

Diskussionspapier des
Instituts für Organisationsökonomik

8/2020

Sind Klausuren überflüssig?

Zum Zusammenhang zwischen
PISA-Ergebnisse, Rechtschreibung und Noten

Alexander Dilger

Discussion Paper of the
Institute for Organisational Economics

**Diskussionspapier des
Instituts für Organisationsökonomik
8/2020**

August 2020

ISSN 2191-2475

Sind Klausuren überflüssig?

Zum Zusammenhang zwischen PISA-Ergebnissen, Rechtschreibung und Noten

Alexander Dilger

Zusammenfassung

Bereits 2002 wurden bei einer BWL-Klausur zwei PISA-Aufgaben gestellt und die Rechtschreibung überprüft, was wiederum mit den Noten bzw. Punkten der Klausur korreliert wurde. Es gibt interessante Ergebnisse und Zusammenhänge, die jedoch nicht völlig eindeutig sind.

JEL Codes: A22, C90, I21, I23, M00

Are Exams Superfluous?

On the Relationship between PISA Results, Orthography and Grades

Abstract

Already in 2002, a written examination in business administration included two PISA tasks and a check of orthography, which in turn were correlated with the grades or points of the examination. There are interesting results and relationships. However, they are not completely unambiguous.

Im Internet unter:

http://www.wiwi.uni-muenster.de/io/forschen/downloads/DP-IO_08_2020

Westfälische Wilhelms-Universität Münster
Institut für Organisationsökonomik
Scharnhorststraße 100
D-48151 Münster

Tel: +49-251/83-24303 (Sekretariat)
E-Mail: io@uni-muenster.de
Internet: www.wiwi.uni-muenster.de/io

Sind Klausuren überflüssig?

Zum Zusammenhang zwischen PISA-Ergebnissen, Rechtschreibung und Noten*

1. Einleitung

Durch die (erste) PISA-Studie im Jahr 2000¹ wurde die Öffentlichkeit insbesondere in Deutschland aufgeschreckt.² Denn dabei stellte sich heraus, dass fünfzehnjährige Neuntklässler bei Aufgaben zur Mathematik, dem Lese- oder auch naturwissenschaftlichen Verständnis im internationalen Vergleich eher bescheiden abschnitten. Was viele schon geahnt hatten, wurde damit eindrucksvoll untermauert, wenn auch nicht zweifelsfrei bewiesen: Die deutschen Schulen sind nicht überragend gut.

Erschreckender noch als das schlechte Abschneiden von Schülern bei altersadäquaten Aufgaben dürfte jedoch die Tatsache sein, dass auch viel ältere, vermeintlich besser gebildete Personen die PISA-Aufgaben nicht lösen konnten. So schimpften Politiker über die Schüler und Lehrer, nur um dann selbst bei einfachen Beispielaufgaben zu versagen. Studierende mit Abitur konnten häufig ebenfalls entsprechende Aufgaben nicht lösen (siehe den 2. Abschnitt dieses Beitrags).

Das schlechte Abschneiden in einem Test ist eine Sache, der Zusammenhang zwischen verschiedenen Leistungsindikatoren eine andere, die in diesem Beitrag näher untersucht werden soll. Vielleicht sind die PISA-Aufgaben ohne Belang, wie auch Intelligenztests häufig vorgeworfen wird, nur die Fähigkeit zum Lösen von Intelligenztests zu testen.³ Wenn jedoch verschiedenartige Indikatoren stets zu ähnlichen Ergebnissen führen, scheint es grundlegende Fähigkeiten von Bedeutung zu geben. Wer dann PISA-Aufgaben nicht lösen kann, ist deshalb trotzdem kein schlechter Mensch oder ohne Fähigkeiten in anderen, z. B. praktischen Bereichen, hat jedoch offensichtlich ein Bildungsdefizit.

Zwei alternative Indikatoren zu den PISA-Ergebnissen sind etwa eine geringe Quote an Rechtschreibfehlern und gute fachliche Noten. Es ist zu vermuten, dass eine positive Korrela-

* Dieses Diskussionspapier ist die schriftliche Version eines Vortrags, der bereits beim 6. Workshop der wissenschaftlichen Kommission Hochschulmanagement im Verband der Hochschullehrer für Betriebswirtschaft e. V. am 8. Februar 2003 in Innsbruck gehalten wurde. Die Einleitung entspricht im Wesentlichen dem Abstract und der übrige Text den Präsentationsfolien, während die Fußnoten und Literaturangaben aktuell ergänzt wurden.

¹ Siehe z. B. Artelt et al. (2001) und OECD/UNESCO Institut for Statistics (2003).

² Siehe z. B. Darnstädt et al. (2001).

³ Vgl. Lossau (2002).

tion zwischen allen drei Größen besteht, weil in jedem Fall grundlegende Fähigkeiten wie Intelligenz oder Bildungsanstrengungen beteiligt sind. Alternativ ist es natürlich denkbar, dass eine substitutive Beziehung vorliegt, also z. B. sprachliche Fähigkeiten negativ mit mathematischen korreliert sind. Schließlich könnten die Indikatoren unkorreliert sein, so dass jede relevante Leistungsdimension für sich getestet werden müsste. Bei hoher Korrelation ist es hingegen fraglich, ob die Vielzahl an Klausuren und Prüfungen, die jeder Schüler und Student im Laufe der Zeit über sich ergehen lassen muss, nicht eigentlich überflüssig ist. Wenn (fast) immer dieselben Personen gut, andere hingegen meistens schlecht beurteilt werden, ist eine ständige Wiederholung dieses Urteils Verschwendung oder sogar Schikane. Gelegentlich geforderte⁴ Universitätseingangsprüfungen sind ebenfalls fraglich, wenn sie zu kaum anderen Ergebnissen führen als eine Analyse der Abiturzeugnisse. Aus entsprechenden Gründen (und wegen fallenden Bewerberzahlen) wurde beispielsweise der bundesweite Medizinerterst 1997 abgeschafft.

Konkret wird zur Überprüfung der Korrelation zwischen PISA-Ergebnissen, Rechtschreibung und Noten ein Datensatz verwendet, der sich auf die im Sommersemester 2002 an der Ernst-Moritz-Arndt-Universität Greifswald durchgeführte Grundstudiumsklausur „Grundzüge der BWL I“ stützt. Es liegen zu 105 Studierenden Daten in allen drei Bereichen vor, die mit statistischen Methoden ausgewertet wurden. Die falsifizierbare Vermutung lautete dabei, dass die Korrelation zwischen den drei Größen sehr hoch ist. In dem Fall könnte über eine Reduktion der Anzahl an Klausuren nachgedacht werden.⁵

Im folgenden zweiten Abschnitt werden die Ergebnisse zu zwei PISA-Aufgaben in der genannten Klausur vorgestellt, im dritten Abschnitt die Analyse der Rechtschreibfehler bei einer weiteren Aufgabe und im vierten Abschnitt die Noten bzw. eigentlich Klausurpunkte. Im fünften Abschnitt wird der Zusammenhang zwischen diesen drei Größen genauer betrachtet. Dieser Beitrag schließt mit einem kurzen Fazit im sechsten Abschnitt.

2. PISA-Ergebnisse

In die analysierte Klausur wurden die beiden nachfolgenden PISA-Aufgaben (aus dem mathematischen Teil) eingefügt. Für jede dieser Aufgabe wurden maximal sieben Punkte verge-

⁴ Siehe z. B. Meyer/Müller-Böling (1996).

⁵ Die Umstellung auf Bachelor- und Masterstudiengänge führte hingegen zu noch deutlich mehr Klausuren, wobei Greifswald einer der letzten Hochschulstandorte ist, wo noch ein Diplomstudiengang BWL existiert.

ben, wobei insgesamt pro Minute empfohlener Bearbeitungszeit ein Punkt erreichbar war. Teillösungen wurden anteilig bepunktet, gegebenenfalls mit dem siebten Punkt als Zuschlag bei vollständig richtiger Lösung. Wie bei der Original-PISA-Untersuchung musste der Lösungsweg nicht angegeben werden, nur das Ergebnis zählte.

„Aufgabe I: Wie können Sie einen Geldbetrag von genau 31 Cent hinlegen, wenn Sie nur 10-Cent-, 5-Cent- und 2-Cent-Münzen zur Verfügung haben? Geben Sie alle Möglichkeiten an!“

„Aufgabe II: Sie werden beauftragt, einen neuen Satz Münzen zu entwerfen. Alle Münzen sollen rund und silberfarben sein, aber verschiedene Durchmesser haben. Forscher haben herausgefunden, dass ein idealer Satz von Münzen folgende Anforderungen erfüllt:

- Der Durchmesser der Münzen sollte nicht kleiner als 15 mm und nicht größer als 45 mm sein.
- Ausgehend von einer Münze muss der Durchmesser der nächsten Münze mindestens 30 % größer sein.
- Die Prägemaschine kann nur Münzen herstellen, deren Durchmesser in Millimeter ganzzahlig ist.

Entwerfen Sie einen Satz von Münzen, der die o. g. Anforderungen erfüllt. Beginnen Sie mit einer 15-Millimeter-Münze. Der Satz sollte so viele Münzen wie möglich enthalten.“

36,2 % der Studierenden lösten Aufgabe I vollständig richtig, 49,5 % erreichten alle sieben möglichen Punkte von Aufgabe II,⁶ beide Aufgaben zur Gänze gelangen 21 %. 13,3 % erzielten null Punkt bei Aufgabe I (9,5 % versuchten sich gar nicht erst an der Aufgabe), 26,7 % erging es so bei Aufgabe II (21,0 % ließen diese Aufgabe aus). Im Durchschnitt wurden 4,56 Punkte für Aufgabe I und 4,05 Punkte für Aufgabe II vergeben.⁷ Da die Aufgaben für fünfzehnjährige Neuntklässler konzipiert sind, erscheint das Ergebnis für Studierende mit Abitur nicht berauschend. Bei deutschen Schülern ist das Ergebnis für diese Aufgaben allerdings noch wesentlich schlechter.⁸

⁶ An der Universität Leipzig konnten 57 % von BWL-Studierenden im Grundstudium diese Aufgabe richtig lösen (siehe Leffers 2002).

⁷ Das liegt immerhin relativ betrachtet deutlich über dem Durchschnitt für die übrigen Aufgaben (siehe den 4. Abschnitt). Diejenigen, die die Aufgaben bearbeitet hatten, erzielten im Schnitt 5,04 Punkte (Aufgabe I) bzw. 5,12 Punkte (Aufgabe II).

⁸ Laut Der Spiegel (2001) konnten nur 1,3 Prozent der deutschen Schülerinnen und Schüler die Aufgabe I vollständig richtig beantworten.

3. Rechtschreibung

Am Anfang unter den allgemeinen Erläuterung stand zu einer Textaufgabe (mit maximal zehn Punkten):

„Zumindest Aufgabe III ist in ganzen Sätzen zu beantworten (bei mindestens vier Sätzen und keinem sprachlichen Fehler gibt es einen Extrapunkt).“

91 Personen (86,7 %) haben diese Aufgabe bearbeitet, drei davon (3,3 % von 91) haben den Extrapunkt erzielt, drei weitere einen halben (für einen „halben“ bzw. geringfügigen Fehler). Die sechs mit einem oder zumindest einem halben Extrapunkt schrieben durchschnittlich 7,3 Sätze, während der Gesamtschnitt bei 8,6 Sätzen (Maximum 26) lag. 7,7 % schrieben weniger als vier Sätze, eine davon fehlerfrei.

25,3 % verwendeten eindeutig die alte Rechtschreibung, 48,4 % eindeutig die neue, während die restlichen 26,4 % sich nicht eindeutig zuordnen ließen. Im Durchschnitt gab es 0,80 Fehler pro Satz (Maximum 3,33), davon 0,32 Zeichensetzungsfehler (Maximum 1,33). Bei alter Rechtschreibung lag die Fehlerquote signifikant höher bei 1,01, für Zeichensetzungsfehler bei 0,49. Mit alter Rechtschreibung erzielte keiner auch nur einen halben Extrapunkt.⁹

4. Noten

Die Klausur umfasste neben den zwei PISA-Aufgaben und der Textaufgabe im Teil „Personal und Organisation“ noch als Aufgabe IV eine Multiple-Choice-Aufgabe mit folgender Erläuterung:

„Pro Frage ist genau eine Antwort richtig. Bitte stets einen Buchstaben ankreuzen (andernfalls wird die Frage nicht gewertet). Es werden bereits vier richtige Antworten vorausgesetzt (Erwartungswert bei zufälligem Ankreuzen), ab der fünften richtig beantworteten Frage gibt es jeweils 2 Punkte pro richtiger Antwort, also maximal 16 Punkte für Aufgabe IV.“

Außerdem hatten 101 der 105 Studierenden zusätzlich einen Marketing- und einen Produktionsteil mit ebenfalls jeweils geplant 40 Minuten und maximal 40 Punkten. Im Marketingteil wurden durchschnittlich 12,85 Punkte erzielt, in Produktion 14,81 Punkte und in Personal und

⁹ Anzumerken ist, dass die neue Rechtschreibung für sich genommen leichter ist, weil es mehr Wahlmöglichkeiten gibt, insbesondere auch bei der Zeichensetzung.

Organisation 18,42 Punkte,¹⁰ insgesamt durchschnittlich 46,08 Punkte. 120 Punkte waren möglich, das erreichte Maximum lag bei 87,5. Die Bestehensgrenze wurde entsprechend herabgesetzt, um die Durchfallquote zu reduzieren.¹¹

In Aufgabe III (der Textaufgabe im Teil Personal und Organisation) wurden durchschnittlich 5,04 Punkte erzielt (ohne Extrapunkte), in Aufgabe IV (der Multiple-Choice-Aufgabe) 4,63. 19 % haben eine, mehrere oder alle MC-Fragen gar nicht beantwortet oder mindestens einmal zwei Antworten angekreuzt, was ebenfalls zu 0 Punkten führte, weshalb im Vergleich dazu reines Raten jeweils eine dominante Strategie gewesen wäre.

Die Korrelation nach Pearson zwischen den Punkten in Marketing und Produktion beträgt 0,44 (signifikant auf dem 1 %-Niveau). Zwischen den Marketing- und Personalpunkten ist der betreffende Wert 0,32 (signifikant auf dem 1 %-Niveau), bei Produktion und Personal 0,27 (1 %-Niveau). Die Korrelation der Gesamtpunktzahl zu denen der Teilklausuren ist noch wesentlich höher (zwischen 0,69 und 0,78), aber z. T. bereits analytisch bzw. arithmetisch bedingt. Entsprechendes gilt für die Personalpunkte und die Punkte der dazugehörigen Aufgaben (zwischen 0,52 und 0,67).

Die Korrelation der Punkte der beiden PISA-Aufgaben beträgt 0,33 (signifikant auf dem 1 %-Niveau). Die Korrelation der Punkte für die Multiple-Choice-Aufgabe mit denen der 2. PISA-Aufgabe beträgt 0,22 (5 %-Niveau), mit der Textaufgabe 0,29 (1 %-Niveau). Die Marketingpunkte sind 0,48 (1 %-Niveau) mit denen der Textaufgabe korreliert, 0,30 (1 %-Niveau) mit den Multiple-Choice-Punkten. Für die Produktionspunkte lauten die entsprechenden Werte 0,22 (5 %-Niveau) und 0,36 (1 %-Niveau). Die Punkte für Marketing sowie Produktion sind (insignifikant) negativ mit denen der 1. PISA-Aufgabe korreliert.

5. Zusammenhang der Indikatoren

Wenn die 1. PISA-Aufgabe perfekt gelöst wurde, dann gibt es insignifikant mehr Punkte in Marketing und in Produktion. Dasselbe gilt, wenn die 2. PISA-Aufgabe vollständig gelöst wurde. Beide PISA-Aufgaben ganz richtig führt zu insignifikant weniger Marketingpunkten.

¹⁰ Bezogen auf die 101 Klausuren mit allen drei Teilen. Die vier Klausuren mit nur dem Teil Personal und Organisation hatten ein durchschnittliches Ergebnis von 15,5 Punkten.

¹¹ Diese war trotzdem ziemlich hoch, was jedoch nicht Gegenstand dieses Beitrags ist.

Die Teilnehmerinnen (alle weiblich) mit einem oder einem halbem Extrapunkt für gute Rechtschreibung schneiden in der Gesamtklausur (um durchschnittlich 10 Punkte) und in allen Teilklausuren wie auch Aufgaben (außer der 2. PISA-Aufgabe) schlechter ab. Die Quote an allgemeinen Fehlern und Zeichensetzungsfehlern ist negativ korreliert mit der Gesamtpunktzahl und positiv korreliert mit der Text- sowie der 2. PISA-Aufgabe. Die Verwendung alter oder neuer Rechtschreibung hat für sich genommen keinen Einfluss auf die Punktzahl. Bei alter Rechtschreibung und unterdurchschnittlich wenigen Zeichensetzungsfehlern ist jedoch die Punktzahl in den Teilklausuren, der Gesamtklausur und der Textaufgabe höher (in der 1. PISA-Aufgabe ist sie niedriger). Bei neuer oder unklarer Rechtschreibung und unterdurchschnittlich vielen Fehlern sind die Punktzahlen insgesamt, im Personal- und Organisationssteil sowie für die beiden PISA-Aufgaben signifikant höher.

Es ließe sich eine Vielzahl weiterer Einzelbefunde berichten, aber die grobe Linie dürfte deutlich geworden sein. Erwähnenswert ist vielleicht noch, dass Frauen (52 %) im Marketing und Männer in Produktion besser beurteilt werden. Studierende im BWL-Diplomstudiengang (72,4 %) schneiden signifikant besser ab als Nebenfächler in der Gesamtklausur, in Produktion und der 2. PISA-Aufgabe. Die Immatrikulationsdauer hat einen W-förmigen Einfluss auf die Punkte (außer in Personal, wo sie nur fallen): Die jüngsten Studierenden sind am besten und die Leistung der Nachzügler schwankt (wohl wegen adverser Selektion und erhöhter Motivation).

6. Fazit

Es gibt einen Zusammenhang zwischen verschiedenen Indikatoren für Leistung, Intelligenz oder auch Vorwissen. Dieser ist jedoch nicht beliebig robust und manchmal sogar negativ. Deshalb muss die Titelfrage nicht zwangsläufig verneint werden. Eher ist das Datenmaterial für die Beantwortung dieser Frage nicht optimal geeignet. Denn innerhalb derselben Klausur konkurrieren ihre Teile und einzelnen Aufgaben um die knappe Zeit. Aussagekräftiger wäre vielleicht eine Zeitreihe mehrerer Klausuren. Ein Bemühen um Rechtschreibung lohnt sich z. B. nicht, während ganz ohne zumindest eine richtige PISA-Aufgabe die Bestehensquote signifikant um 34,5 Prozentpunkte sinkt. Hintergrund könnte sein, dass BWL durchaus quantitative Anteile hat, die ein grundlegendes mathematisches Verständnis voraussetzen.

Literatur

- Artelt, Cordula/Baumert, Jürgen/Klieme, Eckhard/Neubrand, Michael/Prenzel, Manfred/Schiefele, Ulrich/Schneider, Wolfgang/Schümer, Gundel/Stanat, Petra/Tillmann, Klaus-Jürgen/Weiß, Manfred (Hrsg.) (2001): „PISA 2000: Zusammenfassung zentraler Befunde“, Max-Planck-Institut für Bildungsforschung, Berlin. Online unter https://www.pisa.tum.de/fileadmin/w00bgi/www/Berichtsbaende_und_Zusammenfassungen/Zusammenfassung_PISA_2000
- Darnstädt, Thomas/Koch, Julia/Mohr, Joachim/Neumann, Conny/Wensierski, Peter (2001): „Die Pisa-Analyse: Sind deutsche Schüler doof?“, Der Spiegel, 13.12.2001. Online unter <https://www.spiegel.de/lebenundlernen/schule/die-pisa-analyse-sind-deutsche-schueler-doof-a-172357.html>
- Der Spiegel (2001): „Testaufgaben: Was Schüler bei ‚Pisa‘ wissen müssen“, Der Spiegel, 04.12.2001. Online unter <https://www.spiegel.de/lebenundlernen/schule/testaufgaben-was-schueler-bei-pisa-wissen-muessen-a-171102.html>
- Leffers, Jochen (2002): „Blamable Rechenleistungen: Selbst BWL-Studenten scheitern an Pisa-Aufgaben“, Der Spiegel, 19.04.2002. Online unter <https://www.spiegel.de/lebenundlernen/schule/blamable-rechenleistungen-selbst-bwl-studenten-scheitern-an-pisa-aufgaben-a-192528.html>
- Lossau, Norbert (2002): „Intelligenz ist das, was man mit IQ-Tests misst“, Die Welt 27.07.2002. Online unter <https://www.welt.de/print-welt/article402368/Intelligenz-ist-das-was-man-mit-IQ-Tests-misst.html>
- Meyer, Hans Joachim/Müller-Böling, Detlef (1996): „»Leipziger Erklärung«“, in: dieselben (Hrsg.): „Hochschulzugang in Deutschland: Status quo und Perspektiven“, Verlag Bertelsmann Stiftung, Gütersloh, S. 11-14.
- OECD/UNESCO Institute for Statistics (2003): „Literacy Skills for the World of Tomorrow: Further Results from PISA 2000“, Paris. Online unter https://www.oecd-ilibrary.org/literacy-skills-for-the-world-of-tomorrow_5lmqcr2k2mxx.pdf?itemId=%2Fcontent%2Fpublication%2F9789264102873-en&mimeType=pdf

Diskussionspapiere des Instituts für Organisationsökonomik

Seit Institutsgründung im Oktober 2010 erscheint monatlich ein Diskussionspapier. Im Folgenden werden die letzten zwölf aufgeführt. Eine vollständige Liste mit Downloadmöglichkeit findet sich unter <http://www.wiwi.uni-muenster.de/io/de/forschen/diskussionspapiere>.

- DP-IO 8/2020** Sind Klausuren überflüssig?
Zum Zusammenhang zwischen PISA-Ergebnissen, Rechtschreibung und Noten
Alexander Dilger
August 2020
- DP-IO 7/2020** No Home Bias in Ghost Games
Alexander Dilger/Lars Vischer
Juli 2020
- DP-IO 6/2020** The Advances of Community Cloud Computing in the Business-to-Business-Buying Process
Nicolas Henn/Todor S. Lohwasser
Juni 2020
- DP-IO 5/2020** Wirtschaftsethische Überlegungen zum Klimawandel
Alexander Dilger
Mai 2020
- DP-IO 4/2020** Meta-Analyzing the Relative Performance of Venture Capital-Backed Firms
Todor S. Lohwasser
April 2020
- DP-IO 3/2020** From Signalling to Endorsement
The Valorisation of Fledgling Digital Ventures
Milan Frederik Klus
März 2020
- DP-IO 2/2020** Internet-Publikationen gehört die Zukunft
Alexander Dilger
Februar 2020
- DP-IO 1/2020** Ist die Vergütung im Fußball geschlechtergerecht?
Alexander Dilger/Katrin Scharfenkamp
Januar 2020
- DP-IO 12/2019** Von der Promotion zur Professur
Fünfzehn Tipps für den wissenschaftlichen Nachwuchs
Alexander Dilger
Dezember 2019
- DP-IO 11/2019** The Digital Leader
What One Needs to Master Today's Organisational Challenges
Milan Frederik Klus/Julia Müller
November 2019
- DP-IO 10/2019** 9. Jahresbericht des Instituts für Organisationsökonomik
Alexander Dilger/Milan Frederik Klus
Oktober 2019
- DP-IO 9/2019** The Impact of Institutions on Venture Capital
How Transaction Costs, Uncertainty, and Change Affect New Ventures
Felix Hoch/Todor S. Lohwasser
September 2019



Herausgeber:
Prof. Dr. Alexander Dilger
Westfälische Wilhelms-Universität Münster
Institut für Organisationsökonomik
Scharnhorststr. 100
D-48151 Münster

Tel: +49-251/83-24303

Fax: +49-251/83-28429

www.wiwi.uni-muenster.de/io

