebnisse (siehe Tabelle 1) sehen wir, dass die Fragen 1 bis 3 in der FC-Gruppe leicht besser als in der Nicht-FC-Gruppe beantwortet wurden. Hier spielen allerdings die Unterschiede in den Eingangskohorten hinein, da die 2013/14er Kohorte im Schnitt besser war.

Aus Platzgründen präsentieren wir hier überblicksartig nur einige ausgewählte Regressionsergebnisse, die der Hypothese nachgehen, dass die Teilnahme an einer FC-Kohorte im Schnitt mit einer höheren Punktzahl in der Prüfung einhergeht. Wir kontrollieren für die Gesamtpunktzahl als Messung der latenten Fähigkeit, das Geschlecht, den Prüfversuch des Prüflings (1., 2. oder 3. Versuch), den Prüftermin (regulärer Termin versus Wiederholungstermin), die Art der Frage (Multiple Choice versus offene Frage), die Art des Lernziels (Anwendung versus Verstehen) und als alternative Messung für Art der Frage und Lernziel den Schwierigkeitsgrad. Genauere Details befinden sich im Anhang. Die Regressionsanalyse hat das Punktergebnis bei einer Frage als abhängige Variable.

Kurz gesagt: Es gibt keinen eindeutigen Effekt der FC-Gruppe für alle Prüffragen. Der direkte Effekt der Mitgliedschaft in einer FC-Gruppe ist zwar leicht positiv mit 0,02 auf einer Punktskala von 0 bis 4 in dem Grundmodell mit allen

Kontrollvariablen, aber nicht sehr präzise (p-Wert des einseitigen Tests 0,36). Mit anderen Worten: unter Konstanthaltung einer ganzen Reihe von Effekten wirkte sich die Teilnahme an einer FC-Gruppe im Schnitt nicht auf das Punktergebnis aus.

Interessanter wird es, wenn man untersucht, ob die Art der Fragen einen Unterschied ausmacht (über eine Interaktionsanalyse). Auch hier kann man keinen präzisen Effekt finden, wenn man Multiple Choice versus offene Fragen und generell die Schwierigkeit der Frage als Kontextvariable für die FC-Gruppenteilnahme schätzt. Allerdings zeigt sich bei der Unterscheidung nach Lernzielen ein systematisches Muster. Für die Prüffragen, die eher das Verstehen nach Bloom zum Ziel haben, hat der FC keinen positiven Effekt. Für Prüffragen, die eher die Anwendung von Wissen, also ein höherwertiges Ziel, anstreben, hat der FC einen signifikant positiven Effekt (p-Wert = 0,01) von 0,21 auf der Punktskala, das heißt im Schnitt wurden Anwendungsprüffragen unter Kontrolle der vielen Faktoren von Prüflingen, die in FC-Gruppen waren und damit die Möglichkeit gehabt hatten, an der Sitzung teilzunehmen, mit 0,21 Punkten besser absolviert. Weitere Interaktionsanalysen, wie mit dem Termin der Prüfung, dem Prüfversuch (1., 2., 3.) und der Gesamtzahl erreichter Punkte, brachten keine weiteren systematischen Muster zutage. Gerade der letzte Befund ist noch einmal bemerkenswert. Er deutet nämlich darauf hin, dass die FC-Möglichkeit nicht grundsätzlich für „High Performer“ mehr gebracht hat als für „Low Performer“.

In Summe lassen die Befunde ein besseres Prüfergebnis für tiefergehende Fragen bei den Studierenden nachweisen, die in der FC-Gruppe waren. Dieses Ergebnis besteht unabhängig von der generellen Leistungsfähigkeit der Studierenden und ohne Kenntnis darüber, ob der betreffende Studierende tatsächlich an der Sitzung teilgenommen hatte. Der Befund ist umso erstaunlicher, weil nur eine einzige Vorlesungssitzung ausgetauscht worden war. Zudem lässt sich keine Interaktion zwischen dem Format und der individuellen Performanz der Studierenden zeigen, das heißt sehr gute Studierende schnitten durch die Teilnahme an der FC-Sitzung nicht besser ab, bzw. sehr schlechte Studierende nicht noch schlechter. Eine Zweiteilung dieser Vorlesungsform auf unterschiedliche Studierenden-Gruppen konnten wir nicht belegen. Gleichzeitig darf nicht vergessen werden, dass die Effektivität des FC anhand der Prüfungsergebnisse untersucht wurde und nicht durch eine Selbsteinschätzung der Studierenden. Die Ergebnisse anderer FC-Anwendungen verdeutlichen, dass eine Mehrheit der Studierenden in der Selbsteinschätzung angibt, in diesem Format einen höheren Lernerfolg zu haben als im klassischen Vorlesungsformat (Decker/Beier 2014: 11 f.; Lage/Platt/Treglia 2000: 35 ff.).

4. Schlussfolgerung

Der Flipped Classroom als Format für eine Vorlesungssitzung hat in der einführenden Methodenvorlesung nachweislich zu Verbesserungen des Klausurergebnisses im Bereich der Anwendungsfragen geführt. Bei Fragen, die auf die reine Reproduktion und das Verständnis von Wissen abzielen, lässt sich kein präziser Effekt belegen. Ob die Prüfungen offen oder Multiple Choice sind, spielt dabei keine Rolle.

Damit können wir den Mehrwert von FC zur Erlangung anspruchsvoller Kompetenzen bestätigen. Zur reinen Reproduktion von Wissen scheint die Methode hier keinen Vorteil erbracht zu haben. Das Format und die Form der Prüfung müssen das Messen von entsprechenden Kompetenzen aber auch ermöglichen. Wichtig ist auch der Befund, dass der FC im Prüfungsergebnis die besseren Studierenden grundsätzlich nicht stärker macht und die Kluft zu den schwächeren Studierenden nicht vergrößert.

Die vorliegenden Ergebnisse sind natürlich begrenzt aussagefähig, insbesondere da nur eine Sitzung im FC-Format durchgeführt wurde, während für die anderen Sitzungen das Format der klassischen Vorlesung beibehalten wurde. Man kann dies als einen konservativen Test der Effektivität ansehen; jedoch stellt sich die Frage, ob eine einzelne FC-Sitzung in einer klassischen Vorlesung nicht andere Effekte hat als eine Vorlesung, die komplett nach dem FC-Modell arbeitet.

Der curriculare Kontext, eine verpflichtende Methodenveranstaltung in einem sozialwissenschaftlichen BA-Studium, ist aus Sicht einer eher diskursiven Disziplin wie der Politikwissenschaft ein harter Test für FC. Ob sich die Effektivität für andere Kurse und Inhalte der Politikwissenschaft reproduzieren lässt, bleibt eine offene Frage. Man könnte vermuten, dass der FC relativ zum klassischen Format umso mehr Effekt hat, je abstrakter eine Vorlesung ist.

Was sind praktische Stolpersteine bei der Umsetzung?

- Die Erwartungshaltung der Studierenden muss nachhaltig gemanagt werden, damit die Studierenden, die erstaunlich deutliche Erwartungen an eine Vorlesung im klassischen Sinne haben, genau wissen, was sie erwartet und adäquat reagieren.
- Die Vorbereitung der Studierenden ist beim FC anders als bei der klassischen Vorlesung viel notwendiger für Studierende, um erfolgreich teilzunehmen. Wenn Studierende über mehrere Sitzungen hinweg diese Erfahrung machen, sollte sich eine Verbesserung des Vorbereitungsverhaltens einstellen.
- FC ist für Dozierende zeitintensiv, weil jede Vorlesung zwei Vorbereitungszyklen hat, einmal des Vorbereitungsmaterials und einmal des Inhalts der Präsenzsitzungen, und insbesondere in der ersten Umgestaltung von Sitzungen oder gan-

zen Veranstaltungen mit einem Mehraufwand an konzeptioneller Arbeit verbunden.

Der FC ist also kein Allheilmittel, um alle Schwächen des klassischen Vorlesungsformats zu beseitigen. Studierende werden durch dieses Format nicht zu exzellenten Studierenden und die Prüfungsergebnisse verbessern sich nicht allesamt auf wundersame Weise. Die Aufgabe der Lehrenden wird dadurch nicht einfacher, aber vielleicht bereichernder und interaktiver. Der FC ist aber, sofern er bewusst und durchdacht eingesetzt wird, ein starkes Werkzeug zur Aktivierung von Studierenden und zur Verfolgung höherwertiger Kompetenzen und eine Ressource, welche die Zentrierung von der Lehrperson auf die Studierenden richtet.

Literatur

- Anderson, Lorin W./Krathwohl, David R.*, 2001: A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives, New York.
- Bates, Simon/Galloway, Ross*, 2012: The Inverted Classroom in a Large Enrollment Introductory Physics Course: A Case Study, in: HEA STEM Conference. London.
- Bergmann, Jonathan/Sams, Aaron*, 2012: Flip Your Classroom. Reach Every Student in Every Class Every Day, Eugene.
- Bligh, Donald*, 1998: What's the Use of Lectures?, Exeter.
- Bloom, Benjamin S./Engelhart, Max D./Furst, Edward J./Hill, Walker H./Krathwohl, David R.*, 1974: Taxonomie von Lernzielen im kognitiven Bereich, Weinheim.
- Carlisle, C. Martin*, 2010: Using YouTube to Enhance Student Class Preparation in an Introductory Java Course, in: 41st ACM Technical Symposium on Computer Science Education, New York.
- Cooper, James L./Robinson, Pamela*, 2000: The Argument for Making Large Classes Seem Small, in: New Directions for Teaching and Learning (81), 5-16.
- Crouch, H. Catherine/Mazur, Eric*, 2001: Peer Instruction: Ten Years of Experience and Results, in: American Journal of Physics (69), 970-977.
- Crouch, H. Catherine u. a.*, 2007: Peer Instruction: Engaging Students One-on-One, All At Once, in: Physics Education Research 1 (1), 5-50.
- Decker, Christian/Beier, Stephan*, 2014: The Instructor as Navigator: Empirical Evidence of the Implementation of the ICM at HAW Hamburg, in: E.-M.

- Großkurth/J. Handke (Hrsg.), *The Inverted Classroom Model. The 3rd German ICM-Conference-Proceedings*, Berlin.
- European Union*, 2008: Empfehlungen des Europäischen Parlaments und des Rates vom 23. April 2008 zur Einrichtung des Europäischen Qualifikationsrahmens für lebenslanges Lernen (2008/C 111/01), [http://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32008H0506\(01\)&from=EN](http://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32008H0506(01)&from=EN) (Stand: 19.8.2014).
- Fagen, P. Adam/Crouch, H. Catherine/Mazur, Eric*, 2002: Peer Instruction: Results from a Range of Classrooms, in: *The Physics Teacher* 40, 206-209.
- Foertsch, Julie/Moses, Gregory/Strikwerda, John/Litzkow, Mike*, 2002: Reversing the Lecture/Homework Paradigm Using eTEACH® Web-based Streaming Video Software, in: *Journal of Engineering Education* 91 (3), 267-274.
- Forsey, Martin/Low, Mitchell/Glance, David*, 2013: Flipping the Sociology Classroom: Towards a Practice of Online Pedagogy, in: *Journal of Sociology* 49 (4), 471-485.
- Hake, R. Richard*, 1997: Interactive-Engagement Versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses, in: *American Journal of Physics* 66 (64), 64-77.
- Handke, Jürgen*, 2014: The Inverted Classroom Mastery Model – A Diary Study, in: E.-M. Großkurth/J. Handke (Hrsg.), *The Inverted Classroom Model. The 3rd German ICM-Conference-Proceedings*, Berlin.
- Junco, Reynol*, 2012: In-Class Multitasking and Academic Performance, in: *Computers in Human Behavior* 28 (6), 2236-2243.
- Lage, Maureen J./Platt Glenn J./Treglia, Michael*, 2000: Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment, in: *Journal of Economic Education* 31 (1), 30-43.
- Love, Betty/Hodge, Angie/Grandgenett, Neal/Swift, Andrew W.*, 2013: Student Learning and Perceptions in a Flipped Linear Algebra Course, in: *International Journal of Mathematical Education in Science and Technology* 45 (3), 317-324.
- Loviscach, Jörn*, 2012: Videoerstellung für und Erfahrung mit dem ICM, in: J. Handke/A. Sperl (Hrsg.), *Das Inverted Classroom Model: Begleitband zur ersten deutschen ICM-Konferenz*, München.
- Mazur, Eric*, 1997: *Peer Instruction. A User's Manual*, Upper Saddle River.
- Mazur, Eric/Watkins, Jessica*, 2010: Just-in-Time Teaching and Peer Instruction, in: S. Simkins/M. Maier (Hrsg.), *Just in Time Teaching: Across the Disciplines, Across the Academy*, Sterling.

- McGivney-Burrelle, Jean/Xue, Fei*, 2013: Flipping Calculus, in: *Primus* 23 (5), 477-486.
- Moravec, Marin/Williams, Adrienne/Aguilar-Roca, Nancy/O'Dowd, Diane K.*, 2010: Learn before Lecture: A Strategy That Improves Learning Outcomes in a Large Introductory Biology Class, in: *CBE—Life Sciences Education* 9 (4), 473-481.
- Reinhardt, Wolfgang/Sievers, Michael/Magenheim, Johannes/Kundisch, Dennis/Herrmann, Philipp/Beutner, Marc/Zoyke, Andrea*, 2012: PINGO: Peer Instruction for Very Large Groups, in: A. Ravenscroft/S. Lindstaedt/C. Kloos/D. Hernández-Leo (Hrsg.), *21st Century Learning for 21st Century Skills*, Berlin.
- Sams, Aaron*, 2012: Der Flipped Classroom, in: J. Handke/A. Sperl (Hrsg.), *Das Inverted Classroom Model: Begleitband zur ersten deutschen ICM-Konferenz*, München.
- Schäfer, Anna Maria*, 2012: Das Inverted Classroom Modell, in: J. Handke/A. Sperl (Hrsg.), *Das Inverted Classroom Model: Begleitband zur ersten deutschen ICM-Konferenz*, München.
- Simkins, Scott P./Maier, Mark*, 2010: Just-In-Time Teaching: Across the Disciplines, Across the Academy, Sterling.
- Szpunar, K. Karl/Novall Khan, Y./Daniel Schacter, L.*, 2013: Interpolated Memory Tests Reduce Mind Wandering and Improve Learning of Online Lectures. Paper read at National Academy of Sciences, Washington D.C.
- Talbert, Robert*, 2012: Inverted Classroom, in: *Colleagues* 9 (1), 1-3.
- Touchton, Michael*, 2015: Flipping the Classroom and Student Performance in Advanced Statistics: Evidence from a Quasi-Experiment, in: *Journal of Political Science Education* 11 (1), 28-44.
- Weinert, Franz E.*, 2014: Leistungsmessungen in Schulen, Weinheim.
- Wiemeyer, Leonie*, 2013: Clicker-Happy Audience Response System as an Interface Between Preclass Preparation and Inclass Session, in: J. Handke/N. Kiesler/L. Wiemeyer (Hrsg.), *The 2nd German ICM-Conference-Proceedings*, München.

Korrespondenzanschriften:

Prof. Dr. Achim Goerres
Lehrstuhl für Empirische Politikwissenschaft
Institut für Politikwissenschaft
Universität Duisburg-Essen
Lotharstraße 65
47057 Duisburg
E-Mail: achim.goerres@uni-due.de

Kärger, Caroline, M.A.
Lehrstuhl für Internationale Beziehungen und Entwicklungspolitik
Institut für Politikwissenschaft
Universität Duisburg-Essen
Lotharstraße 65
47057 Duisburg
E-Mail: caroline.kaerger@uni-due.de

Dr. rer. pol. Daniel Lambach
Vertretungsprofessor für Internationale Beziehungen
Lehrstuhl für Internationale Beziehungen und Entwicklungspolitik
Institut für Politikwissenschaft
Universität Duisburg-Essen
Lotharstraße 65
47057 Duisburg
E-Mail: daniel.lambach@uni-due.de

Anhang

Tabelle: Übersicht über Regressionsergebnisse mit Prüfergebnis nach Punkten für eine Frage als abhängige Variable

Modell	R ² in %	Kontextvariable für FC-Gruppe	p-Wert eines Modellvergleichstests	Punktschätzung FC-Gruppe (p-Wert einseitiger Test, dass Koeffizient > 0)	Ausprägung der Kontextvariable	Punktschätzung FC-Gruppe (p-Wert einseitiger Test)	Ausprägung der Kontextvariable
Grundmodell 1: Geschlecht, Gesamtzahl erreichter Punkte, erster oder zweiter Prüfetermin, Prüfversuch des Prüflings, MC-Frage oder offen, Lernziel Verstehen oder Anwendung, FC-Gruppe oder nicht	46,4	Nein	-	0,02 (0,36)	-	-	-
Grundmodell 2: Grundmodell 1 ohne Anwendung und ohne MC, dafür mit Schwierigkeit der Prüffrage	43,0						
Grundmodell 1 plus Interaktion FC-Gruppe mit Prüfversuch	46,4	Prüfversuch	0,14				
Grundmodell 1 plus Interaktion FC-Gruppe mit Zweittermin	46,4	Zweittermin	0,37				
Grundmodell 1 plus Interaktion FC-Gruppe mit MC-Frage oder nicht	46,4	MC-Frage	0,39				
Grundmodell 1 plus Interaktion FC-Gruppe mit Anwendung vs. Verstehen	46,7	Anwendung vs. Verstehen	0,002	-0,16 (0,98)	Frage Verstehen als Lernziel	0,21 (0,01)	Frage Anwendung als Lernziel
Grundmodell 2 plus Interaktion FC-Gruppe mit Schwierigkeit	43,1	Schwierigkeit der Prüffrage	0,25				

Grundmodell 1 plus Interaktion FC-Gruppe mit Gesamtzahl erreichter Punkte	46,4	Gesamtzahl erreichter Punkte	0,31				
---	------	------------------------------	------	--	--	--	--

Anmerkung: OLS-Regression mit robusten Standardfehlern. Abhängige Variable: Punktergebnis bei einer Frage, pro Prüfling mindestens vier Beobachtungen, 1.508 Beobachtungen. Kollinearität und Autokorrelation überprüft. Heteroskedastizität mit robusten Standardfehlern berücksichtigt.