

## Ein Märchen aus tausend und einer Nacht? Kommentar zu dem Artikel von Frederike Wuermelings „Paßt die Türkei zur EU und die EU zu Europa?“

Kai Arzheimer\*

*Abstract:* In ihrem in Heft 2/2007 der Kölner Zeitschrift erschienen Artikel argumentiert Frederike Wuermeling, daß die vergleichsweise niedrige Zustimmung der Türken zu den EU-Prinzipien kausal auf das niedrige Bruttosozialprodukt und den hohen Muslimanteil zurückzuführen ist. Tatsächlich weist Wuermelings Analyse jedoch eine Vielzahl von methodischen und theoretischen Defiziten auf. Die Kernaussagen ihres Beitrages sind deshalb durch die Daten nicht gedeckt.

*Keywords:* Türkei, Europäische Union, Europäische Wertestudie, Religionsfreiheit, Demokratie, Gleichberechtigung, Rechtsstaatlichkeit, Mehrebenenanalyse

*Abstract:* In a recent contribution to this journal, Frederike Wuermeling argues that there is a causal link between the rather low level of support for the EU's core principles amongst the Turkish public on the one hand and the country's low GDP and its large muslim population on the other hand. Her analysis, however, suffers from a whole host of serious methodological and theoretical flaws. Therefore, the core claims of her paper are untenable.

*Keywords:* Turkey, European Union, European Values Survey, freedom of religion, democracy, equal rights, constitutional legality, multilevel analysis

Preprint. Die endgültige Fassung ist im Heft 1/2008 der Kölner Zeitschrift für Soziologie und Sozialpsychologie erschienen:  
<http://dx.doi.org/10.1007/s11577-008-0006-9>

---

\*Department of Government, University of Essex, Wivenhoe Park, Colchester CO4 3SQ. Ich danke Sarah Kirschmann (Mainz) für die Beschaffung der Makro-Daten und Ersin Oezsahin (Konstanz) für die Übersetzung der türkischen Items aus dem EVS-Fragebogen sowie für zahlreiche wertvolle Hinweise und Anregungen.

Replikationsdaten für diesen Beitrag: Kai Arzheimer, 2008, Replication data for: Ein Märchen aus tausend und einer Nacht ... Demokratische Einstellungen in der Türkei, [hdl:1902.1/11171](https://hdl.handle.net/1902.1/11171) UNF:3:eHkqqsOl69GCE2ap0dRWMw==

## I. Einleitung

Frederike Wuermelings in Heft 2/2007 dieser Zeitschrift erschienener Beitrag über die Verankerung der EU-Grundprinzipien Religionsfreiheit, Demokratie, Gleichberechtigung und Rechtsstaatlichkeit in der türkischen Bevölkerung ([Wuermeling 2007](#)) wird ohne Zweifel die bislang recht verhaltene *wissenschaftliche* Debatte über einen möglichen Beitritt der Türkei zu Europäischen Union beleben. Eher ungewöhnlich ist allerdings die Resonanz, die der Beitrag bereits jetzt außerhalb der Fachöffentlichkeit gefunden hat. Noch bevor das Heft an die Abonnenten ausgeliefert wurde, wurde die Kernaussage des Papiers – die Türkei als islamisch geprägter Staat sei derzeit noch nicht reif für einen EU-Beitritt – von Massenmedien wie dem Kölner Stadtanzeiger, der österreichischen „Presse“, dem „Neuen Volksblatt“ und sogar der Deutschen Welle aufgegriffen und verbreitet. In einer zweiten Rezeptionswelle erreichten – wiederum vor Erscheinen des Heftes – die eingängigsten Thesen aus der Berichterstattung von Stadtanzeiger und „Presse“ sowie aus der von der Universität Köln herausgegebenen Pressemitteilung die einschlägigen Diskussionsforen und „blogs“ im Internet.<sup>1</sup> Was man dort schon immer zu wissen glaubte – „je höher der Anteil an Muslimen, desto geringer das Akzeptanzniveau“ demokratischer Werte – wird nun quasi mit dem Gütesiegel der Kölner Soziologie verbreitet.

Bedenklich ist dies deshalb, weil Würmelings Artikel einerseits substantiell wenig Neues enthält – die von ihr verwendeten Daten wurden 1999-2001 erhoben (siehe Abschnitt II.3.) und bereits mehrfach unter genau diesem Aspekt ausgewertet ([Gerhards 2004, 2007](#)) – und andererseits in den Teilen, in denen er von Jürgen Gerhards Beiträgen abweicht, erheblich theoretische und methodische Mängel aufweist, die im nächsten Abschnitt skizziert werden.

## II. Probleme

---

<sup>1</sup>Vgl. statt vieler „Jihad Watch Deutschland“ [http://fredalanmedforth.blogspot.com/2007\\_07\\_01\\_archive.html](http://fredalanmedforth.blogspot.com/2007_07_01_archive.html), „Politically Incorrect“ <http://www.politicallyincorrect.de/2007/07/tuerkei-nur-ein-drittel-fuer-gleichberechtigung/> und „Weckstube – gegen die Islamisierung Deutschlands“ <http://weckstube.info/archives/26>.

## 1. Theoretische Basis und Hypothesen

Der Titel von Frederike Wuermelings Studie – „Paßt die Türkei zur EU und die EU zu Europa“ – ist aus mehreren Gründen irreführend. Denn erstens wird der zweite Teil der im Titel aufgeworfenen Frage im Text nur am Rande angesprochen. Zweitens läßt sich die Doppelfrage nicht im eigentlichen Sinne beantworten, da es keinen Maßstab dafür gibt, wieviel Zustimmung auf den von Wuermeling konstruierten Skalen gegeben sein muß, damit von einer Passung gesprochen werden kann. Drittens schließlich steht ein Vergleich der Ländermittelwerte untereinander oder mit einem externen Maßstab gar nicht im Zentrum des Artikels.

Von den bislang vorliegenden Analysen der Daten aus der European Values Study (EVS) unterscheidet sich der Beitrag von Wuermeling vielmehr vor allem dadurch, daß die Autorin ein Mehr-Ebenen-Modell spezifiziert, mit dessen Hilfe sie eine Reihe von *Hypothesen über das Zustandekommen dieser Länderunterschiede* überprüfen will. Diese Hypothesen wiederum gewinnt Wuermeling aus zwei sehr weitgefaßten Erklärungsansätzen. Unter dem Rubrum „Kultureller Kontext“ diskutiert Wuermeling unter Rückgriff auf [Weber \(2006, zuerst 1904/05\)](#) und [Huntington \(1996\)](#) die Frage, ob die EU-Prinzipien in der Türkei deshalb abgelehnt werden, weil es sich um ein islamisch geprägtes Land handle ([Wuermeling 2007: 189-193](#)). Diese Überlegungen münden in der Hypothese „je größer der Anteil der Muslime in einem Land, desto geringer ist das Akzeptanzniveau der EU-Grundprinzipien“ ([Wuermeling 2007: 192](#)). Eine zweite Gruppe von Überlegungen, die in einem losen Zusammenhang mit [Inglehart \(1997\)](#) und [Bell \(1973\)](#) stehen, faßt Wuermeling unter der Überschrift „Modernisierungsgrad“ zusammen. Hier lautet die Kernhypothese „Je höher der ökonomische Modernisierungsgrad eines Landes, desto höher das Akzeptanzniveau der EU-Grundprinzipien“ ([Wuermeling 2007: 194](#)).

Wuermelings theoretische Einlassungen werden jedoch ihren Gewährsleuten nicht gerecht. So ist Ingleharts Ansatz trotz aller Modifikationen und obwohl Inglehart sich in den letzten Jahren schwerpunktmäßig mit Makro-Analysen befaßt hat (u. a. [Inglehart und Baker 2000](#)), im Kern dem methodologischen Individualismus verpflichtet: Individuelle Wertorientierungen werden auf individuelle Sozialisationserfahrungen während einer prägenden Jugendphase zurückgeführt. Daß etwa beispielsweise die aktuelle Wirtschaftslage einen Einfluß auf Wertorientierungen haben könnte, wurde von Inglehart deshalb stets energisch bestritten.

Eine präzise Erklärung dafür, warum ein höheres Bruttoinlandsprodukt (BIP) *im Jahr der Umfrage* (und nicht etwa das individuelle oder Haushaltseinkommen) unter Kontrolle diverser Mikro-Variablen zu einer stärkeren Befürwortung der EU-Prinzipien führen soll, bleibt Wuermeling schuldig, zumal sie selbst davon ausgeht, daß die von ihr analysierten Einstellungen nur mit einer „Zeitverzögerung“ ([Wuermeling 2007: 193](#)) auf veränderte Randbedingungen reagieren. Wenn überhaupt, dann sollten

die wirtschaftlichen Bedingungen in der formativen Phase (eine Makro-Variable, die mit der Kohortenzugehörigkeit variiert) oder die Wirtschaftslage in der vorangegangenen Dekade einen Einfluß auf die Akzeptanz der EU-Prinzipien haben.

Selbst dann stellt sich allerdings die Frage, ob das BIP, d. h. die schiere Leistungsfähigkeit einer Volkswirtschaft, ein geeigneter Indikator für die von Inglehart und Bell gemeinten Prozesse sein kann. Nicht umsonst sprechen die von Wuermeling angeführten Autoren von „Post-Modernisierung“ beziehungsweise „Post-Industrialisierung“ als einer neuen Phase des ökonomischen und gesellschaftlichen Wandels, die durch einen (partiellen) Niedergang des industriellen Sektors, den Übergang zu „Post-Fordistischen“ Produktionsstrukturen und einen Bedeutungsgewinn (gehobener) Dienstleistungen gekennzeichnet ist (Inglehart und Baker 2000: 22). Vor diesem Hintergrund erscheint es geradezu absurd, Inglehart und Bell als Kronzeugen für eine wie auch immer geartete kausale Wirkung des gegenwärtigen Bruttoinlandsproduktes auf individuelle demokratische Einstellungen anzurufen.

Ähnlich unverbindlich ist Wuermelings Rekurs auf Webers Überlegungen zum Zusammenhang zwischen protestantischer Ethik und kapitalistischen Produktionsstrukturen. Wie der Rekonstruktionsversuch von Coleman (1994) zeigt, handelt es sich bei der „Protestantischen Ethik“ im Kern ebenfalls um eine Mehr-Ebenen-Erklärung, die sich in stark verkürzter Weise so darstellen läßt: (1) Protestantisch-Calvinistische Werte und Normen der „innerweltlichen Askese“ werden von der Kirche und anderen Institutionen propagiert, (2) die entsprechenden Einstellungen werden von den Akteuren internalisiert und (3) gewinnen so Einfluß auf deren Verhalten im Erwerbsleben, was (4) durch die Aggregation individueller Handlungen zur Entstehung einer leistungsfähigen, kapitalintensiven Volkswirtschaft führt, so daß (5) auf der Makro-Ebene ein Zusammenhang zwischen Protestantismus und Kapitalismus besteht (Coleman 1994: 8).

Von diesem Argument übernimmt Wuermeling lediglich die erste Hälfte, derzufolge Religionen Einstellungen auch in solchen Bereichen beeinflussen können, die in keinem unmittelbaren Zusammenhang mit den Glaubenslehren stehen. Wie und auf welcher Ebene aber soll „der Islam“ politische Einstellungen beeinflussen? Ist hier tatsächlich der nationale politische Kontext entscheidend, oder werden Einstellungen nicht eher innerhalb von Familien, (religiösen) Gemeinden oder Regionen geformt? Werden räumliche Kontexte im 21. Jahrhundert nicht in für die Fragestellung Wuermelings entscheidender Weise von transnationalen Medienkontexten (etwa den vor allem von Saudi-Arabien finanzierten religiös-politischen Satellitenprogrammen) überlagert? Ist schließlich vor dem Hintergrund der strengen islamischen Lehren ernsthaft davon auszugehen, daß „der Islam“ als kultureller Faktor die über ein Korruptions-Item gemessene Akzeptanz der Rechtsstaatlichkeit beeinträchtigt? Diese und ähnliche Fragen werden von Wuermeling nicht einmal angerissen.

Da Wuermeling die *individuelle* Zugehörigkeit zu einer Religionsgemeinschaft nicht

berücksichtigt, ist es überdies unmöglich zu entscheiden, ob der von ihr gefundene Effekt des Moslemanteils (1) ein echter Kontexteffekt ist (muslimisch geprägte Gesellschaften reduzieren die individuelle Akzeptanz von Demokratie, Rechtsstaatlichkeit etc.), (2) durch die Komposition der Stichproben zustande kommt (in überwiegend muslimischen Gesellschaften ist die Wahrscheinlichkeit größer, daß ein Befragter das Individualmerkmal „Moslem“ aufweist und deshalb unabhängig von seinem nationalen-religiösen Kontext die EU-Prinzipien in einem geringeren Umfang akzeptiert) oder (3) auf eine Kombination beider Effekte zurückgeht.

Auch wenn man von all diesen fundamentalen Schwächen absieht, ist die von Wuermeling gewählte Spezifikation äußerst problematisch. Da, wie bereits erwähnt, die islamische Prägung über den Anteil der Muslime an der Gesamtbevölkerung gemessen wird, handelt es sich um eine metrische Größe auf der Ebene des nationalen Kontextes, die einen linearen Effekt auf individuelle Einstellungen haben soll.

Die Implikationen dieser Modellierung sind bei näherer Betrachtung nachgerade absurd: Beispielsweise müßte man davon ausgehen, daß in Westdeutschland, einer gefestigten westlichen Demokratie, die 1951/57 zu den Gründungsmitgliedern der Europäischen Gemeinschaften zählte, die Akzeptanz der EU-Prinzipien *ceteris paribus* niedriger sein müßte als im postsozialistischen Ostdeutschland, das erst durch die Wiedervereinigung Teil der EU wurde, da im Westen der Anteil der Muslime mit 4,7 Prozent um fast fünf Punkte höher liegt als in den neuen Ländern. In Frankreich (6,8 Prozent Muslime), einem weiteren Gründungsmitglied der EG, das 1789 Demokratie und Rechtsstaatlichkeit zu Staatsprinzipien erhob, 1791 als erstes Land der Welt die Strafverfolgung von Homosexuellen abschaffte und bis heute auf einer strikten Trennung von Staat und Religion besteht, sollte die Zustimmung zu den EU-Prinzipien ebenfalls niedriger sein als etwa in Estland (0,3 Prozent Muslime) oder der Slowakei (0 Prozent Muslime), die vor und nach dem Beitritt zur EU von der Kommission immer wieder wegen ihrer Minderheitenpolitiken gerügt wurden, oder in Polen (0 Prozent Muslime), das wegen der Diskriminierung von Homosexuellen vom Europäischen Gerichtshof für Menschenrechte verurteilt und vom Menschenrechtskommissar des Europarates scharf kritisiert wurde. Umgekehrt müßte man davon ausgehen, daß während der von Korruption, Autokratie und ethnischen Konflikten geprägten Spätphase des osmanischen Reiches die Zustimmung zu den EU-Prinzipien *ceteris paribus* höher gewesen wäre als in der modernen Türkei, weil der Anteil der Muslime an der Bevölkerung vor dem Ersten Weltkrieg lediglich bei etwa 70 Prozent lag.

Die Vorstellung, daß die bloße, mehr oder minder stark ausgeprägte Gegenwart von Muslimen in einem Land einen „kulturellen Kontext“ darstellt, der individuelle Einstellungen erklären könnte, ist offensichtlich nicht haltbar. Wenn überhaupt, dann sollte die religiös-konfessionelle Heterogenität eines Landes berücksichtigt oder aber eine dichotome Variable für eine *Dominanz* des Islam gebildet werden. In diesem Fall ergibt sich aber ein neues Problem: Unter allen in der europäischen Wertestudie

untersuchten Ländern ist nur die Türkei islamisch geprägt. Bulgarien als das Land, das mit rund 12 Prozent den zweitgrößten muslimischen Bevölkerungsanteil aufweist, wird von allen Autoren, die mit diesen Kategorien operieren, klar dem orthodoxen Kulturkreis zugeordnet (vgl. z. B. [Inglehart und Norris 2004](#): Kapitel 6).

Eine dichotome Variable für die Prägung durch den Islam wäre deshalb perfekt mit einem Ländereffekt für die Türkei korreliert. Empirisch trifft dies jedoch auch auf die von Wuermeling verwendete metrische Variable zu: Die Korrelation zwischen dem Muslimanteil und einem Indikator für das Befragungsland Türkei beträgt 0,99. Damit ist es unmöglich, zwischen einem (wie auch immer zu erklärenden) Effekt des Anteils der Muslime an der Bevölkerung und dem Einfluß anderer Merkmale der Systemebene, die für die Türkei spezifisch sind, zu unterscheiden. Die wichtigste Aussage des Artikels – „Je höher der Anteil an Muslimen im Land, desto geringer das Akzeptanzniveau der EU-Grundprinzipien“ – ([Wuermeling 2007](#): 211) ist deshalb von den Daten nicht gedeckt.

## 2. Replizierbarkeit

Im Sinne einer intersubjektiven Überprüfbarkeit sollten empirische Arbeiten dem „replication standard“ ([King 1995](#)) genügen, d. h. es sollte ohne Mühe möglich sein, die berichteten Ergebnisse exakt zu reproduzieren. Frederike Wuermelings Beitrag genügt diesen Minimalanforderungen leider nicht. In keinem Fall werden die Namen der verwendeten Variablen genannt, obwohl diese in den Fragebögen der European Values Study neben den Fragetext gedruckt sind, so daß eine Zuordnung besonders leicht möglich wäre.<sup>2</sup> Besonders ärgerlich ist dies im Falle der „vier Variablen“ die „sich für die Erstellung eines Demokratie-Indizes (sic) an[bieten]“ ([Wuermeling 2007](#): 197). Tatsächlich folgen im EVS-Fragebogen *acht* Variablen (v216-v223) aufeinander, die sich auf demokratische und nicht-demokratische Systeme beziehen. Vor diesem Hintergrund gerät die Identifikation der von Wuermeling verwendeten Variablen zum Ratespiel.

Auch die Bildung der Indizes für Religionsfreiheit und Befürwortung demokratischer Prinzipien ist unzureichend dokumentiert. Wie zwei Items mit je fünf Antwortvorgaben beziehungsweise vier Items mit je vier Antwortvorgaben von der Autorin zu einer fünf- beziehungsweise „vierstufigen Likertskala“ ([Wuermeling 2007](#): 197) kombiniert werden, ist alles andere als offensichtlich. Auch darüber, wie sie Kirchengangshäufigkeit (acht Stufen) und Gebetshäufigkeit (sieben Stufen) zu einer „7-stufigen Skala der Intensität der individuellen Religiosität“ ([Wuermeling 2007](#): 198) kombiniert, gibt die Autorin keine Auskunft. Letzteres ist besonders problematisch, da das zweite Item in Schweden und Slowenien überhaupt nicht erfragt wurde, so daß die

---

<sup>2</sup>Hier und im folgenden beziehe mich auf die vom Zentralarchiv unter der Nummer 3811 bereitgestellte Version der Daten und der Dokumentation.

entsprechenden Werte in diesen Ländern für *alle* Befragten fehlen. Darüber hinaus können die in den Tabellen genannten Fallzahlen nicht korrekt sein, da trotz listenweisem Ausschluß fehlender Werte stets die Gesamtzahl der Befragten im Datensatz (35.462) ausgewiesen ist.

### 3. Datenbasis und Operationalisierungen

Die Autorin weist selbst darauf hin, daß die von ihr analysierten Daten der EVS vergleichsweise alt sind, rechtfertigt ihre Verwendung aber damit, daß „Grundeinstellungen ... keinem schnellen zeitlichen Wandel unterliegen, da sie eng mit Wertvorstellungen verbunden sind“ (Wuermeling 2007: 196). Diese Annahme steht allerdings in einem gewissen Widerspruch zu den Erfahrungen mit dem relativ raschen Wandel der politischen Kultur etwa in Westdeutschland, dem Versuch der türkischen Eliten, mit Blick auf eine EU-Mitgliedschaft gesellschaftliche Wandlungsprozesse zu forcieren, dem raschen Nachrücken jüngerer und nach Einschätzung von Wuermeling selbst liberalerer Generationen in der Türkei, und nicht zuletzt der von ihr aufgestellten Hypothese, daß der über das Pro-Kopf-Inlandsprodukt operationalisierte *gegenwärtige* gesellschaftliche Wohlstand mit einer größeren Akzeptanz der EU-Prinzipien einhergeht.

Problematischer als das Alter der Daten sind allerdings die von Wuermeling verwendeten Operationalisierungen und hier insbesondere der von ihr konstruierte Index der Religionsfreiheit, der für eine valide Messung nicht geeignet ist, wie im folgenden gezeigt wird.<sup>3</sup> Als normativen Maßstab für die EU-Prinzipien verwendet Wuermeling die vom europäischen Konvent erarbeitete Grundrechtecharta, die sich wiederum auf die Menschenrechtskonvention des Europarates und die Allgemeine Erklärung der Menschenrechte der Vereinten Nationen stützt.

Religionsfreiheit im Sinne dieser Dokumente läßt sich, anders als von Wuermeling (2007: 196) behauptet, nicht auf „Säkularisierung ... [als] Prozess der institutionellen Trennung von Kirche/Religion und Staat (bzw. öffentlichen Ämtern)“ reduzieren. Nach dem Text der Charta hat Religionsfreiheit vielmehr mindestens sieben Aspekte: (1) das Recht, die Religion zu wechseln oder aufzugeben, sowie (2) das Recht auf private und (3) öffentliche Ausübung der eigenen Religion (Artikel II-70<sup>4</sup>), (4) das Recht der Eltern, ihre Kinder im Sinne der eigenen religiösen Überzeugungen zu erziehen (Artikel II-74), (5) das Verbot der Diskriminierung aufgrund religiöser Überzeugungen (Artikel II-81) und schließlich (6) der Schutz der religiösen Vielfalt durch die

---

<sup>3</sup> Auch die Indikatoren für die Akzeptanz der beiden Prinzipien „Gleichberechtigung“ und „Rechtsstaatlichkeit“ sind nicht ideal, da sie auf sehr spezifische Teilaspekte (geschlechtsspezifische Berufsrollen beziehungsweise Schmiergelder) abzielen (siehe Wuermeling 2007: 197).

<sup>4</sup> Hier und im folgenden wird die Numerierung des Verfassungsvertrages verwendet, wie er von Bundestag und Bundesrat beschlossen wurde (Bundestagsdrucksache 15/4900).

Europäische Union (Artikel II-82). Säkularisierung im Sinne im Sinne von Wuermeling ist nicht mehr als eine Voraussetzung für die negative Religionsfreiheit (1) und die weltanschauliche Neutralität der Union (5).

Ohne Zweifel steht es um die faktische Anerkennung der Religionsfreiheit durch die öffentliche Gewalt in der Türkei nicht zum besten. Zwar garantiert die türkische Verfassung in Artikel 14 die Religionsfreiheit, und drei wichtige religiöse Minderheiten (Juden, orthodoxe und armenische Christen) sind offiziell als Religionsgemeinschaften anerkannt. Andere Religionen werden jedoch diskriminiert (Freedom House 2007), und auch die Situation der orthodoxen Christen ist prekär.

Die beiden von Wuermeling verwendeten Items v129 und v131 sind jedoch nur sehr bedingt geeignet, um die Akzeptanz der Religionsfreiheit durch die Bevölkerung zu messen. V129 erfaßt die Zustimmung zu der Aussage „Politiker, die nicht an Gott glauben, sind ungeeignet für ein öffentliches Amt“. Wenn überhaupt, dann besteht hier ein Zusammenhang mit den Aspekten (1) und (5), d. h. dem Recht, keiner Religion anzugehören und dem Verbot der Diskriminierung auf Grund religiöser und weltanschaulicher Überzeugungen. Allerdings richtet sich diese „Diskriminierung“ gegen die sehr spezifische Subgruppe der Bevölkerung, die sich öffentlichen Wahlen stellt, bei denen in allen Ländern der Welt selbstverständlich und legitimerweise auch Charaktereigenschaften der Kandidaten eine Rolle spielen. Das Adjektiv „ungeeignet“ (in der englischen Ausgangsversion des Fragebogens „unfit“) ist primär eine moralische beziehungsweise inhaltlich-politische, nicht aber eine juristisch-staatsrechtliche Bewertung.<sup>5</sup> Daß eine starke Präferenz für fromme Politiker mit der Bereitschaft einhergeht, in anderen Bereichen des öffentlichen Lebens Atheisten, Agnostiker und Angehörige religiöser Minderheiten zu diskriminieren, ist durchaus wahrscheinlich, ergibt sich aber nicht aus den hier analysierten Daten.

In der Türkei gewinnt das Item durch den dauerhaften Konflikt zwischen religiös orientierten Politikern und der Militärführung überdies eine besondere und nicht ohne weiteres auf andere Länder übertragbare Bedeutung. Die Generale haben in der Vergangenheit bekanntlich mehrfach in die Politik eingegriffen, weil sie die laizistisch-nationalistischen Staatsprinzipien in Gefahr sahen. Die türkischen EVS-Daten wurden 2001 erhoben, d. h. in dem Jahr, in dem die islamische „Tugendpartei“ verboten wurde, deren Nachfolgerin AKP dann die Parlamentswahl von 2002 gewann. Vor diesem Hintergrund ist es schwer vorstellbar, daß das Item ausschließlich oder auch nur überwiegend die Akzeptanz für die Prinzipien der Grundrechtecharta erfaßt. Dies gilt aus offensichtlichen Gründen noch weitaus stärker für das Item v131: „Es wäre besser für die Türkei, wenn mehr Menschen mit einer starken religiösen Überzeugung öffentliche Ämter innehätten“.

Zusammenfassend läßt sich festhalten, daß die von Wuermeling verwendeten Items

---

<sup>5</sup>Dies gilt ausdrücklich auch für die türkische Version des Fragebogens.



eine Präferenz für oder gegen religiös orientierte Politiker erfassen, die in der Türkei eine besondere Bedeutung hat, aber keine Aussagen über Einstellungen gegenüber der Trennung von Kirche und Staat zulassen. Der aus den Items v129 und v131 gebildete Index ist deshalb ungeeignet, die Akzeptanz der Religionsfreiheit valide zu messen.

Zum Abschluß dieses Abschnitts sollen noch zwei vergleichsweise kleine Operationalisierungsprobleme angesprochen werden: Das Lebensalter der Befragten wird durch die Dekade des Geburtsjahres (Wuermeling 2007: 198) erfaßt, was im Ergebnis zur einer komplexen und vermutlich nicht intendierten nicht-metrischen Transformation führt. Beispielsweise wird zwei Befragten A und B, die Anfang 1970 beziehungsweise Ende 1979 geboren sind, jeweils ein Wert von 7 zugewiesen, obwohl sich ihr Lebensalter um fast zehn Jahre unterscheidet. Eine dritte Person C, die Anfang 1980 geboren wurde, erhält hingegen den Wert 8, selbst wenn er oder sie nur wenige Tage älter ist als B. Umgekehrt behandelt Wuermeling die im Datensatz vorhandenen numerischen Codes für die Größe des Wohnortes als eine metrische Variable, obwohl sich die Abstände zwischen den Kategorien dramatisch unterscheiden.

#### 4. Meßmodelle

Wuermelings Analyse enthält für zwei der vier abhängigen Variablen sowie für eine der unabhängig Variablen ein implizites Meßmodell: Indem die Autorin für die Akzeptanz von Religionsfreiheit und Demokratie sowie für die individuelle Religiosität jeweils einen Index konstruiert, geht sie stillschweigend davon aus, daß die von ihr in diesen Indizes zusammengefaßten Indikatoren (1) tatsächlich in allen Kontexten dasselbe messen und (2) die Zusammenhänge zwischen den Items über die 29 untersuchten Länder beziehungsweise Gebiete hinweg identisch sind. Die einfachste Möglichkeit, diese Annahmen zu überprüfen, ist die länderweise Berechnung von Cronbachs Alpha, das der (für die Zahl der Indikatoren korrigierten) mittleren Korrelation zwischen den Items einer Skala entspricht.

Als Minimum für eine reliable Messung wird dabei in der Literatur ein Wert von 0,7 genannt. Dieser Schwellenwert ist naturgemäß willkürlich, aber recht großzügig angesetzt. Im Falle einer Skala mit zwei beziehungsweise vier Indikatoren genügen bereits mittlere Korrelation von 0,54 beziehungsweise 0,38, um den Schwellenwert zu erreichen, was einer gemeinsamen Varianz von lediglich 29 beziehungsweise 14 Prozent entspricht. Für den Index der Religionsfreiheit ist jedoch in 12 der 29 Kontexte nicht einmal diese minimale Anforderung erfüllt. In ähnlicher Weise gilt dies auch für den als unabhängige Variable verwendeten Index der individuellen Religiosität. Den niedrigsten Wert für Alpha erzielt hier mit 0,29 ausgerechnet die Türkei, gefolgt von Rumänien und Griechenland (0,52 beziehungsweise 0,59). Bei näherer Betrachtung ist dies nicht allzu erstaunlich, da beide Items als Indikatoren für die intrinsische beziehungsweise extrinsische Sub-Dimension (vgl. Allport und Ross 1967: 434-435)

der Religiosität betrachtet werden können, die in Abhängigkeit vom kulturellen Kontext mehr oder minder unabhängig voneinander sind.

In allen drei Ländern kommt der relativ schwache Zusammenhang zwischen beiden Items dadurch zustande, daß eine größere Zahl von Befragten selten eine Kirche beziehungsweise Moschee besucht, aber dennoch relativ häufig betet. Besonders ausgeprägt ist dieses Muster in der Türkei, wo selbst unter den Befragten, die angeben, niemals an Gottesdiensten teilzunehmen, 63 Prozent täglich mindestens einmal beten. Eine plausible Erklärung für diesen Befund liegt darin, daß im Islam dem fünfmal täglich zu verrichtenden Gebet in Richtung Mekka („Namaz“) ein weitaus höherer Stellenwert zukommt als dem Besuch einer Moschee. Weiter kompliziert wird der Sachverhalt dadurch, daß im türkischen Fragebogen, der sich ja nicht nur an Muslime richtet, nicht auf dieses spezifische Gebet Bezug genommen, sondern ein allgemeineres Wort („Dua“) benutzt wird.<sup>6</sup> Damit sind Mißverständnisse und andere befragten-beziehungswise interviewerspezifische Effekte vorprogrammiert.

Die Reliabilität des aus den Demokratie-Items gebildeten Indexes ist noch weitaus geringer und erreicht in *keinem* der untersuchten Länder den Schwellenwert von 0,7. Mit Abstand am niedrigsten ist der Wert von Alpha wiederum in der Türkei, wo lediglich ein Wert von 0,39 erreicht wird. Dies erklärt sich zu einem gewissen Grad dadurch, daß das „starker Führer“-Item in der Türkei praktisch nicht mit den beiden Fragen korreliert ist, die sich direkt auf die Demokratie beziehen, während in den übrigen Ländern mit  $r = -0,28$  beziehungsweise  $r = -0,23$  zumindest ein moderater negativer Zusammenhang besteht. Auch der Zusammenhang zwischen dem „starker Führer“-Item und der Frage zur Bewertung einer Militärdiktatur ist in der Türkei mit  $r = 0,17$  deutlich schwächer ausgeprägt als im Mittel der übrigen Länder, wo immerhin ein Wert von 0,34 erreicht wird.

Zu einem ähnlichen Ergebnis gelangten auf der Basis des World Values Survey 1997 bereits [Tessler und Altinoglou \(2004: 35\)](#): Damals bestand überdies eine substantielle positive Beziehung zwischen der Befürwortung der Akzeptanz der Demokratie und dem Vertrauen in die Armee. Eine mögliche Erklärung für diese eher ungewöhnlichen Befunde liegt in der ambivalenten Rolle, die die Militärführung in der Geschichte der Türkei gespielt hat. Einerseits intervenierten die Offiziere zwischen 1960 und 1998 viermal im größeren Maßstab, um unerwünschte politische Entwicklungen zu verhindern ([Tessler und Altinoglou 2004: 22-23](#)), was mit demokratischen Prinzipien unvereinbar ist. Andererseits waren diese Eingriffe nicht auf die Einrichtung einer dauerhaften Militärdiktatur gerichtet und wurden vom Militär stets mit seiner Sonderrolle als Hüter der türkischen Demokratie und der republikanischen Prinzipien begründet, die von der Verfassung garantiert und zumindest von Teilen der Bevölkerung akzeptiert ist ([Tessler und Altinoglou 2004: 35](#)).

---

<sup>6</sup>Ich danke Ersin Oezsahin für die Erläuterung beider Begriffe.

Unabhängig von Spekulationen über die möglichen Ursachen läßt sich festhalten, daß die von Wuermeling konstruierte Skala in keinem der untersuchten Länder reliabel ist. Auch die Ergebnisse bezüglich der Akzeptanz des EU-Prinzips Demokratie sind deshalb nicht zu verwerten.

Aus anderen Gründen ist auch die Messung der Unterstützung für die Gleichberechtigung der Frau problematisch. Während die Indizes für die Akzeptanz der Religionsfreiheit und der Demokratie trotz der skizzierten Validitäts- und Reliabilitätsproblematik zumindest eine große Zahl von Meßwerten aufweisen und deshalb als quasi-metrisch betrachtet werden könnten, wird die Akzeptanz der Gleichberechtigung mit einem einzigen Item gemessen, bei dem lediglich drei Antwortmöglichkeiten vorgegeben waren: (1) „stimme zu“, (2) „lehne ab“ und (3) „weder noch“. Selbst wenn man davon absieht, daß Option (3) vermutlich von etlichen Befragten als „weiß nicht“-Kategorie mißverstanden beziehungsweise benutzt wurde, ist die Annahme einer metrischen Skalierung hier offensichtlich nicht haltbar und die von Wuermeling gewählte lineare Spezifikation unangemessen.

Damit verbleibt als einzig unproblematische der vier abhängigen Variablen das Korruptions-Item. Bezeichnenderweise wird jedoch gerade diese Variable von Wuermeling *nicht* in multivariater Perspektive untersucht, weil die Türkei bei diesem EU-Prinzip von allen Ländern die höchsten Zustimmungsraten aufweist. Aus der Sicht der vergleichenden Einstellungsforschung ist ein solcher negativer (beziehungsweise politisch positiver) Fall, in dem sich die von Wuermeling aufgestellten Hypothesen *prima facie*<sup>7</sup> nicht bestätigen lassen, mindestens ebenso interessant wie die relativ niedrigen türkischen Mittelwerte für die anderen Indikatoren. Gute *wissenschaftliche* Gründe, gerade diesen Fall nicht näher zu betrachten, lassen sich nur schwer vorstellen.

## 5. Modellierung

Selbst ohne alle bislang skizzierten Probleme wären die von Wuermeling geschätzten Modelle fehlspezifiziert. Bei den in Wuermelings Tabelle 2 (Wuermeling 2007: 206) präsentierten Ergebnissen handelt es sich um Schätzungen auf der Basis von Hierarchisch-Linearen- oder Mehr-Ebenen-Modellen (Snijders und Bosker 2000; Hox 2002), bei denen der Achsenabschnitt nicht als fix betrachtet wird, sondern zufällig über die untersuchten Länder variiert (random-intercept-Modell). Solche Mehr-Ebenen-Modelle tragen auf elegante Weise der Tatsache Rechnung, daß in vielen sozialwissenschaftlichen Anwendungen (1) die Fälle nicht unabhängig voneinander, sondern in Kontexten gruppiert sind, was zu über-optimistischen Standardfehlern führt,

---

<sup>7</sup>Korrekte Messungen und eine angemessene Spezifikation des Modells vorausgesetzt wäre es zumindest denkbar, daß die Effekte der Kontextvariablen von stärkeren Effekten der Individualvariablen überlagert werden.

(2) Merkmale der Kontexte selbst einen Einfluß auf die abhängige Variable haben können und (3) die Parameter des zu schätzenden Regressionsmodells (Achsenabschnitt und Steigungskoeffizienten) über die Kontexte hinweg ähnlich, aber nicht notwendigerweise identisch sind (kontextuelle Heterogenität). Für die vorliegenden Daten sind die von Wuermeling spezifizierten Modelle jedoch ungeeignet.

Ein erster Punkt betrifft dabei die von der Autorin vorgenommene Zentrierung der Individualvariablen an ihren jeweiligen Ländermittelwerten. Eine solche gruppenspezifische Zentrierung ist unüblich, weil sie die Bedeutung der geschätzten Effekte massiv und in einer oft schwer nachzuvollziehenden Weise verändert. Zudem ist sie mit einem Informationsverlust verbunden, sofern nicht auch die Gruppenmittelwerte als erklärende Variable im Modell enthalten sind (Kreft *et al.* 1995; Hofmann und Gavin 1998; Hox 2002: 62). Eine gruppenspezifische Zentrierung der Individualvariablen wäre nur dann anzuraten, wenn es sehr gute theoretische Gründe dafür gäbe anzunehmen, daß die Wirkung der betreffenden Variable einem „frog pond“-Mechanismus folgt (Kreft *et al.* 1995: 18; Hox 2002: 62). Mit Bezug auf die Türkei müßte man beispielsweise argumentieren, daß die Wirkung des Merkmals „Berufstätigkeit“ auf die Akzeptanz der EU-Prinzipien vom allgemeinen Beschäftigungsniveau in dem betreffenden Land abhängt. Eine solche theoretische Begründung für die gewählte Spezifikation ist nur schwer vorstellbar und in dem Beitrag von Wuermeling nicht zu erkennen.

Ein zweiter, grundsätzlicherer Punkt betrifft die Voraussetzungen, die für die Schätzung von Modellen mit Zufallskoeffizienten gelten. Erstens muß es sich bei den analysierten Kontexten tatsächlich um eine *Zufallsstichprobe aus einer (großen) Population von Kontexten* handeln (Kreft *et al.* 1995: 2), die als prinzipiell austauschbar betrachtet werden können. Insbesondere müssen innerhalb der Kontexte grundsätzlich dieselben Zusammenhänge bestehen und alle realen Abweichungen als zufällig interpretierbar sein. Zweitens sollte unabhängig davon die Zahl der Kontexte möglichst groß sein. In der Literatur werden als Minimum Zahlen zwischen 30 und 50 Kontexten genannt (Snijders und Bosker 2000: 140). Wenn wie bei Wuermeling Varianzkomponenten geschätzt werden, sollte die Zahl der Kontexte nochmals deutlich größer sein (Hox 2002: 175).

Im Fall der EU-Staaten sind diese Voraussetzungen schon deshalb offensichtlich nicht erfüllt, weil es sich gerade nicht um eine Stichprobe, sondern um die (mit Ausnahme Zyperns) vollständige Population der Länder handelt, über die eine Aussage gemacht werden soll (vgl. Berk 2004: 42-56 für eine weiterführende Diskussion). In der einschlägigen Literatur wird für Konstellationen wie die hier betrachtete deshalb ausdrücklich empfohlen, einen möglichen Einfluß der nationalen Kontexte durch fixe Effekte zu modellieren (siehe z. B. Duch und Stevenson 2005: 400 (FN 19)) beziehungsweise separate Modelle zu schätzen.

Darüber hinaus besteht hier keine echte Notwendigkeit für die Spezifikation eines

Modells mit Zufallskoeffizienten. Typischerweise werden Modelle mit Zufallskoeffizienten genutzt, wenn die Zahl der Kontexte sehr groß, die Zahl der Personen pro Kontext aber relativ klein (normalerweise zwischen 2 und 100) ist, so daß es nicht möglich ist, für jeden Kontext ein eigenes Modell zu spezifizieren. Statt dessen wird zur Schätzung der kontextspezifischen Parameter in mehr oder minderem großem Umfang Informationen aus anderen Kontexten herangezogen („borrowing strength“), was zu präziseren, aber verzerrten Schätzungen führt.

Die Notwendigkeit, Informationen aus anderen Kontexten heranzuziehen, nimmt jedoch selbstverständlich mit der Zahl der Befragten in einem Kontext ab. Ist diese wie hier mit rund 1.000 Personen pro Land sehr groß und gleichzeitig die Zahl der Kontexte vergleichsweise gering, so sollte auf weniger voraussetzungsreiche konventionelle Verfahren zurückgegriffen werden (Snijders und Bosker 2000: 43-44).

Im vorliegenden Fall könnte als Äquivalent zu den von Wuermeling spezifizierten Modellen etwa ein „dummy variable model“ geschätzt werden, das für jedes Land einen eigenen Achsenabschnitt enthält, der an die Stelle des random intercept und der Kontextvariablen tritt, aber ansonsten auf der Annahme basiert, daß die Effekte der unabhängigen Variablen in den einzelnen Ländern trotz des unterschiedlichen Ausgangsniveaus identisch sind. Da die verbleibenden zufälligen Einflüsse auf der Länderebene nicht notwendigerweise konstant und voneinander unabhängig sind, sollten überdies Huber-White-Standardfehler berechnet werden, die gegenüber solchen Klumpeneffekten robust sind. Auf diese Weise würde allen Aspekten der hierarchischen Datenstruktur Rechnung getragen.

Im Falle des Gleichberechtigungs-Items müßte mit Blick auf das Skalenniveau darüber hinaus eine nicht-lineare, d. h. ordinale oder multinomiale Spezifikation gewählt werden. Auf der Grundlage der Mittelwerte beziehungsweise Antwortwahrscheinlichkeiten, die sich aus diesem relativ einfachen Modell ergeben, lassen sich einige aufschlußreiche Gedankenexperimente durchführen. Setzt man beispielsweise für die Türkei das im Mittel deutlich höhere skandinavische Bildungsniveau in das Modell ein, so werden die Kompositionseffekte, die sich aus den unterschiedlichen Bevölkerungsstrukturen ergeben, ausgeglichen und die Kontexteffekte isoliert. Im Ergebnis steigt der erwartete Anteil derjenigen, die dem Gleichberechtigungs-Item zustimmen, erheblich an, liegt aber immer noch 50 Prozentpunkte unter den skandinavischen Werten.<sup>8</sup>

Daß es für diese extremen Unterschiede zwischen Skandinavien auf der einen und der Türkei (sowie vielen kontinentaleuropäischen Ländern) auf der anderen Seite eine Erklärung auf der Systemebene geben muß, ist offensichtlich. Ebenso offensichtlich ist jedoch, daß es für die Identifikation dieser Faktoren einer intensiven Beschäftigung

---

<sup>8</sup>Stata-Skripte zur Replikation dieser Befunde werden hier zur Verfügung gestellt: [hdl:1902.1/11171](https://hdl.handle.net/1902.1/11171)  
[UNF:3:eHkqqsOl69GCE2ap0dRWMw==](https://nfdi.ub.uni-leipzig.de/urn:nbn:de:hbz:5:1-63882-p0001-7) .

mit den betreffenden Ländern bedarf, die sich nicht auf einige Regressionsmodelle auf der Grundlage der EVS-Daten und zweier Makro-Indikatoren beschränken kann.

Eine Vielzahl von Erweiterungen des Modells wären leicht zu realisieren, ohne daß auf Zufallskoeffizienten zurückgegriffen werden müßte. Dies würde aber eine gründliche theoretische Auseinandersetzung mit den Determinanten der Einstellungen zu den Geschlechterrollen voraussetzen, die von der Autorin nicht geleistet wird.

### *III. Fazit*

Frederike Wuermelings Studie zur Verankerung demokratischer Prinzipien in der Türkei wird die wissenschaftliche Debatte über einen möglichen Türkei-Beitritt und das „demokratische Defizit“ der EU neu beleben. Substantiell geht Wuermelings Artikel jedoch nicht über die altbekannten Mittelwert-Differenzen hinaus, da ihre multivariaten Analysen aufgrund der zahlreichen inhaltlichen und methodischen Defizite nicht zu verwerten sind. Die Kernaussagen ihres Beitrages sind deshalb nicht haltbar. Daß Wuermelings Beitrag dennoch politisch instrumentalisiert wurde, ist der Autorin nicht anzulasten, wohl aber, daß sie eine Vielzahl von methodischen und inhaltlichen Problemen übersehen oder ignoriert hat.

### *Literatur*

- Allport, Gordon W., und Ross, J. Michael, 1967: Personal Religious Orientation and Prejudice. *Journal of Personality and Social Psychology* 5: 432–443.
- Bell, Daniel, 1973: *The Coming of the Post-Industrial Society. A Venture in Social Forecasting*. New York: Basic Books.
- Berk, Richard A., 2004: *Regression Analysis. A Constructive Critique*. Thousand Oaks, London, New Delhi: Sage.
- Coleman, James S., 1994: *Foundations of Social Theory*. Cambridge, London: The Belknap Press of Harvard University Press.
- Duch, Raymond M., und Stevenson, Randy, 2005: Context and the Economic Vote. A Multilevel Analysis. *Political Analysis* 13: 387–409.
- Freedom House, 2007: Turkey. [http://www.freedomhouse.org/modules/mod\\_call\\_dsp\\_country-fiw.cfm?year=2007&country=7291](http://www.freedomhouse.org/modules/mod_call_dsp_country-fiw.cfm?year=2007&country=7291) (20.07.2007): Freedom House.
- Gerhards, Jürgen, 2004: Europäische Werte – Passt die Türkei kulturell zu Europa? *Aus Politik und Zeitgeschichte* 2004: 14–20.

- Gerhards, Jürgen, 2007: *Cultural Overstretch? The enlargement of the European Union and the cultural differences between old and new member states and Turkey*. London, New York: Routledge.
- Hofmann, David A., und Gavin, Mark B., 1998: Centering Decisions in Hierarchical Linear Models: Implications for Research in Organizations. *Journal of Management* 24: 623–641.
- Hox, Joop, 2002: *Multilevel Analysis. Techniques and Applications*. Mahwah: Lawrence Erlbaum.
- Huntington, Samuel, 1996: *The Clash of Civilizations and the Remaking of World Order*. New York: Simon and Schuster.
- Inglehart, Ronald, 1997: *Modernization and Postmodernization. Cultural, Economic, and Political Change in 43 Societies*. Princeton: Princeton University Press.
- Inglehart, Ronald, und Baker, Wayne E., 2000: Modernization, Cultural Change, and the Persistence of Traditional Values. *American sociological review* 65: 19–51.
- Inglehart, Ronald, und Norris, Pippa, 2004: *Sacred and Secular. Religion and Politics Worldwide*. Cambridge: Cambridge University Press.
- King, Gary, 1995: Replication, Replication. *PS: Political Science and Politics* 28: 443–499.
- Kreft, Ita G. G., Leeuw, Jan de, und Aiken, Leona S., 1995: The Effect of Different Forms of Centering in Hierarchical Linear Models. *Multivariate Behavioral Research* 30: 1–22.
- Snijders, Tom A. B., und Bosker, Roel J., 2000: *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*. London, Thousand Oaks: Sage.
- Tessler, Mark, und Altinoglou, Ebru, 2004: Political Culture in Turkey: Connections Among Attitudes Toward Democracy, the Military and Islam. *Democratization* 11: 21–50.
- Weber, Max, 2006, zuerst 1904/05: *Die protestantische Ethik und der Geist des Kapitalismus, Vollständige Ausgabe*. Herausgegeben und eingeleitet von Dirk Kaesler. München: Beck.
- Wuermeling, Frederike, 2007: Passt die Türkei zur EU und die EU zu Europa? *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59: 185–214.