

Measuring Polyarchy

Michael Coppedge and Wolfgang H. Reinicke

The authors have developed a scale based on Robert Dahl's concept of polyarchy. The scale measures the degree to which national political systems meet the minimum requirements for political democracy, where real-world "democracies" rather than abstract ideals are the standard. The Polyarchy Scale is constructed from indicators of freedom of expression, freedom of organization, media pluralism, and the holding of fair elections. The scale is (1) well grounded in democratic theory, (2) world-wide in scope, (3) demonstrably valid, (4) solves problems of weighting indicators, and (5) is easy to interpret and replicate. Some limitations in the scale's applicability are discussed and suggestions are made for improvements and future research.

The Concept of Polyarchy

The term "polyarchy" was originally coined in Dahl and Lindblom (1953), but was developed most fully in Dahl (1971). Polyarchy is defined as the set of institutional arrangements that permits public opposition and establishes the right to participate in politics. In these two respects—public contestation and inclusiveness—polyarchy is similar to the concept of democracy. However, polyarchy is not, and was not intended to be, exactly equivalent to democracy, so it is important to clarify at the outset the various ways in which polyarchy differs from some other conceptions of democracy.

First, the term "democracy" inevitably calls to mind a host of ideals that no actual political system has ever approached. But we also refer to certain actual systems as "democracies," which is often confusing. In order to "maintain the distinction between democracy as an ideal system and the institutional arrangements that have come to be regarded as a kind of imperfect approximation of an ideal," Dahl and Lindblom introduced the term "polyarchy" to denote the latter (Dahl 1971, 9n).

Michael Coppedge is assistant professor in the Latin American Studies Program at the Paul H. Nitze School of Advanced International Studies (SAIS), 1740 Massachusetts Ave., N.W., Washington, D.C. 20036. He is completing a book on party factions and presidential democracy in Venezuela.

Wolfgang H. Reinicke is a Ph.D. candidate at Yale University and a research fellow at the Brookings Institution, 1775 Massachusetts Ave., N.W., Washington, D.C. 20036. He is currently completing a dissertation on the politics of global integration in banking and finance.

Second, like democracy, polyarchy is a quality of a political system; but unlike democracy (at least as it is usually conceived), polyarchy is also a dimension. There are degrees of polyarchy, ranging from full polyarchy to the absence of polyarchy, or hegemony. Those who are looking for a measure that will permit them to make comparisons among democratic states may therefore find a measure of polyarchy odd, in that it says more about the many gradations of failure to achieve full polyarchy than it says about the most democratic extreme.

Third, because polyarchy is concerned with imperfect approximations rather than ideals, the standard for the most democratic regimes is rather low. The concept of polyarchy is limited to the most basic institutional requirements for democracy, specifically those that had been met in most Western European countries by the end of the First World War. A country can qualify as a full polyarchy even if it does not allow workplace or communitarian democracy, proportional representation, referenda, or party primaries.

Fourth, polyarchy does not take into account varying degrees of democracy at different levels of the polity; it is concerned with the national regime only (Dahl 1971: 10–14). And fifth, no particular level of socioeconomic equality is required for a country to be fully polyarchic. Political and social scientists working in some traditions of scholarship would include social and economic equality in the definition of democracy, but in *Polyarchy*, Dahl discussed socioeconomic equality only as a condition that favors the development of polyarchy. So it should be clear that polyarchy reflects political, not social or economic, democracy. Our scale of polyarchy, then, measures the degree to which national political systems meet the minimum requirements for political democracy, where real-world “democracies” rather than abstract ideals are the standard.

As formulated in *Polyarchy* (Dahl 1971: 3), these minimum requirements for political democracy were:

1. Freedom to form and join organizations
2. Freedom of expression
3. The right to vote
4. Eligibility for public office
5. The right of political leaders to compete for support
6. Alternative sources of information
7. Free and fair elections
8. Institutions for making government policies depend on votes and other expressions of preference.

These requirements serve as both a conceptual and an operational definition of polyarchy, which greatly simplifies the problem of constructing a valid measure.

Polyarchy has by no means replaced democracy as an analytical concept in political science, but it is one of “democracy’s” closest competitors. Polyarchy is a concept familiar to most American political scientists. A sizable body of theoretical speculation and empirical research concerning polyarchy has been built up, not only in *Polyarchy*, but in Dahl’s more recent works and in publications by other political scientists.¹ The Polyarchy Scale developed in this article will provide a useful tool for evaluating many of the previously-untested hypotheses about polyarchy.

A Comprehensive Data Base

An ideal measure of polyarchy would be based on all existing nations. Until recently, comprehensive studies were not possible due to the lack of adequate political data on the smaller, newer, and least-developed nations. For example, in an earlier attempt to construct a Guttman scale of polyarchy, Norling and Williams (Dahl 1971: App. A) were handicapped by inadequate data. As a consequence, they had to exclude from their analysis twenty-one countries that were independent by 1968. Another consequence was that data were missing for a large number of the countries that *were* included in the analysis, which biased the ranking in favor of the countries about which less was known.²

Today the data are much more detailed and complete. Thanks to systematic data collection efforts by the Department of State, Raymond Gastil of Freedom House, Charles Humana, and others, there have been multiple sources of information on the political systems of every country in the world since the early 1980s. Using the information they had gathered, we were able to base our scale on the universe of the 170 independent nations that existed in 1985.³ Only dependent territories were excluded. Usually these sources provide detail sufficient to rate each country on each of the relevant criteria. The information is surprisingly current: even though our research was done in the summer of 1986, there was enough recent information to make our analysis a snapshot of the situation as of mid-1985.

Defining the Variables

With a few substantively minor adjustments, the component variables in the Polyarchy Scale correspond directly to Dahl's eight institutional requirements. Thus, we coded one variable for the extent of the suffrage, one for freedom of expression, one for freedom of organization, and one for the existence of alternative sources of information. Three of the remaining four institutional requirements were easily combined into a single variable measuring free and fair elections. It proved impossible to create a separate variable for the remaining institutional requirement, "Eligibility for public office," because information on this requirement is still incomplete. However, the information available suggests that eligibility for office differs very little from eligibility to vote, so little is lost by omitting this guarantee. The suffrage variable was eventually dropped from the final scale because it contributed very little empirically to the measurement of polyarchy. This amendment is discussed in greater detail below.

Following are descriptions of the variables and their categories.

The variable measuring free and fair elections is FAIRELT, which contains three categories:

1. Elections without significant or routine fraud or coercion.
2. Elections with some fraud or coercion.
3. No meaningful elections: elections without choice of candidates or parties, or no elections at all.

The freedom of organization variable, FREORG, has four categories:

1. Some trade unions or interest groups may be harassed or banned but there are no restrictions on purely political organization.
2. Some political parties are banned and trade unions or interest groups are harassed or banned, but membership in some alternatives to official organizations is permitted.
3. The only relatively independent organizations that are allowed to exist are nonpolitical.
4. No independent organizations are allowed. All organizations are banned or controlled by the government or the party.

The variable measuring freedom of expression, FREXT, contains three categories:

1. Citizens express their views on all topics without fear of punishment.
2. Dissent is discouraged, whether by informal pressure or by systematic censorship, but control is incomplete. The extent of control may range from selective punishment of dissidents on a limited number of issues to a situation in which only determined critics manage to make themselves heard. There is some freedom of private discussion.
3. All dissent is forbidden and effectively suppressed. Citizens are wary of criticizing the government even privately.

The fourth variable, ALTINF, is a measure of availability of alternative sources of information. It contains four categories:

1. Alternative sources of information exist and are protected by law. If there is significant government ownership of the media, they are effectively controlled by truly independent or multi-party bodies.
2. Alternative sources of information are widely available but government versions are presented in preferential fashion. This may be the result of partiality in and greater availability of government-controlled media; selective closure, punishment, harassment, or censorship of dissident reporters, publishers, or broadcasters; or mild self-censorship resulting from any of these.
3. The government or ruling party dominates the diffusion of information to such a degree that alternative sources exist only for nonpolitical issues, for short periods of time, or for small segments of the population. The media are either mostly controlled directly by the government or party or restricted by routine prior censorship, near-certain punishment of dissident reporters, publishers, and broadcasters, or pervasive self-censorship. Foreign media may be available to a small segment of the population.
4. There is no public alternative to official information. All sources of information are official organs or completely subservient private sources. The media are considered instruments of indoctrination. Foreign publications are usually unavailable or censored, and foreign broadcasts may be jammed.

The indicator of the right to vote is SUFF, which has the following categories:

1. Universal adult suffrage.
2. Suffrage with partial restrictions.
3. Suffrage denied to large segments of the population.
4. No suffrage.

It should be noted that quite a few countries that do not hold elections nevertheless provide for universal adult suffrage. We coded the suffrage variable according to the legal provisions of each country, leaving their interpretation to a later stage of analysis.

TABLE 1
1985 Distribution of Countries on Two-Dimensional Scale of Polyarchy

Degree of Public Contestation		Extent of Suffrage			
		Full	Partial	Very Restricted	None
Scale Types	Full 0	41	0	1	0
	1	10	0	0	0
	2	11	0	0	1
	3	1	0	1	0
	4	14	1	0	0
	5	8	0	0	0
	6	6	0	1	2
	7	15	0	0	4
	8	13	0	0	6
	9	5	0	0	2
None 10	20	0	1	6	
Totals		144	1	4	21

Coding the Variables

For the coding, we relied on sources that differ widely in the quality, format, and completeness of the information they provide. Some cover all countries, others omit the smaller ones. A few sources present information in precoded variables, others in a descriptive format. Some sources are useful for measuring all of the requirements for polyarchy, while others are useful for only one or a few. Rather than attempting to combine this disparate information into some untested index or weighing scheme, we decided to utilize it to code our own new variables. Fresh coding allowed us to draw on different sources and differently formatted data for each variable. All ratings were assigned initially by one coder consulting multiple sources. The sources consulted for each variable are listed in appendix A.

Unidimensionality

Since Dahl intended polyarchy to be a two-dimensional concept, we constructed two separate measures. The measure of the dimension of inclusiveness was SUFF, the ‘‘right to vote’’ variable. Public contestation was measured by constructing an ordered typology in the form of a Guttman scale.⁴ This scale ranged from systems with full public contestation (Type 0) to those that allow no contestation (Type 10).

A frequency cross tabulation of the measurements on the two dimensions (Table 1) led us to question the usefulness of inclusiveness as a criterion for polyarchy. Eighty-five percent of all countries in 1985 provided for universal suffrage, whether they held meaningful elections, approval elections, or no elections at all. Furthermore, all but two of the countries which had less than full suffrage in 1985 did not have meaningful elections in the first place. These two countries—South Africa and Western Samoa—should not be ranked as high as countries with the same polyarchy scores, but it is much simpler to treat these two cases as anomalies than to create a separate conceptual dimension to take them into account.

Some readers might feel that it would be worthwhile to keep the inclusiveness dimension in order to distinguish between nondemocratic regimes with limited hypothetical suffrage and nondemocratic regimes with no hypothetical suffrage at all. We disagree, but the ratings on the suffrage variable are included in the Appendix for the benefit of these readers. Dropping the suffrage dimension leaves us with a unidimensional scale of polyarchy that is identical to the scale of public contestation. Its coefficients of reproducibility and scalability are .900 and .829, respectively.⁵

Two other measures of democracy do not deal adequately with the question of unidimensionality. Gastil (1986) compiles separate indexes for political rights and civil liberties even though they are very highly correlated. Bollen (1988) insists on dividing Dahl's institutional requirements for polyarchy into separate conceptual dimensions ("political rights" and "political liberties") even though in building his earlier Index of Political Democracy (Bollen 1980), he demonstrated that the components of political democracy were unidimensional.⁶ Common sense dictates that a one-dimensional phenomenon be measured with a one-dimensional indicator. Once the unidimensionality of the phenomenon has been established, insisting on two-dimensional indicators is like trying to measure length in acres.

A Solution to the Problem of Weights

One problem that frequently arises in scale and index construction is assigning weights to the variables. How important is having a free press compared to having freedom of organization? Do you penalize a country more for having fraudulent elections than you do for sporadic press censorship? Most scaling techniques offer no definitive solutions to these problems. One great advantage of Guttman scaling is that it makes it possible to rank countries without having to decide on the relative weights of the component variables, as long as the countries' ratings match the perfect scale types. In effect, the variables are weighted equally. However, the weights are irrelevant to the scale; they could just as easily be grossly unequal, and the ranking of the scale types would remain completely unchanged. For example, any country that fits type 0 perfectly will always rank higher than any country that fits type 1 perfectly, no matter how the component variables are weighted.

Owing to this special property of Guttman scaling, the Polyarchy Scale produces an unequivocal ranking of 137 of the world's 170 independent nations for the year 1985. For the remaining thirty-three countries that do not fit the scale types perfectly, ranking is more problematic. There is no *a priori* way to weight the variables in order to decide whether these countries rank above, below, or on par with the countries that have the same scale score and match the perfect scale types.

Nevertheless, we can make a convenient distinction between two types of deviating cases. One group consists of twenty-six countries whose ratings differ only slightly from the corresponding perfect scale type: no rating is more than one category removed from the expected level, and no more than two ratings are off. We call these cases "approximately equivalent variants." Doubling the weight of one of the variables would move these countries only one rank up or down, so their ranks can safely be considered approximately correct.

The other group consists of anomalies: countries with one or more ratings that are two or more levels removed from the norm, or have three or more odd ratings altogether. The ranks of these countries change more drastically as the variable weights are modified, so that the ranks of anomalies cannot safely be considered approximately correct. Take, for example, the small European principalities of Andorra, Liechtenstein, and Monaco. In media pluralism, freedom of expression, and freedom of organization, they are like Type 0 polyarchies. But the effective leaders of their governments are not elected; they are either born to the office or appointed by other unelected officials. We prefer not to guess where these odd cases should be ranked. Fortunately, there are only a few—Andorra, Liechtenstein, Monaco, Vatican City, and Syria. South Africa and Western Samoa are also considered anomalies because of their restricted suffrage. Appendix B shows the Polyarchy Scale scores and their interpretations, along with a listing of the perfect scale types, approximately equivalent variants, and the anomalies.

Implications for the Measurement of Polyarchy

The most obvious implication of this scale is that polyarchy can be measured, albeit roughly. It is impossible to rank each country one by one on a unidimensional scale that spans the whole range from polyarchy to nonpolyarchy, because as soon as we try to rank countries within types, we encounter a multidimensional reality that does not allow simple comparisons such as rankings.⁷ This is analogous to measuring a thread with a ruler. On a human scale, a thread is effectively one-dimensional and can be measured with a one-dimensional instrument. On a microscopic scale, with microscopic instruments, the three-dimensional structure of the thread is harder to ignore.

We have not developed a finely calibrated instrument suitable for measuring the complex differences among countries at the same level of polyarchy. What we have produced is a coarsely calibrated instrument: an ordered typology with ten to twelve types that apply to many countries at once. This Polyarchy Scale should be very useful for identifying countries that are similar in the degree of polyarchy they possess, and for making broad comparisons of dissimilar types.

The scale is also useful as a checklist: it tells the researcher who is making comparisons involving polyarchy, which characteristics of political systems to pay attention to, and which to ignore. Broadly speaking, someone who is interested in how polyarchic a country is compared to other countries in the world should examine four factors: elections, freedom of organization, freedom of expression, and alternatives to official sources of information.

The Polyarchy Scale also provides more specific items for a checklist. Success in constructing a scale depends as much on how each variable is divided into categories as it does on selecting the correct variables. Based on the category descriptions for each variable, then, the important things to know in measuring polyarchy are:

1. whether or not elections are held that offer voters a meaningful choice of parties or candidates;
2. whether or not election outcomes are affected by significant fraud or coercion;

3. whether all, some, or no political organizations are banned;
4. if all political organizations are banned, whether some or no nonpolitical organizations are allowed to function independently;
5. whether freedom of expression is complete, nonexistent, or somewhere in between;
6. whether the media are pluralistic or government-dominated;
7. if the media are pluralistic, whether official views receive preferential or balanced treatment; and
8. if the media are government-dominated, whether control is complete or incomplete

The process of scale construction also revealed that several distinctions we assume to be relevant to polyarchy are not, in fact, useful for measuring polyarchy. There were five distinctions that we included in original working definitions of our variables, but eliminated because they frustrated our preliminary attempts to construct a scale. Their incompatibility with the other criteria under the constraints of unidimensionality does not mean they are unimportant distinctions for other concepts or issues; it simply means that they vary independently of the more basic components of polyarchy.

The original version of FAIRELT (fair elections) included categories that distinguished between countries that held approval elections (without a choice of candidates) and countries that held no elections at all for the effective leaders of the government. It also distinguished between degrees of fraud or coercion that probably determined the winner of the election, and those that only changed the margin of victory. FREORG (freedom of organization) began with seven categories, but had to be collapsed into four to produce a scale. This adjustment eliminated two distinctions: that between one-party and no-party states, and that between states that allow freedom of purely political organization but harass or ban trade unions or other interest groups, and those that guarantee freedom of both political and nonpolitical organization. The original version of FREXT (freedom of expression) contained two middle categories instead of one, in order to distinguish between exceptional and routine suppression of dissent, where control was nevertheless incomplete.

After collapsing a few categories in the original variables, we found that it was possible to generate more than one scale that met the minimum criterion of a coefficient of reproducibility of at least .900. The final scale was chosen because it does the best job of discriminating between countries that are subjectively dissimilar while still meeting the reproducibility requirement (at exactly .900). The substantive implication of the collapsed categories is that the following distinctions are not useful for measuring polyarchy:

1. elections without choice vs. no elections;
2. one-party states vs. no-party states;
3. exceptional vs. routine suppression of dissent;
4. electoral fraud or coercion that changes the winner of an election vs. fraud or coercion that changes only the margin of victory;
5. in states with full freedom of political organization, full vs. partial freedom of nonpolitical organization;

For example, the first two distinctions listed above were designed to distinguish

between authoritarian and totalitarian regimes. Their irrelevance means that some authoritarian and totalitarian regimes can be roughly equally “unpolyarchic.”

Reliability

In this kind of research the two principal causes of error, and therefore the greatest threats to reliability, are the biases of the sources consulted and subjectivity in the coders' scoring.⁸ We minimized contamination by the sources' biases by using multiple sources of information whenever possible. For some of the ratings, it was necessary to rely on a single source, but the number of countries affected is small: over ninety-one percent of the ratings are based on more than one source. The average number of sources consulted for all the ratings was 2.3.

Our original procedure was for each coder to pool the information obtained from all of the sources before assigning a rating. Where there were superficial disagreements among the sources, the coder tried to read between the lines to see a reality that was compatible with all of the descriptions. Where these differences were irreconcilable, two coders reviewed the sources and jointly agreed on the correct rating, sometimes after consulting additional sources.⁹

In our original procedure, there was no way to test the assumption that the use of multiple sources would improve the scale's reliability. In order to check the reliability of our ratings, we asked a research assistant to repeat the coding procedure, this time keeping records on consensus or disagreement among sources.¹⁰ Sixty-two percent of her ratings were based on complete consensus among all sources consulted. Another 23 percent were based on at least two-thirds consensus, and for the remaining fifteen percent, it was impossible to assign a rating without pooling information from the sources or consulting additional sources. To the extent that consensus among sources is an indicator of reliability, then, approximately three-fifths of our ratings can be considered very reliable, 85 percent can be considered at least fairly reliable, and about fifteen percent are of uncertain reliability.

Reliability is also affected by coding errors. Our original check on coding errors was to discuss together all ratings on which there seemed to be irreconcilable disagreement among the sources. As it turned out, we discussed fewer than fifty ratings, either because of consensus among sources or, perhaps in part, our own willingness to resolve apparently superficial disagreements by reading between the lines.

We also checked on coding errors by comparing our ratings with the ratings assigned by the research assistant. As mentioned earlier, she was unable to assign a rating to fifteen percent of the cases because of disagreement among sources. But 86.5 percent of the remaining ratings from the attempted replication agreed perfectly with ours. If the comparison is limited to just the ratings for which there was full consensus among sources, the rate of agreement between the original and replicated ratings rises to 93 percent.

Obviously, for a measure that passes judgment on entire nations, perfect reliability is desirable. While perfection is not attainable, there are simple ways to improve the reliability of the Polyarchy Scale that could easily be tried in the future. The most

pressing need is for more explicit coding criteria. The authors of this article shared many assumptions regarding coding decisions that did not find their way into the written category descriptions. As a result, the replicated ratings were based on somewhat different assumptions. We believe that most of the differences between the original and the replicated ratings would disappear if the coding criteria were more detailed. Coders should be thoroughly trained, and at least two coders should be used. In addition, where disagreements among sources or between coders prevent the assignment of a rating, coders should discuss the discrepancies and consult experts or additional sources.

All of these improvements must be made before every single rating can be regarded as highly reliable. However, that degree of reliability is not necessary for research involving comparisons of large numbers of countries, since most statistical analyses incorporate techniques for dealing with measurement error. We therefore feel safe in claiming that the Polyarchy Scale is already reliable enough for research on the relationship between polyarchy and other characteristics of whole political systems.

Validity

Validity is always difficult to establish, which explains why most attempts to measure democracy have claimed only face validity, which is entirely subjective. The validity of our scale is, however, one of its strongest points, with the understanding that it is a measure of polyarchy rather than democracy. Creating a measure with high construct validity is a question of (1) choosing the correct indicators, and (2) combining them correctly.¹¹ The fact that polyarchy is an operationally defined construct enabled us to satisfy the first requirement: with minor exceptions, there is a one-to-one correspondence between the guarantees that Dahl used to define polyarchy and the component variables that we used to measure his concept.

We did not code a separate variable for guarantee number 8, "Institutions for making government policies depend on votes and other expressions of preference," because in a more recent reformulation of these criteria, Dahl stipulated that officials who have "control over government decisions about policy," rather than the policies themselves, should depend on votes (Dahl 1982, 10–11). This amendment made the eighth guarantee essentially equivalent to number 7, "free and fair elections." It was not necessary to construct a variable for number 5, "Right of political leaders to compete for support," because this requirement is implicit in three others—"free and fair elections," "freedom to form and join organizations," and "freedom of expression." Aside from these adjustments, the component variables are by definition valid indicators of polyarchy.

But there are two sides to choosing the correct indicators. So far we have shown that we included indicators of all of the aspects of polyarchy; we must also be careful to include indicators of just those aspects and no others. This second requirement was satisfied by the process of constructing the Guttman scale, for we began our scale construction with variables that were more finely calibrated than the ones that appear in the final scale.¹² That is, our original variables included "extra" infor-

mation that we presumed to be useful for measuring polyarchy as a unidimensional concept, but which scalogram analysis showed to be unrelated to polyarchy. By eliminating this extra information, we pared down our variables to the most valid possible set of indicators.

Additional support for the validity of the scale comes from its association with other constructs thought to be correlated with democracy. The Pearson product-moment correlations between our scale (with the seven anomalies excluded) and Gastil's indexes of political rights and civil liberties are .938 and .934, respectively.¹³

The Polyarchy Scale should be used with caution in multiple regression since it is ordered categorical data rather than ratio data. The fact that it lacks an absolute zero point and there is no guarantee that the categories are evenly spaced violates one of the assumptions of regression. Log-linear and logit models are more appropriate, as they are for most other measures of democracy. Given the finite universe of 170 countries, the other variables in such models should not have a total of more than three or four categories, unless the Polyarchy Scale is collapsed into a smaller number of categories for the purposes of analysis.

Straightforward Interpretation

The reproducibility of a Guttman scale means that if one knows the score of a country on the scale, one can reproduce, or reconstruct, the country's ratings on each of the component variables. (Actually, with the minimum coefficient of reproducibility of .900, one can reproduce ninety percent of the ratings.) This one-to-one correspondence between scores and combinations of characteristics, or "perfect scale types," makes Guttman scale scores very easy to interpret.

For example, Chile's scale score of 5 tells one much more than the fact that Chile ranks higher than Jordan, lower than Mexico, and about on par with Liberia; it also tells one that in Chile, no meaningful elections are held, some independent political organizations are banned, some public dissent is suppressed, and there is preferential presentation of official views in the media. And by comparing these conditions with the standard conditions in the adjacent scale types, one can determine what the substantive difference is between Chile and the countries that are ranked higher or lower. For example, the relevant difference between Chile and Mexico is that there are meaningful elections in Mexico, even though they are marred by fraud or coercion; and the difference between Chile and Jordan is that some independent political organizations exist in Chile, while in Jordan they are all banned.

Easy Replication

The final attractive characteristic of the Polyarchy Scale is that it is relatively easy to replicate. Fresh coding of the variables means that it is not necessary to wait for an updated edition of the *Political Handbook* in order to decide how polyarchic a country is. Anyone having access to the sources and to the category definitions could rate countries individually or in groups. With a more detailed coding guide, relatively little training would be required to do an adequate job of this, thanks to the

information collected by other scholars. Our principal sources update their information annually, so as long as they continue to do so, the Polyarchy Scale could also be updated annually. Since it relies on multiple sources, it could continue to be updated even if some of the sources are discontinued or replaced by others. Replication by other researchers, preferably with teams of coders, is the best check on reliability.

Conclusions

The Polyarchy Scale that we have constructed is theoretically well-grounded, empirically comprehensive, demonstrably valid, and easy to interpret and replicate. While its reliability needs to be improved, it is already useful for large-scale comparative research on polyarchy. The scale is somewhat disappointing in that it locates such a disproportionate number of countries in the top category. This result is inevitable, given the fact that polyarchy is concerned with the *minimum* requirements for political democracy.

Still, one would like to be able to compare the quality of democracy among the full polyarchies. A "Polyarchy Plus" scale would be constructed by adding a new variable or variables, or subdividing the top category of one of the existing variables, in order to incorporate an additional criterion that is useful for making distinctions among the full polyarchies. Unfortunately, there are many such criteria that could be incorporated into the scale. Incorporating all of these would almost certainly violate unidimensionality; incorporating just one or some of them would require making difficult and controversial judgments. Until a scholarly consensus develops on the essential characteristics of higher levels of democracy, the Polyarchy Scale is probably as precise as we can make it.

To conclude on a cautionary note, we believe it is unwise to discuss measurement of democracy or polyarchy without paying attention to regime stability. While stability and polyarchy are very distinct aspects of a political system, instability has an undeniable impact on the quality of political life. A precarious or unconsolidated polyarchy is very different from a fully consolidated one. We therefore urge those who may contemplate using our scale to remember that there are important characteristics of democracy that it does not reflect.

Appendix A

Sources Consulted for Coding Variables

Fair Elections (FAIRELT): Gastil 1986, Statesman's Yearbook 1986, Banks and Textor 1985, Europa 1986a-d; and in some cases, Phillips 1984, Delury 1983, DOS 1986, Watch Committees 1986, Humana 1986, and McHale 1983.

Freedom of Expression (FREXT): Humana 1986, Gastil 1986, DOS 1986, and Watch Committees 1986.

Media Pluralism (ALTINF): Humana 1986, Gastil 1986, DOS 1986, and Watch Committees 1986.

Freedom of Organization (FREORG): Humana 1986, Gastil 1986, DOS 1986; and in some cases Europa 1986a-d, Phillips 1984, Statesman's Yearbook 1986, Banks and Textor 1985, McHale 1983, Delury 1983, and Alexander 1982.

Extension of Suffrage (SUFF): Banks and Textor 1985, Delury 1983, Alexander 1982, McHale 1983, Fukui 1985, and Europa 1986a-d.

Appendix B

The Polyarchy Scale 1985

The series of numbers in parentheses are the combinations of ratings received by the countries in that column. The first digit is the rating on FAIRELT, the second is for FREORG, and the third is for FREXT, and the fourth is for ALTINF.

Scale

Score N

0 41 Perfect Scale Type (1 1 1 1):

Meaningful fair elections are held, there is full freedom for political organization and expression, and there is no preferential presentation of official views in the media.

Argentina	Finland	Luxembourg	San Marino
Australia	France	Nauru	Spain
Austria	Germany, West	Netherlands	Sweden
Barbados	Grenada	New Zealand	Switzerland
Belgium	Honduras	Norway	Trinidad & Tobago
Belize	Iceland	Papua New Guinea	Tuvalu
Brazil	Ireland	Portugal	United Kingdom
Canada	Italy	St. Christopher & Nevis	United States
Colombia	Japan	St. Lucia	Uruguay
Costa Rica	Kiribati	St. Vincent & The Grens.	Venezuela
Denmark			

Scale

Score N

- 1 10 Perfect Scale Type (1 1 1 2):
 Meaningful fair elections are held, and there is full freedom for political organization and expression, but there is preferential presentation of official views in the media.
- | | |
|----------|-----------------|
| Botswana | Ecuador |
| Cyprus | Mauritius |
| Dominica | Solomon Islands |
- Dominican Republic
 Approximately Equivalent:
 Israel, Peru (1 1 2 1)
 Fiji (2 1 1 1)
- 2 9 Perfect Scale Type (1 1 2 2):
 Meaningful fair elections are held and there is full freedom for political organization, but some public dissent is suppressed and there is preferential presentation of official views in the media.
- Antigua & Barbuda
 Bahamas
 Bolivia
 India
 Thailand
 Approximately Equivalent:
 Vanuatu (1 1 2 3)
 Greece, Jamaica (2 1 1 2)
 Panama (2 2 1 1)
- 3 0 (1222/2122): Undefined due to lack of cases.
- 4 16 Perfect Scale Type (2 2 2 2):
 Elections are marred by fraud or coercion, some independent political organizations are banned, some public dissent is suppressed, and there is preferential presentation of official views in the media.
- | | |
|-------------|-------------|
| Egypt | Philippines |
| El Salvador | Singapore |
| South Korea | Sri Lanka |
| Malta | Turkey |
- Mexico
 Approximately Equivalent:
 Senegal (2 2 1 2)
 Malaysia, Nicaragua, Zimbabwe (1 2 2 3)
 Maldives, Morocco (3 1 2 2), Lebanon (3 2 2 1)

Scale

Score N

- 5 7 Perfect Scale Type (3 2 2 2):
No meaningful elections are held, some independent political organizations are banned, some public dissent is suppressed, and there is preferential presentation of official views in the media.
- Bangladesh
Chile
Approximately Equivalent:
Gambia, Indonesia, Guyana (2 2 2 3)
- Guatemala
Liberia
- 6 9 Perfect Scale Type (3 3 2 2):
No meaningful elections are held, only nonpolitical organizations are allowed to be independent, some public dissent is suppressed, and there is preferential presentation of official views in the media.
- Jordan
Kuwait
Approximately Equivalent:
Lesotho, Madagascar, Pakistan, Paraguay (3 2 2 3)
Bhutan (3 4 1 2)
- Sudan
Tunisia
- 7 19 Perfect Scale Type (3 3 2 3):
No meaningful elections are held, only non-political organizations are allowed to be independent, some public dissent is suppressed, and alternatives to the official media are very limited.
- Brunei
Burkina Faso
C. A. R.
Chad
Djibouti
Gabon
Approximately Equivalent:
Iran (3 2 3 3)
- Ghana
Guinea
Haiti
Poland
Rwanda
Sierra Leone
- Suriname
Swaziland
Taiwan
Tonga
Yugoslavia
Zambia
- 8 18 Perfect Scale Type (3 4 2 3):
No meaningful elections are held, all organizations are banned or controlled by the government or official party, some public dissent is suppressed, and alternatives to the official media are very limited.
- Bahrain
Benin
Burundi
Cameroon
Cape Verde
Comoros
Approximately Equivalent:
Qatar (3 3 2 4)
- Congo
Hungary
Ivory Coast
Kenya
Mauritania
Nepal
- Nigeria
Seychelles
Uganda
United Arab Emirates
North Yemen

Scale

Score N

9 7 Perfect Scale Type (3 4 2 4):

No meaningful elections are held, all organizations are banned or controlled by the government or official party, some public dissent is suppressed, and there is no public alternative to official information.

Algeria	Tanzania
Equatorial Guinea	Zaire
Oman	

Approximately Equivalent:

Malawi (3 3 3 4)

Niger (3 3 3 4)

Scale

Score N

10 27 Perfect Scale Type (3 4 3 4):

No meaningful elections are held, all organizations are banned or controlled by the government or official party, all public dissent is suppressed, and there is no public alternative to official information.

Afghanistan	Germany, East	Mozambique
Albania	Guinea-Bissau	Romania
Angola	Iraq	Sao Tome & Principe
Bulgaria	Kampuchea	Saudi Arabia
Burma	Korea, North	Somalia
China	Laos	Togo
Cuba	Libya	USSR
Czechoslovakia	Mali	Viet Nam
Ethiopia	Mongolia	South Yemen

Anomalies excluded from the ranking are:

Andorra, Liechtenstein, Monaco, Vatican City (3 1 1 1); Syria (3 2 3 4); and South Africa and Western Samoa (1 2 2 2 and 1 1 1 1 with limited suffrage).

Appendix C

1985 Country Ratings on Criteria for Polyarchy (Numbers correspond to category numbers in text.)

NATION	FAIRELT	FREORG	FREXT	ALTINF	SUFF
AFGHANISTAN	3	4	3	4	4
ALBANIA	3	4	3	4	1
ALGERIA	3	4	2	4	1
ANDORRA	3	1	1	1	1
ANGOLA	3	4	3	4	3
ANTIGUA & BARBUDA	1	1	2	2	1
ARGENTINA	1	1	1	1	1
AUSTRALIA	1	1	1	1	1
AUSTRIA	1	1	1	1	1
BAHAMAS	1	1	2	2	1
BAHRAIN	3	4	2	3	4
BANGLADESH	3	2	2	2	1
BARBADOS	1	1	1	1	1
BELGIUM	1	1	1	1	1
BELIZE	1	1	1	1	1
BENIN	3	4	2	3	1
BHUTAN	3	4	1	2	1
BOLIVIA	1	1	2	2	1
BOTSWANA	1	1	1	2	1
BRAZIL	1	1	1	1	1
BRUNEI	3	3	2	3	4
BULGARIA	3	4	3	4	1
BURKINA FASO	3	3	2	3	1
BURMA	3	4	3	4	1
BURUNDI	3	4	2	3	1
CAMEROON	3	4	2	3	1
CANADA	1	1	1	1	1
CAPE VERDE	3	4	2	3	1
CENTRAL AFRICAN REP.	3	3	2	3	1
CHAD	3	3	2	3	1
CHILE	3	2	2	2	1
CHINA	3	4	3	4	1
COLOMBIA	1	1	1	1	1
COMOROS	3	4	2	3	1
CONGO	3	4	2	3	1
COSTA RICA	1	1	1	1	1
CUBA	3	4	3	4	1
CYPRUS	1	1	1	2	1
CZECHOSLOVAKIA	3	4	3	4	1
DENMARK	1	1	1	1	1
DJIBOUTI	3	3	2	3	1

NATION	FAIRELT	FREORG	FREXT	ALTINF	SUFF
DOMINICA	1	1	1	2	1
DOMINICAN REPUBLIC	1	1	1	2	1
ECUADOR	1	1	1	2	1
EGYPT	2	2	2	2	1
EL SALVADOR	2	2	2	2	1
EQUATORIAL GUINEA	3	4	2	4	1
ETHIOPIA	3	4	3	4	4
FIJI	2	1	1	1	1
FINLAND	1	1	1	1	1
FRANCE	1	1	1	1	1
GABON	3	3	2	3	1
GAMBIA	2	2	2	3	1
GERMANY, EAST	3	4	3	4	1
GERMANY, WEST	1	1	1	1	1
GHANA	3	3	2	3	4
GREECE	2	1	1	2	1
GRENADA	1	1	1	1	1
GUATEMALA	3	2	2	2	1
GUINEA	3	3	2	3	4
GUINEA-BISSAU	3	4	3	4	4
GUYANA	2	2	2	3	1
HAITI	3	3	2	3	1
HONDURAS	1	1	1	1	1
HUNGARY	3	4	2	3	1
ICELAND	1	1	1	1	1
INDIA	1	1	2	2	1
INDONESIA	2	2	2	3	1
IRAN	3	2	3	3	1
IRAQ	3	4	3	4	1
IRELAND	1	1	1	1	1
ISRAEL	1	1	2	1	1
ITALY	1	1	1	1	1
IVORY COAST	3	4	2	3	1
JAMAICA	2	1	1	2	1
JAPAN	1	1	1	1	1
JORDAN	3	3	2	2	1
KAMPUCHEA	3	4	3	4	1
KENYA	3	4	2	3	1
KIRIBATI	1	1	1	1	1
KOREA, NORTH	3	4	3	4	1
KOREA, SOUTH	2	2	2	2	1
KUWAIT	3	3	2	2	3
LAOS	3	4	3	4	4
LEBANON	3	2	2	1	2
LESOTHO	3	2	2	3	4

NATION	FAIRELT	FREORG	FREXT	ALTINF	SUFF
LIBERIA	3	2	2	2	1
LIBYA	3	4	3	4	1
LIECHTENSTEIN	3	1	1	1	1
LUXEMBOURG	1	1	1	1	1
MADAGASCAR	3	2	2	3	1
MALAWI	3	3	3	4	1
MALAYSIA	1	2	2	3	1
MALDIVES	3	1	2	3	1
MALI	3	4	3	4	1
MALTA	2	2	2	2	1
MAURITANIA	3	4	2	3	4
MAURITIUS	1	1	1	2	1
MEXICO	2	2	2	2	1
MONACO	3	1	1	1	1
MONGOLIA	3	4	3	4	1
MOROCCO	3	1	2	2	1
MOZAMBIQUE	3	4	3	4	4
NAURU	1	1	1	1	1
NEPAL	3	4	2	3	1
NETHERLANDS	1	1	1	1	1
NEW ZEALAND	1	1	1	1	1
NICARAGUA	1	2	2	3	1
NIGER	3	3	3	4	4
NIGERIA	3	4	2	3	4
NORWAY	1	1	1	1	1
OMAN	3	4	2	4	4
PAKISTAN	3	2	2	3	1
PANAMA	2	2	1	1	1
PAPUA NEW GUINEA	1	1	1	1	1
PARAGUAY	3	2	2	3	1
PERU	1	1	2	1	1
PHILIPPINES	2	2	2	2	1
POLAND	3	3	2	3	1
PORTUGAL	1	1	1	1	1
QATAR	3	3	2	4	4
ROMANIA	3	4	3	4	1
RWANDA	3	3	2	3	1
ST. CHRISTOPHER & NEVIS	1	1	1	1	1
ST. LUCIA	1	1	1	1	1
ST. VINCENT & GREN.	1	1	1	1	1
SAN MARINO	1	1	1	1	1
SAO TOME & PRINCIPE	3	4	3	4	1
SAUDI ARABIA	3	4	3	4	4
SENEGAL	3	2	1	2	1

NATION	FAIRELT	FREORG	FREXT	ALTINF	SUFF
SEYCHELLES	3	4	2	3	1
SIERRA LEONE	3	3	2	3	1
SINGAPORE	2	2	2	2	1
SOLOMON ISLANDS	1	1	1	2	1
SOMALIA	3	4	3	4	1
SOUTH AFRICA	1	2	2	2	3
SPAIN	1	1	1	1	1
SRI LANKA	2	2	2	2	1
SUDAN	3	3	2	2	4
SURINAME	3	3	2	3	4
SWAZILAND	3	3	2	3	1
SWEDEN	1	1	1	1	1
SWITZERLAND	1	1	1	1	1
SYRIA	3	2	3	4	1
TAIWAN	3	3	2	3	1
TANZANIA	3	4	2	4	1
THAILAND	1	1	2	2	1
TOGO	3	4	3	4	1
TONGA	3	3	2	3	1
TRINIDAD & TOBAGO	1	1	1	1	1
TUNISIA	3	3	2	2	1
TURKEY	2	2	2	2	1
TUVALU	1	1	1	1	1
UGANDA	3	4	2	3	4
UNITED ARAB EMIRATES	3	4	2	3	1
UNITED KINGDOM	1	1	1	1	1
UNITED STATES	1	1	1	1	1
URUGUAY	1	1	1	1	1
USSR	3	4	3	4	1
VANUATU	1	1	2	3	1
VATICAN CITY	3	1	1	1	4
VENEZUELA	1	1	1	1	1
VIET-NAM	3	4	3	4	1
WESTERN SOMOA	1	1	1	1	3
YEMEN, NORTH	3	4	2	3	1
YEMEN, SOUTH	3	4	3	4	1
YUGOSLAVIA	3	3	2	3	1
ZAIRE	3	4	2	4	1
ZAMBIA	3	3	2	3	1
ZIMBABWE	1	2	2	3	1

Notes

We are deeply indebted to Robert Dahl for conceiving this project, entrusting it to us, and giving us encouragement and intellectual guidance along the way. We have also benefited from comments by Kenneth Bollen, Martin Lipset, and Larry Diamond, and from the research assistance of Anna Pappavlachopoulou. Responsibility for errors, however, is ours alone.

1. The *Social Sciences Citation Index* lists 182 references to *Polyarchy* since its publication.
2. Another problem with the scaling attempt by Norling and Williams was that the data did not match Dahl's

criteria for polyarchy very closely. Norling and Williams relied on updated versions of ten variables from Banks and Textor's 1963 *Cross-Polity Survey* that seemed to be related to the requirements for polyarchy. But they never tested the Banks and Textor variables to see whether they were conceptually valid substitutes for Dahl's eight criteria. Instead, they simply tried to scale the ten variables directly, with results that were optimistically judged "moderately satisfactory." Their coefficient of reproducibility was only .829, even though it was calculated by a controversial method of counting errors that tends to raise the coefficient artificially (McIver and Carmines 1981, 42-45).

3. We did not use Gastil's Indexes of Civil Rights and Political Liberties directly, since they are summary measures; instead, we used the descriptive information that he provides as justification for the ratings.
4. For an introduction to Guttman scaling, see Gorden, 1977.
5. Those who are accustomed to Guttman scales based on dichotomous variables may find our polychotomous variables unorthodox, but there is no cause for concern; the coefficients reported in the text *underestimate* the unidimensionality of the scale, if they are inaccurate at all. Each n-chotomous variable can be disaggregated into n-1 dichotomous variables, generating an equivalent scale of ten dichotomous variables. The coefficients of reproducibility and scalability of this new scale are ever more impressive, at .957 and .869, respectively. Conventionally, the coefficient of reproducibility should be at least .900 and the coefficient of scalability should be at least .600 (Dunn-Rankin 1983, 106-107).
6. Bollen's Index, like ours, did not incorporate a component of inclusiveness.
7. In technical terms, if we try to construct a more precise scale by increasing the number of categories in each variable so that each institutional guarantee is measured more accurately, the coefficient of reproducibility falls below .900.
8. Some commonly used indicators of reliability, such as coefficient alpha, are not appropriate for the Polyarchy Scale. These measures were developed for psychological testing, in which reliability depends much more on the internal consistency of the instruments than it does on measurement error. For a task such as measuring polyarchy, measurement error due to the subjectivity of the sources and the coding is much more important. The coefficient of reliability is best regarded in this situation as an upper limit on the reliability of the scale based on its internal consistency—i.e., the number of variables and the degree to which they all measure the same thing. (In this sense, reliability is also measured by the coefficient of scalability, reported in the discussion of validity.) However, reliability coefficients can be calculated. For the Polyarchy Scale, the value of KR-20, the Kuder-Richardson version of coefficient alpha for dichotomous items, is .946. To calculate it, we decomposed our four multiple-choice variables into ten dichotomous variables. A coefficient of .946 indicates the extremely high degree of reliability that one would like to have for a measure that would be used to make important decisions, such as an IQ score that would be used to assign children to special education classes (Nunnally 1967, 226).
9. Actual procedures varied slightly from variable to variable, depending on the quality of the source information. For example, one of the authors made use of an extraneous variable previously coded by the other author, on "Bans on Political Parties," because two of the categories exactly matched categories in the "freedom of organization" variable. An unusually high proportion (22 percent) of the freedom of expression ratings were based on only one source because Charles Humana's *World Human Rights Guide* (Humana 1986) contained especially complete, hard information. These variations occurred only because it was a new research project, and procedures evolved as they were being carried out. In future measurements, standardized, consistent procedures can be followed.
10. The research assistant was a foreign-born, U.S.-educated master's student in international relations. Her minimal training consisted of reading portions of Dahl's *Polyarchy* and a preliminary version of this paper (with the ratings removed). She consulted the same sources that we used originally.
11. For a lucid discussion of construct validity, see Nunnally (1967, 83-102).
12. See the discussion under "Implications for the Measurements of Polyarchy."
13. The ratings can be found in Raymond Gastil, ed., *Freedom in the World: Political Rights and Civil Liberties, 1987-1988* (Lanham, Md.: University Press of America, 1988), Table 6.

References

- ALEXANDER, ROBERT J., ed.
1982 *Political parties of the Americas: Canada, Latin America, and the West Indies*. Westport, CT: Greenwood Press.
- BANKS, ARTHUR S., ed.
1985 *Political handbook of the world*. New York: McGraw-Hill.
- BANKS, ARTHUR S. and ROBERT B. TEXTOR
1963 *A cross-polity survey*. Cambridge: M.I.T. Press.
- BOLLEN, KENNETH A.
1988 Political democracy: Conceptual and measurement traps, paper prepared for conference on "The Measurement of Democracy" held at the Hoover Institution, Stanford University, May 27-28, 1988.
1980 Issues in the comparative measurement of political democracy, *American Sociological Review* 45 (June): 370-390.

- DAHL, ROBERT A.
 1971 *Polyarchy: Participation and opposition*. New Haven: Yale University Press.
 1982 *Dilemmas of pluralist democracy*. New Haven: Yale University Press.
- DAHL, ROBERT A. and CHARLES LINDBLOM
 1953 *Politics, economics, and welfare*. New York: Harper and Brothers.
- DELURY, GEORGE E., ed.
 1983 *World encyclopedia of political systems and parties*. Harlow, Essex, UK: Longman.
- DEPARTMENT OF STATE
 See United States Government
- DUNN-RANKIN, PETER
 1983 *Scaling methods*. Hillsdale, N.J.: L. Erlbaum.
- EUROPA
 1986 *The Europa yearbook*. London: Europa Publications.
 1986a *South America, Central America and the Caribbean*. London: Europa Publications.
 1986b *The Far East and Australasia*. London: Europa Publications.
 1986c *Africa south of the Sahara*. London: Europa Publications.
 1986d *The Middle East and North Africa: A survey and directory of the countries of the Middle East*. London: Europa Publications.
- FUKUI, HARUHIRO, ed.
 1985 *Political parties of Asia and the Pacific*. Westport, CT: Greenwood Press
- GASTIL, RAYMOND L.
 1986 *Freedom in the world: Political rights and civil liberties, 1985-1986*. Westport, CT: Greenwood Press.
- GORDEN, RAYMOND L.
 1977 *Unidimensional scaling of social variables: Concepts and procedures*. New York: Free Press.
- HUMANA, CHARLES
 1986 *World human rights guide*. New York: Facts on File.
- McHALE, VINCENT E., ed.
 1983 *Political parties of Europe*. Westport, CT: Greenwood Press.
- McIVER, JOHN P. and EDWARD G. CARMINES
 1981 *Unidimensional scaling*. Beverly Hills: Sage Publications.
- NUNNALLY, JUM C.
 1967 *Psychometric theory*. New York: McGraw-Hill.
- PHILLIPS, CLAUDE S.
 1984 *The African political dictionary*. Santa Barbara, Calif.: ABC-Clío.
- The statesman's year-book
 1986 *Statistical and historical annual of the states of the world*. London and New York: Macmillan, St. Martin's Press.
- United States Government, Department of State
 1986 *Country reports on human rights practices for 1985*. Washington: United States Government Printing Office.
- Watch Committees, The Lawyers Committee for Human Rights
 1986 *Critique: Review of the Department of State's country reports on human rights*. New York: Fund for Free Expression for the Watch Committees and the Watch Committees and the Lawyers Committee for Human Rights.