

Antwort auf die Replik von Marc Bühlmann, Wolfgang Merkel, Lisa Müller und Bernhard Weßels zum Forumsbeitrag von Thomas Müller und Susanne Pickel

Thomas Müller · Susanne Pickel

Marc Bühlmann, Wolfgang Merkel, Lisa Müller und Bernhard Weßels greifen in ihrer Replik eines der zentralen Anliegen der aktuellen Demokratieforschung auf – den berechtigten Wunsch nach Verbesserung der Beurteilung etablierter Demokratien. In unserem Artikel „Wie lässt sich Demokratie am besten messen? Zur Konzeptqualität von Demokratie-Indizes“ in Heft 3 (2007) der PVS haben wir diesem Bedürfnis sowie einer Reihe anderer Forderungen im Zusammenhang mit der validen und reliablen Bestimmung der demokratischen Qualität politischer Systeme Rechnung getragen.

Mit unserem Beitrag zielten wir vor allem auf die Tatsache ab, dass zugunsten von Ergebnissen, die zwischen Demokratien und Autokratien diskriminieren, zuweilen die Technik der Indexkomposition die Sorgfalt vermissen lässt, die für die durchgängige Validität und Reliabilität der Messung nötig wäre. Nicht nur die Herleitung der Systemkriterien aus der Theorie muss erneut mit der Empirie vereinigt werden, auch die methodische Sorgfalt der Indexkonstruktion sollte immer wieder einer kritischen Prüfung unterzogen werden.

Bisweilen unterliegt die ausführliche Prüfung und Würdigung von Demokratiemaßen – auch wenn man „nur“ die gängigen berücksichtigt – deutlichen Einschränkungen durch die Vorgaben des Publikationsmediums. Der für die Beurteilung der Qualität eines bei der PVS eingereichten Manuskripts wichtige Begutachtungsprozess hat zeitliche Verschiebungen zur Folge. So haben wir uns entschieden, Anmerkungen zur *Eignung der untersuchten Demokratiemessung für die aktuelle Demokratieforschung* an anderer Stelle (Pickel/Müller 2006) zu veröffentlichen. Einen neuen, in Teilen bereits im Internet zugängigen Index zur Bestimmung der Qualität etablierter Demokratien (Neuer Index Demokratie [NID], Lauth 2006, 2007) haben wir in einem jüngeren Konferenzpapier evaluiert (Pickel/Müller 2007b). Aus diesen Untersuchungen ergibt sich eine grundsätzliche

Thomas Müller, B. A. (✉)
Feldstraße 80, 17489 Greifswald, Deutschland
E-Mail: info@thomas-mueller.eu

Dr. Susanne Pickel
Universität Greifswald, Institut für Politikwissenschaft, Baderstr. 6/7, 17487 Greifswald, Deutschland
E-Mail: spickel@uni-greifswald.de

Erkenntnis, die dem von Bühlmann, Merkel, Müller und Weßels festgestellten Problem der Bestimmung von Qualitätsunterschieden entspricht: „Es kann nur gemessen werden, was theoretisch konzipiert wurde.“ Als Grundlage werden in den gängigen Indizes eben nicht diejenigen Indikatoren verwendet, die besonders gut geeignet wären, Demokratien *unterschiedlicher Qualität* voneinander zu unterscheiden. Ihr primäres Ziel ist die Trennung von Demokratien und Autokratien. Hingegen zielen sowohl der „Index Defekter Demokratie“ als auch der Fragekatalog von „Democratic Audit“ auf die Qualitätsbestimmung von Demokratien.¹ Die Konstrukteure des Democratic Audit sprechen sogar davon, dass es keine perfekte Demokratie gebe und sich somit in jedem realen demokratischen System verbesserungswürdige Punkte fänden. Beide Evaluierungsversuche sind jüngeren Datums, nehmen sich explizit der Frage der Qualität der Demokratie an und sind vielleicht schon allein deshalb darauf angewiesen, ihr Evaluierungskonzept besonders sorgfältig zu gestalten. Möchte man in diesem Sinne die feinen Unterschiede bestimmen, wie sie z. B. zwischen länger etablierten Demokratien bestehen, muss auch das Maß fein kalibriert sein; möchte man gröbere Unterschiede erkennen, reicht ein weniger feiner Maßstab aus.

Die Ergebnisse, die mit groben Messinstrumenten erzielt werden, sind umso niederschmetternder, als sie nicht nur Defizite bei der Qualitätsmessung etablierter Demokratien offenbaren, sondern zusätzlich die Unfähigkeit, problematische, defekte oder hybride Systeme eindeutig zu bestimmen. Dies zeigt sich besonders deutlich, wenn man ihre Werte kombiniert. Die Ergebnisse der vier gängigen Indizes Freedom House, Polity IV, ID (Vanhanen) und BTI (Bertelsmann Transformation Index) variieren nicht nur untereinander stark,² ihre Länderberichte zeigen zusätzlich Unstimmigkeiten der internen Bewertung auf. Sie decken sich teilweise nicht mit den in quantitatives Datenmaterial übersetzten Werten. Im Bereich der etablierten Demokratien unterliegen sie den bekannten Einschränkungen der Qualitätsbestimmung – haben die Demokratien ein bestimmtes Qualitätsniveau überschritten, so sind die Indizes zu keinem „fine tuning“ mehr fähig. Dies verhindert z. B. auch die Erfassung von kleineren Qualitätsverlusten innerhalb dieser hohen Margen und spiegelt eine Stabilität der Demokratiequalität vor, wie sie nicht unbedingt vorhanden ist.

In unmittelbarem Zusammenhang mit dem Bedarf an feinen Messinstrumenten steht die Frage nach der *Reichweite der Demokratiemodelle* der gängigen Indizes. Diese geht in der Tat kaum über das allgemein bekannte Angebot Dahl'scher Kriterien hinaus, die ja bekanntlich als Minimalkriterien – noch dazu von Polyarchie – gedacht waren (Dahl 1971).

Ein Weg, die Dahl'schen Minimalkriterien nicht nur zu ergänzen, sondern kritisch zu hinterfragen, ist eine differenzierte Erforschung des Verhältnisses von Gleichheit, Freiheit und Kontrolle. Freiheit und Gleichheit gehen in gegenwärtigen Demokratien vielfältige Verschränkungen ein und bilden nur noch selten real existierende Endpunkte eines Kontinuums zwischen libertärer Demokratie und sozialistischem Demokratiever-

1 Im Fall des Index Defekter Demokratien werden Autokratien nicht berücksichtigt. Der Democratic Audit geht davon aus, dass jedes politische System diesem Fragenkatalog unterworfen werden kann. Primär sollte er jedoch der Verbesserung von zumindest teilweise demokratischen Systemen dienen.

2 Russland, Burkina Faso und Sierra Leone werden von allen vier Maßen unterschiedlich beurteilt. Vanhanen bietet keine Länderberichte an.

ständnis ab. Kontrolle – als Maß für die Responsivität der Regierenden gegenüber den Regierten einerseits und als Maß zur Erfassung von Rechtsstaatlichkeit andererseits – bringt für die Legitimitätsbestimmung politischer Systeme wesentliche Aspekte in die Qualitätsbestimmung neuer Demokratieindizes ein (Pickel 2008; Lauth 2006, 2007).

Insofern sind die Erkenntnisse des NID, der sich auf die Säulen Wettbewerb, Partizipation, Stabilität und Kontrolle stützt, sowie das von Bühlmann, Merkel, Müller und Weßels vorgestellte Demokratiebarometer mit seinen Komponenten Freiheit, Gleichheit und Kontrolle differenzierter angelegt und fordern zur Evaluation anhand unserer Kriterien sowie zu einer ergänzenden Bewertung anhand auf die theoretische Komposition abzielender Maßstäbe (Fuchs/Roller 2008) geradezu auf.

Das von uns formulierte Ziel der Konzeptualisierungsphase, einen Index zu konstruieren, der *sparsam* mit der Zahl der Indikatoren umgeht, ruft indes nicht zum Verzicht auf relevante Demokratieindizes auf – „ein nicht sehr sparsames Konzept [kann] eben gerade alle relevanten Merkmale umfassen“ (Müller/Pickel 2007a: 519). Indem sich die Überprüfung der Konzeptualisierung vorrangig mit der technischen Konstruktion der Indizes beschäftigt – dies betrifft Konzeptspezifikation (sind alle relevanten Demokratieindizes erfasst?) und Konzeptlogik (sind die Merkmale logisch organisiert?) –, sind unsere Erkenntnisse über die weitgehende methodische Sauberkeit der Indexkonstruktionen in der Konzeptualisierungsphase schlüssig. Diese Konstruktionsschritte **verhalten sich demokratiethoretisch** zunächst neutral. Man kann den Blick jedoch nicht davor verschließen, dass das Dahl'sche Demokratieindizeskonzept oftmals implizit als Demokratieindizesmodell angenommen wird (Müller/Pickel 2007a: 524). Bedenklich erscheint uns – hierin stimmen wir mit Bühlmann, Merkel, Müller und Weßels überein – die „versteckte“ Verwendung von Merkmalen liberaler Demokratievorstellungen, die teilweise als so selbstverständlich angenommen werden, dass sie scheinbar keiner Erwähnung mehr bedürfen. Nicht nur, dass die Validität und Reliabilität solcher Maße infrage gestellt werden muss, ein analytisches Entwicklungspotenzial ist nicht zu erwarten, wenn alle Demokratien seit 1971 mehr oder minder den gleichen (wenig reflektierten) Kriterien unterworfen werden. In dieser Konsequenz ist die Eignung des Dahl'schen Konzeptes als „**Basismodell**“ zu hinterfragen und die Ergänzung bzw. eine erneute kritische Durchsicht der Dahl'schen Kriterien einzufordern (Müller/Pickel 2007a: 524-525; Dahl 1971, 1989). Das theoretische Demokratieindizesmodell der jeweiligen Indexkonstrukteure ist dementsprechend zu dokumentieren und zu begründen, ein Aspekt, dessen *inhaltliche*, demokratiethoretische Komponente allenfalls Randthema unserer Untersuchung war. Hier sei auf die jüngste Arbeit von Dieter Fuchs und Edeltraud Roller (2008) verwiesen.

Allerdings bleibt zu beachten, dass bei aller Kritik an den etablierten Indizes ihre Zielrichtung nicht aus dem Blick geraten sollte: Zu allererst wurden sie konzipiert, um politische Systeme zwischen Autokratie und Demokratie verorten zu können. Dies gelingt ihnen mit Ausnahmen hinlänglich. Mit der Zunahme demokratischer Systeme erwachsen neue Aufgaben: Wie stabil sind Demokratien? Welche Qualitätsunterschiede bestehen zwischen Demokratien? Verändert sich ihre Qualität? Zwei neue Maße, die dieser Aufgabe gerecht werden (könnten), liegen nunmehr vor. Mit Spannung sind empirische Ergebnisse dazu zu erwarten, die mit Sicherheit Gegenstand unserer weiteren Forschung auf dem Gebiet der Evaluierung von Maßen zur Bestimmung der Qualität politischer Systeme im Allgemeinen und Demokratien im Besonderen sein werden.

Literatur

- Dahl, Robert A.*, 1971: Polyarchy. Participation and Opposition. New Haven/London: Yale University Press.
- Dahl, Robert A.*, 1989: Democracy and its Critics. New Haven/London: Yale University Press.
- Fuchs, Dieter/Roller, Edeltraud*, 2008: Die Konzeptualisierung der Qualität von Demokratie. Eine kritische Diskussion aktueller Ansätze, in: *Brodocz, André/Llanque, Marcus/Schaal, Gary* (Hrsg.), Bedrohungen der Demokratie. Wiesbaden: VS Verlag für Sozialwissenschaften [im Erscheinen].
- Lauth, Hans-Joachim*, 2006: Die Qualität der Demokratie im interregionalen Vergleich. Probleme und Entwicklungsperspektiven, in: *Pickel, Gert/Pickel, Susanne* (Hrsg.), Demokratisierung im internationalen Vergleich. Neue Erkenntnisse und Perspektiven. Wiesbaden, 89-110.
- Lauth, Hans-Joachim*, 2007: Die Qualität der Demokratie. Der NID als pragmatischer Vorschlag für die komparative Forschung, in: *Behnke, Joachim/Behnke, Nathalie/Schnapp, Kai-Uwe* (Hrsg.), Datenwelten. Datenerhebung und Datenbestände jenseits klassischer Umfrageformen. Baden-Baden: Nomos [im Erscheinen].
- Müller, Thomas/Pickel, Susanne*, 2007a: Wie lässt sich Demokratie am besten messen? Zur Konzeptqualität von Demokratie-Indizes, in: Politische Vierteljahresschrift 48 (3), 511-539.
- Müller, Thomas/Pickel, Susanne*, 2007b: Performanz der Demokratiemessung. Überlegungen zur validen und reliablen Bestimmung politischer Systeme als Demokratien. Unveröffentlichtes Vortragsmanuskript. Tagung des Arbeitskreises „Demokratieforschung“ der DVPW 11.–13.10.2007 in Lüneburg.
- Neuer Index der Demokratie (NID)*, in: http://www.fernuni-hagen.de/polwiss/institut/projekte/39936_lg5.html; 18.12.2007.
- Pickel, Gert*, 2008: Die subjektive Verankerung von Demokratie in Osteuropa – Stand, Gründe und Konsequenzen von Bevölkerungseinstellungen in den jungen Demokratien Osteuropas. Wiesbaden: VS Verlag für Sozialwissenschaften [im Erscheinen].
- Pickel, Susanne/Müller, Thomas*, 2006: Systemvermessung – Schwächen der Konzepte und Verzerrungen der empirisch-quantitativen Bestimmung von Demokratie und Autokratie, in: *Pickel, Gert/Pickel, Susanne* (Hrsg.), Demokratisierung im internationalen Vergleich. Neue Erkenntnisse und Perspektiven. Wiesbaden: VS Verlag für Sozialwissenschaften, 135-172.