

*Coppedge 2012 – Democratization and Research Methods*

Conceptualization -> Operationalization -> Measurement -> Indicator

Questions for discussion:

- 1) See Held's list in Coppedge pp. 15-16 (Figure 2.1). How could we to operationalize these 76 characteristics? In order to use these to define democracy, we would have to make a lot of judgment calls about exactly *how much* a particular characteristic matters. Is this a problem?
- 2) Coppedge thinks that the tradeoff he identifies between thick and thin concepts is only a result of resource scarcity. He writes: "If political scientists had more resources, we could probably develop concepts of democraticness that would be richly descriptive of all countries (and all historical periods) to an equal degree." Is this view correct, or might there be conceptual reasons for the tradeoff between thick and thin that will resist accumulation of additional information?
- 3) The philosophical tradition dealing with universals may be useful to us here. Universal concepts ("cat", "chair", "homework") cover a variety of particular instantiations. Imagine taking a "thick" approach to such concepts along the lines taken (Coppedge suggests) by David Held. We could develop a list of 72 particular characteristics of "cat" – but many of them would be mere accidental characteristics ("uses a litter box", "is a household pet") in the same way that many elements on Linz's list ("Workplace democracy", "proportional representation") seem to be incidental to democracy. Are universals necessarily thin concepts?
- 4) Munck advises us to avoid both maximalist and minimalist definitions of our concepts. Does this mean he is urging a compromise between thick and thin theories?
- 5) Does genuine multidimensionality vitiate a concept? In other words, does it make sense to search for democracy in the world if we believe that it is made up of (pace Dahl) contestation and inclusiveness? Shouldn't we just search for contestation and inclusiveness? See figure 2.3.

Coppedge explores the consequences of disaggregating measures of democracy. He distinguishes between **thick and thin conceptions of democracy**, and he argues that the ultimate goal "is to have concepts and indicators that are useful: ones that establish an easy and natural correspondence between the symbols in our minds and the observable features of the real political" pp.13-14

"there is a trade-off between a concept's intension (the number of defining attributes it has) and its extension (the range of countries to which it can be applied) (1970)."

"Unfortunately, **different preferences for thin or thick concepts lead scholars to talk past one another**: when qualitative and quantitative analysts say "democracy," they literally mean different things." p.23

"To these three dimensions - contestation, breadth of inclusion, and fullness of inclusion - I would add two more: the division of powers and the scope of democratic authority." p.26

"In measurement theory, accuracy has two aspects - validity and reliability. Validity [is] the extent to which an indicator measures what one claims it measures...[and] reliability is the degree to which a measurement procedure produces the same measurements every time, regardless of who is performing it...reliability is also a function of the unidimensionality of the components of an indicator because, in practice, **unidimensionality is a matter of degree.**" p.29

"The limitations of existing democracy indicators are partly the result of the multidimensionality of democracy." p.29

”Paradoxically, therefore, **one way to measure democracy better is to stop measuring democracy and simply measure its component dimensions instead.** This disaggregated strategy has the additional advantage of making it possible to explore empirically the interrelationships among dimensions, which would open up a fascinating new avenue for research.” p.31

“Most measurement of democracy now is ordinal, so if we wish to develop a single indicator of democracy in several dimensions, we will have to find ways of measuring dimensions at the interval level or higher. One way to do this is to reformulate the attributes of democracy in terms of probabilities.” p.33

“...the dimensions that structure a thick concept are best thought of as handy bundles of a larger number of potential dimensions. Such bundles probably hold together only for selected periods and places. **The more diverse the sample and the longer the expanse of time it covers, the more likely it is to resist reduction to a small number of dimensions.**” p.35

“Thick concepts are most easily measured with nominal or ordinal data, whereas thin concepts lend themselves more naturally to interval or ratio data.” P.38

“there have been heated debates about whether democracy should be measured by dichotomous or continuous indicators. On one side, **Sartori (1987) and Przeworski (2000) argue that democracy is inherently dichotomous...**On the other side, **Dahl (1971), Bollen (1990), and others argue that there are degrees of democracy and a continuum of democraticness,** ranging from very democratic to highly undemocratic regimes. For them, higher levels of measurement improve the accuracy and reliability of democracy indicators. I agree more with Collier and Adcock (1999), who argue that almost any concept can be considered as either categorical or continuous.” p.40

**“I suspect that we are not likely to achieve much improvement in reliable and valid measurement until we begin working with a thicker, multidimensional concept of democracy.”** p.42

### ***Munck 2009 – Measuring Democracy***

Conceptualization -> Measurement -> Aggregation

“...academics must recognize that good data are produced through the combination of methods, theory, and empirics, and strive to combine these skills.” p.12

“the most useful—if admittedly flexible—methodological suggestion that can be offered is that scholars should avoid the extremes of including too much or too little in a definition.” p.16

“Most constructors of indices subscribe to a procedural definition of democracy and thus avoid the problem of maximalist definitions...In contrast, **the problem of minimalist definitions is quite widespread.**” p.17

“no matter how many levels of abstraction are introduced, attributes at the last level of abstraction, generically labeled as “leaves,” are used as the starting point for the task of measurement.” p.21

“Moving beyond the initial step of identifying **what** attributes are deemed to be constitutive of a concept, analysts must also consider **how** these attributes are related to each other and, more specifically, take explicit steps to ensure the *vertical organization of attributes by level of abstraction.*” p.21

“the identification of multiple attributes of a concept essentially amounts to a process of disaggregation, which immediately raises the question of **how the disaggregate data might be aggregated**” p.21

“it is necessary that less abstract attributes be placed on the proper branch of the conceptual tree—that is, immediately subordinate to the more abstract attribute it helps to flesh out and make more concrete. On the other hand, attributes at the same level of abstraction should tap into mutually exclusive aspects of the attribute at the immediately superior level of abstraction.” p.22

“...even when concepts have been extremely well articulated, these leaves are rarely observable themselves. Hence, it is necessary to form measurement models relating unobservable “latent variables” to “observable variables” or indicators.” p.23

“The first decision in the formation of measures is the *selection of indicators* that operationalize the leaves of a concept tree.” p.24

Munck identifies two common problems for measurement formation. The first is “the failure to recognize the manifold empirical manifestations of a conceptual attribute and to properly use multiple indicators,” and the second is “the failure to appreciate the inescapable nature of measurement error.”

“the common practice of using data already coded by others is strongly associated with a tendency to simply sidestep the need to justify the choice of indicators.” p.25

Munck thinks that scholars do not “give much attention to the subjective judgments that shape the selection of “objective” indicators.” p.25

“there is no foundation to the widespread perception that the selection of measurement levels is something that is decided solely by reference to a priori assumptions.” p.25

“the choice about measurement level might be seen as an attempt to avoid the excesses of introducing distinctions that are either too fine-grained, which would result in statements about measurement that are simply not plausible in light of the available information and the extent to which measurement error can be minimized, or too coarse-grained, which would result in cases that are well known to be different being placed together.” p.26

Munck recommends that “analysts should *record and make public*: (1) their *coding rules*, which should include, at the very minimum, a list of all indicators, the selected measurement level for each indicator, and sufficiently detailed information so that independent scholars can interpret the meaning of each scale; (2) the *coding process*, which should include the list of sources used in the coding process, the number of coders, and the results of any intercoder reliability tests; and (3) the *disaggregate data* generated on all indicators... With respect to the coding process, existing indices do quite poorly... In sum, existing indices of democracy have not tackled the challenge of measurement very well. More generally, it is fair to state that **existing indices fail on numerous grounds**. They do little to select indicators that reflect a sensitivity to context, problems of equivalence, and measurement error. They tend to rely on a fairly unsophisticated approach to the selection of measurement level. And they do not take adequate steps to ensure replicability.” pp.27-29

“In the conventional view, taken largely from psychometrics, aggregation involves a delicate balancing act... Aggregation is seen as justified inasmuch as the lower levels of aggregation tap into a unidimensional phenomenon (empirical question)” p.30

“A different rationale for aggregating disaggregate scores is that the disaggregate scores, taken by themselves, are parts of a whole and that their meaning is understood only when the parts are considered in context or, in other words, when the parts are combined to form a whole... Thus, in tackling the decision to generate an index, analysts should first specify how they think about the relationship between the indicators used to measure a concept and the concept being measured and, specifically, which of these two perspectives they adopt... With regard to these various tasks, existing datasets on democracy once again are less than adequate.” pp.31-33

“if the aggregation of two attributes is at issue and one’s theory indicates that there is no interaction among the attributes, one would simply add the scores of both attributes. If one’s theory indicates that the two attributes are both necessary conditions, one could multiply both scores. And, if one’s theory indicates that the two attributes are sufficient conditions, one could take the score of the attribute with the highest value.”

“To highlight only the most notable strengths and weaknesses, **praise is most justified in the cases of ACLP (Alvarez et al. 1996)**, who are particularly insightful concerning the selection of indicators and especially clear and detailed concerning coding rules; **Coppedge and Reinicke (1991)**, who display a concern with coder reliability and sensitivity on the question of levels of aggregation; and **Hadenius (1992)**, who offers a compelling conceptualization of democracy, an appropriate choice of indicators, and a sophisticated use of aggregation rules. **Datasets that are unfortunately so problematic as to require explicit mention include those compiled by Freedom House (2007b), Gasiorowski (1996), and Vanhanen (2000, 2007)**, which exemplify problems in all three areas of conceptualization, measurement, and aggregation (see table 2.4)... no single index offers a satisfactory response to all three challenges of conceptualization, measurement, and aggregation... the best indices are also fairly restricted in their scope (see table 2.1), while the indices with the broadest scope, with the partial exception of Polity IV, are not among the strongest on issues of conceptualization, measurement, and aggregation.” pp.35-36

“an implication of the discussion is that democracy is a relative value and that an analysis of democracy cannot limit itself to democracy. The very justification for promoting democracy actually hinges on the answer to an empirical question—Are there trade-offs between democracy and other political values?—that calls for an analysis that goes beyond democracy.” p.120

“proposals to replace Schumpeter’s minimalist definition [of democracy] with broader definitions, including those advanced in the recent literature on the quality of democracy, have been quite weak, giving credibility to the argument that, conceptual shortcomings notwithstanding, a minimalist definition of democracy is preferable because its parsimony makes it analytically clearer than its alternatives and hence more suitable for purposes of empirical analysis.” p.123

“...even when a robust procedural definition of democracy is adopted, there are other political values beyond democracy. For example, given the potential for abuses of state power, it might be wise to protect the rule of law by placing some issues out of reach of the demos. The implications of this point for academic research and political action that takes democracy as its master concept and value are vast. Once political values other than democracy are acknowledged, the normative valuation that lies at the heart of research on democracy and the promotion of democracy—more democracy is always better—becomes a matter for discussion and has to be justified in relation to other political values.” pp.128-129

### ***Cheibub, Gandhi and Vreeland – Democracy and Dictatorship Revisited***

Cheibub and his co-authors believe that democracy measures aren’t interchangeable, and they set out to assess the strengths and weaknesses of existing measures. They argue that a dichotomous, minimalist

measure is the optimal approach, and they assert that alternatives to this measure are themselves bimodal, gaining little analytical purchase despite their additional detail.

“We have, thus, a six-fold classification of political regimes that emphasizes the institutions capable of removing the government from power. This classification is rooted in the distinction between democracies and dictatorships, that is, those regimes in which executive and legislative offices are filled through contested elections and those in which they are not. Among democracies, we distinguish those that are parliamentary (only the legislature can remove the government), presidential (only the president can remove the government) and mixed or semipresidential (the legislature can remove the government and there is a directly elected head of state). In dictatorships we distinguish those that are monarchic (family and kin networks remove the government), military (the armed forces remove the government) and civilian (a residual category often characterized by the presence of a political party as the institution capable of determining the fate of existing governments)... Because it is based on observational data, it is reproducible, a characteristic that is not present in any of the existing alternative measures of political regimes.” p.97

***Collier and Adcock – “Democracy and Dichotomies”***

“In our view, generic claims that the concept of democracy should inherently be treated as dichotomous or graded are incomplete. The burden of demonstration should instead rest on more specific arguments linked to the goals of research. We thus take the pragmatic position that how scholars understand and operationalize a concept can and should depend in part on what they are going to do with it.” p.537

“The choice of a dichotomy in effect places this distinction at what is traditionally viewed as the lowest level of measurement (Stevens 1946, Roberts 1976:492.93). This choice thereby appears to underutilize more fine-grained information that may routinely be available about differences among regimes. Yet both Sartori and Przeworski and collaborators are convinced that this lowest level of measurement is more valid in conceptual terms. This puzzle points to a question: What, indeed, are the grounds for viewing this as a valid dichotomy, and not a false dichotomy?” p.538

“It is widely believed that, in relation to many phenomena in the real world, it is possible to identify an inherently correct name or system of names. People believe it is possible to identify *the* name of a thing, the one that tells what it really is.. (Brown 1958:17; see also Lakoff 1987:9, 118.21). One possible consequence could be that the idea of bounded wholes is uncritically embraced for the wrong reasons.” p.544

“in the face of changing social reality, shifting definitions of the subject matter, and evolving theoretical understanding and empirical knowledge, conceptualizations that initially serve to justify a dichotomy based on a particular cutpoint can subsequently break down... More broadly, in many domains of knowledge, what initially appeared to be clear distinctions sometimes break down and are reconceptualized.” p.545

“Overall, it can be argued that in a variety of disciplines, (a) dichotomies are of central importance, (b) the cutpoints that establish dichotomies may evolve over time, and (c) scholars face an ongoing choice between retaining the use of dichotomies, based on these potentially evolving cut-points, or shifting to a graded approach that employs multiple cut-points.” p.545

“the specific goals of standardization and rigor are most productively pursued in conjunction with a realistic focus on how other goals influence the use and application of concepts... it is more realistic to aim for an accumulation of knowledge grounded in mutual comprehension among scholars who self-consciously recognize their conceptual decisions as real choices.” p.562

### **Collier and Levitsky – Democracy with Adjectives**

Tradeoff between analytic differentiation and conceptual validity.

“that a recognition of the contested status of a given concept [like democracy] opens the possibility of understanding each meaning within its own framework.” p.433

“Sartori builds on a basic insight about the organization of concepts: a significant aspect of the relationship between the meaning of concepts and the range of cases to which they apply can be understood in terms of a "ladder of generality. 15 This ladder is based on a pattern of inverse variation between the number of defining attributes and number of cases. Thus, concepts with fewer defining attributes commonly apply to more cases and are therefore higher on the ladder of generality, whereas concepts with more defining” attributes apply to fewer cases and hence are lower on the ladder.” p.434

“Moving down the ladder of generality provides useful differentiation, and the subtypes just noted play an important role in the recent literature. Yet subtypes formed in this manner may leave the analyst vulnerable to conceptual stretching, because they presume the cases under discussion are definitely democracies...Sartori's proposal for avoiding conceptual stretching is to move up the ladder of generality to concepts that have fewer defining attributes and correspondingly fit a broader range of cases...Although climbing the ladder of generality helps to avoid conceptual stretching, it has an important drawback. Because these subtypes remain more general than the concept of democracy, this approach leads to a loss of conceptual differentiation. Thus, taken together, Sartori's two strategies can advance one or the other of these goals, but not both at once.” pp.435-436

Classical subtypes – “Classical subtypes are understood as instances of the root definition<sup>18</sup> of democracy in relation to which they are formed, at the same time that they are differentiated vis-?-vis other classical subtypes of this concept. Thus, "parliamentary democracy," "multiparty democracy," and "federal democracy" are all considered definitely democratic (by whatever standard the author is using), at the same time that each is considered a particular type of democracy.” p.435

Diminished subtypes – “...in contrast to the classical subtypes discussed above, diminished subtypes are not full instances of the root definition of "democracy" employed by the author who presents the subtype. For example, "limited-suffrage democracy" and "tutelary democracy" are understood as less than complete instances of democracy because they lack one or more of its defining attributes...Because diminished subtypes represent an incomplete form of democracy, they might be seen as having fewer defining attributes, with the consequence that they would be higher on the ladder of generality and would therefore provide less, rather than more, differentiation. However, the distinctive feature of diminished subtypes is that they generally identify specific attributes of democracy that are missing, thereby establishing the diminished character of the subtype, at the same time that they identify other attributes of democracy that are present. Because they specify missing attributes, they also increase differentiation, and the diminished subtype in fact refers to a different set of cases than does the root definition of democracy” p.438

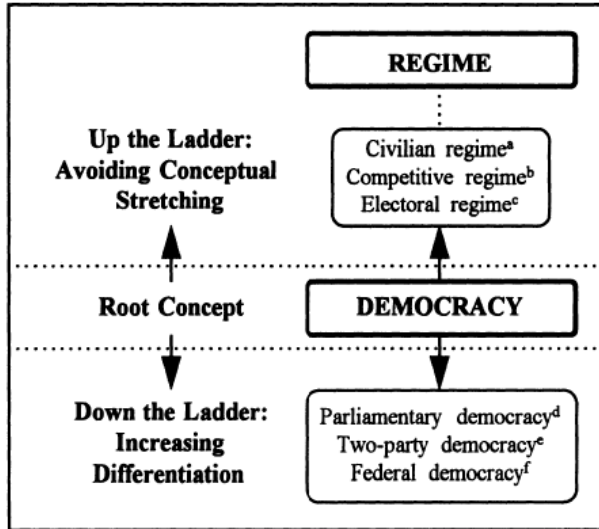


FIGURE 1  
THE LADDER OF GENERALITY:  
INCREASING DIFFERENTIATION VERSUS AVOIDING  
CONCEPTUAL STRETCHING

Coppedge: Put Figure 2.2 and Table 2.5 and Table 4.5 on TV