CONTENTS

1

Introduction

AN ONTOLOGICAL-SEMANTIC APPROACH TO CONCEPTUALIZATION AND MEASUREMENT

> To define a thing, is to select from among the whole of its properties those which shall be understood to be designated and declared by its name; the properties must be very well known to us before we can be competent to determine which of them are fittest to be chosen for this purpose.
>
> The essence of a thing … is that without which the thing could neither be, nor be conceived to be.
>
> Every proposition consists of two names [concepts]: and every proposition affirms or denies one of these names, of the other.… Here, therefore, we find a new reason why the signification of names, and the relation generally, between names and the things signified by them, must occupy the preliminary stage of the inquiry we are engaged in.
>
> J. S. MILL

JOHN STUART MILL began his *System of logic* with a "book" devoted to concepts. Starting with concepts was a logical choice since they are key building blocks for constructing theoretical propositions. Propositional logic involves the proper manipulation of symbols. For this to have usefulness in science, these symbols need substantive and empirical content.

Concepts are measured en route to statistical and empirical evaluations of theoretical propositions. So concepts must be quantified and measured in many social science applications. However, as Lazarsfeld and Barton state, one must conceptualize before moving to empirical analyses:

> Before we can investigate the presence or absence of some attribute … or before we can rank objects or measure them in terms of some variable, *we must form the concept of that variable*. (Lazarsfeld and Barton 1951, 155; emphasis is mine)

Philosophers, lawyers, political and social theorists debate normative concepts such as democracy, justice, and human rights. Concepts are fundamental to description. For anthropology, ethnography, grounded theory, and similar methodologies, developing concepts is a core theoretical and empirical activity. Concepts are thus core in causal theories, normative philosophy, and empirical description.

This book provides a unified framework for working with, constructing, and evaluating concepts that applies in these different domains. These domains often overlap. To use an example dear to my heart, which will appear as an ongoing example, "peace" is normative, descriptive, and a domain of causal theories and empirical testing of those theories (Goertz, Diehl, and Balas 2016).

Concepts are fundamentally about meaning, semantics, and ontology. Thus a methodology of concepts must to a significant degree be about semantics and ontology (which is a theory of being):

> Concepts are an answer to "what is" questions.

To ask *what* democracy *is*, or *what* poverty *is*, etc., is to inquire about ontology and definition. Concepts are about meaning and semantics.

As such, this volume uses terminology rarely seen in the literature on indicators, measurement, etc. Concepts are about definitions, semantics, ontology, meaning, and the like. *Downstream* we often want to have numeric, quantitative measures of these concepts—the

so-called indicators. J. S. Mill started with "names" in his *System of logic*, not indicators of names.

It should be stressed that this ontology can vary and be contested on various grounds. For example, gender analyses of concepts contest the bias of their traditional ontologies, be it democracy, the welfare state, etc. This is closely related to normative issues that underlie many concepts. The concept and data for transitional justice are inherently normative. The ontology might vary depending on purpose and use. Finally, I'm personally working on the concept of civil or intrastate negative peace (Goertz et al. 2019), which has different versions depending on whether I'm focusing more on it as a dependent variable or as an independent and causal variable. This becomes quite evident in the guidelines regarding independent and dependent variables at the end of chapter 3.

The basic framework, described in detail in the next chapter, provides a methodology for analyzing, critiquing, and creating complex concepts (which might be given quantitative expression). *Complex concepts are multidimensional and multilevel.* Multidimensional appears semantically in definitions that include multiple attributes and features. They appear in data sets with multiple coding rules. Multidimensional is typically quite obvious; for example, in the next chapter see the analysis of the *Multidimensional* Poverty Index (MPI; Alkire et al. 2015).

The multilevel character of concepts in contrast has received almost no explicit attention. As discussed in the next chapter, the logic varies as one moves up and down the different levels. The logic of definitions is one of completeness and nonredundancy. Going back to Aristotle—and in philosophy in general—a good definition gives the set of necessary and jointly sufficient conditions for a concept. The logic of the data level is usually *multiple indicators of one defining dimension*. This is a logic of redundancy with no real need for completeness. One wants many—or enough—indicators of a one-dimensional concept: indicators are good when they are redundant.

This implies that the mathematics of the levels tends to differ significantly, which then means that final quantitative expression of the

complex concept should typically incorporate different logics and the appropriate mathematics when generating the final numeric value.

In spite of the primordial importance of concepts, they have received relatively little methodological attention. At the same time there is a booming industry in complex *indicator* and *index* construction, such as the Multidimensional Poverty *Index*. International institutions such as the World Bank, the United Nations, OECD, and the EU—not to forget many prominent NGOs—generate hundreds of complex indicators. There has been a surge of books on indicators, such as Merry (2016) and Laurent (2018). Kelley and Simmons have an ongoing project examining dozens of "Global Performance Indicators" (2015); see also the Broome et al. project, e.g., 2018, for another long list of global indicators, as well as Weaver's global indices project.[1] All of these indicators are complex multilevel and multidimensional concepts.

An equally long list, and perhaps an even bigger industry, involves measures of physical and mental health or, more accurately, measures of illness, sickness, and disability. For example, I will refer on occasion to the DSM manual, which is the bible for clinical psychology and psychiatry (American Psychiatric Association, various years). It consists of a set of concepts, aka mental illnesses, along with sets of symptoms or indicators of the illnesses. Governments generate and use for crucial policy decisions many health indicators. For children there is the US Department of Health and Human Services Children's Bureau's child well-being measure, UNICEF's State of the World's Children measure, and Stirling Children's Wellbeing Scale, among others (see Alexandrova 2017 and Hausman 2015 for nice discussions of health concepts and scales). These are all very complex—multidimensional and multilevel—indices.

The terms "index" and "indicator" come from the idea of pointing, for example with one's index finger. These are all indicators of *something*. This book focuses on that "something," which is a concept.

1. http://www.ipdutexas.org/global-indices-project.html

One core feature of the basic framework is that conceptually it works top down. It asks for definitions which then connect downward to empirical indicators and data. Quantitative measures work from the bottom up: one starts with the data–indicators and then moves up to the top level. There is no real separation between conceptualization and measurement: they are fused.

In the appendix to the next chapter I contrast this ontological-semantic approach to conceptualization and measurement with latent variable models. These present a radically different approach, focused much more on indicators and measurement and much less on ontology and semantics. *Latent variables are cause indicators.* The ontological–semantic approach does not have causal relationships within the basic framework. The core criteria of completeness and nonredundancy for concepts make no sense in the latent variable framework. In the basic framework the focus is first and above all on the definitions and concept structure and then secondarily on the empirical indicators; for latent variables the focus is on "measurement models."

More generally, *all data sets rest on concepts.* It is hard to imagine the data being good if the underlying conceptualization is problematic, as is the case for example with terrorism. It is not uncommon for there to be disconnects between conceptualization and measurement, as we will see for the Polity democracy concept-measure. In short, this volume provides a methodology for the analysis of data sets of all sorts. The basic framework sees "coding rules" as conceptualizations with a certain aggregation structure. To understand data sets one needs a semantic and conceptual interpretation of the coding rules. "Coding rules"—always in the plural—implies multidimensional concepts.

Mill in the epigraph starts with "to define" and he continues with "its properties": this is about semantics, meaning, and ontology. He follows with "a thing" that is something in the real world, identifying and locating those "things" in the world. He then moves to "propositions," which are causal claims about the world. This volume focuses on the central role that concepts play in description and causal hypotheses as well as in normative analyses.

The Conceptual Juggling Act

> My notion would be, that anything which possesses any sort of power to affect another, or to be affected by another, if only for a single moment, however trifling the cause and however slight the effect, has real existence; and I hold that the definition of being is simply power.
>
> PLATO

> A property carries its [causal] capacities with it, from situation to situation.
>
> NANCY CARTWRIGHT

> The failure to explain is caused by a failure to describe.
>
> BENOÎT MANDELBROT

Developing valid concepts for social science involves the juggling of multiple conceptual balls. I use the juggling metaphor because typically the focus is on one or two of the balls and the others are left to fall to the ground. Some people focus on one or two balls in their research, while others focus on other balls. This volume argues that one needs to keep all of the balls in the air in one's mind when thinking about concepts, and eventually downstream for measurement.

Figure 1.1 illustrates the nature of the juggling problem. All these aspects of concept development and analysis will appear prominently in the chapters to come. Given the emphasis on measurement in many contexts, it is worth noting that these balls typically do not appear in research design or measurement books. In contrast, the first three chapters of this volume focus in particular on these conceptual balls and it is only in later chapters that measurement appears. This does not mean that measurement will not be implicitly, or sometimes explicitly, present in the discussion, because one of the key issues is linking conceptualization with measurement.

In the middle of the figure lies the semantic ball. A core question is what does one mean by a concept? This means that conceptualization is about definitions. For example, in some areas, such as political and

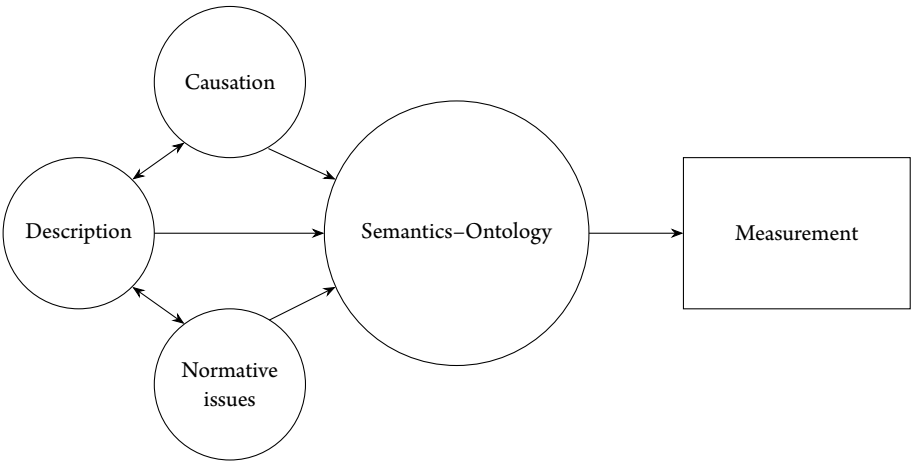

FIGURE 1.1. Juggling conceptual balls.

moral philosophy, the majority of the analysis is about the definition of some key political concept.

If one focuses exclusively on definitions, one risks getting a completely nominal view of concepts and their ontology. Hence it is absolutely critical to keep the three balls to the left of semantics in the air when discussing semantics. How does one determine what one should put into the semantics of concepts? As the arrows in figure 1.1 illustrate, there are three important factors or issues that should determine the semantic content of concepts in social science.

At the core of the methodology of concepts is the connection between the semantics of concepts and real-world phenomena. One dimension of this is the degree to which concepts and measures correctly capture and describe the world. For example, the degree to which Central American countries were democracies in the early 20th century is very debatable (Bowman et al. 2005). These debates combine the various possible conceptualizations of democracy with an application of those concepts to the reality of these countries in that time period. In general these are *both* conceptual and empirical questions.

Hence one conceptual ball is descriptive validity or what might be called empirical validity. A core goal of all sciences—natural and

social—is to accurately describe the world. This means that this book rejects a nominalist view of concepts and, in contrast, takes a realist view of them. Good concepts accurately describe the world.

Often in psychology one hears about "constructs," which leaves the ontological character of the concept ambiguous. Remaining with psychology, depression is a real thing and the goal of the researcher is to conceptualize and describe depression accurately. In short, concepts have descriptive purposes and can be evaluated in terms of their empirical accuracy. One of my core examples will be concepts and measures of poverty: perhaps the central goal of these indices is to accurately describe the level of poverty in the world.

Figure 1.1 indicates via the arrow that descriptive validity influences the semantics of concepts. Hence, empirical accuracy should be a factor in determining semantics. The real world should influence definitions, because concepts are about ontologies in the world.

Concepts and numeric measures often lie at the beginning of the process of causal analysis. This was Mill's reason for beginning with a discussion of names. Causal inference depends critically on getting concepts and measurement right.

Figure 1.1 includes an arrow from causation to semantics. To state the obvious, concepts and measures are used in causal mechanisms, hypotheses, and theories. Issues of causation must, or at least should, play a role in determining the semantics of concepts. For example, in chapter 3 a guideline states that one should use features of causal mechanisms to conceptualize and define concepts. Not surprisingly, this also makes the methodology presented here a realist one. Concepts need to tap into the workings of the world; that needs to inform conceptualization. Hence it is completely appropriate to use the epigraph from Nancy Cartwright to begin this section. Valid concepts identify causal features of the world that constitute their causal powers.

Causation and causal inference appear at various points in this volume. Causal mechanisms play a key role in deciding upon defining dimensions when the concept in question is an explanatory factor. Normally the researcher will choose attributes that are parts of causal mechanisms and that have significant causal capacities. This is the

point that Cartwright is making: core constitutive features of the phenomenon travel with their causal powers. Those causal powers are key to the conceptualization.

It is useful to distinguish causal relationships within concepts from causal relationships between concepts. Causal relationships between concepts are when they function as independent or dependent variables in some theory or causal mechanism. Causation is also an important feature of analysis *within* concepts. Figure 1.1 is about causal relationships within concepts.

Causal relationships within concepts lie at the core of the latent variable approach to measurement. Typically, the unobserved latent variable causes the indicators. I like to call this the disease–symptom model, since diseases cause symptoms. One *infers* diseases based on symptoms. But one would be wrong to confuse the symptoms with the disease: the disease *is not*—an ontological claim—the symptoms.

One might also argue that there are causal relationships between defining dimensions of concepts. This comes up quite often in the literature on poverty and human well-being:

> Development consists of the removal of various types of unfreedoms that leave people with little choice and little opportunity of exercising their reasoned agency. The removal of substantial unfreedoms, it is argued here, is *constitutive* of development.… The intrinsic importance of human freedom, in general, as the preeminent objective of development is strongly supplemented by the instrumental effectiveness of freedoms of particular kinds to promote freedoms of other kinds. The linkages between different types of freedoms are empirical and causal, rather than constitutive and compositional. (Sen 1999, xii)

Sen illustrates how ontological, aka constitutive, dimensions may have causal relationships between them. At the end Sen is making strong causal claims about the relationships between the constitutive elements of "development as freedom."

On the dependent variable side the defining dimensions influence causal hypotheses because they are part of what is being explained. For

example, if one includes women's suffrage as part of democracy then one might need to think about the causes of women's suffrage on the independent variable side of things.

Issues of causal inference lie at the core of many aspects of concept construction because of the various roles that concepts play in constructing and evaluating causal hypotheses.

Figure 1.1 includes a normative ball being juggled as well. Obviously, political and moral philosophies focus on the relationship between normative issues, semantics, and conceptualization. But many other research questions implicitly or explicitly involve normative issues. The whole area of international indicators and indices usually has normative dimensions: for example, one should increase transparency, one is concerned with the very poor, one wants to understand peace.

Unsurprisingly, social scientists have tended to avoid, dance around, and not explicitly address normative issues in conceptualization. It has really been philosophers of social science who have seriously explored the connection between normative issues and the various other conceptual balls. Yet these connections are fundamental in both the social as well as health sciences.

Much of the capability approach to human well-being has been normative as well as descriptive. What is frequently absent is how this can be linked to causal questions of various sorts. For example Robeyns discusses the linkage between the capability approach and explanatory issues:

> Nevertheless, the notions of functionings and capabilities in themselves can be employed as elements in explanations of social phenomena, or one can use these notions in descriptions of poverty, inequality, quality of life and social change.… This raises the question of whether the capability approach should aspire to do this kind of explanatory capabilitarian analysis. Another very important task of the capability approach is to reach out to those disciplines in order to make bridges between the normative and the explanatory analyses—one valuable element of the truly post-disciplinary agenda to which the capability approach aims to contribute. (Robeyns 2017, 142–43)

The huge literature and industry on physical and mental health indicators illustrates the importance of normative considerations. For example, in the history of the DSM, homosexuality was for decades considered a mental disorder. DSM-I (American Psychiatric Association 1952) defined homosexuality as a disorder as follows:

> OOQ-x63 Sexual deviation:
> This diagnosis is reserved for deviant sexuality which is not symptomatic of more extensive syndromes, such as schizophrenic and obsessional reactions. The term includes most of the cases formerly classed as "psychopathic personality with pathologic sexuality." The diagnosis will specify the type of the pathologic behavior, such as homosexuality, transvestism, pedophilia, fetishism and sexual sadism (including rape, sexual assault, mutilation). (DSM-I, 39–40)

Obviously, today such a claim would not be widely accepted. This is also highly problematic from a normative perspective. In retrospect—and even at the time—it seems amazing that homosexuality is in a list together with rape and mutilation. It took Stonewall and gay activism in the 1970s to finally incite the DSM to remove homosexuality as a mental disorder in the DSM-III supplement (American Psychiatric Association 1987).

Similar issues arise when deafness and other "handicaps" are considered illness or not-healthy. For example, "according to the HUI(3), the value of the health state of being deaf and having no other health deficiencies on a 0–1 scale is .465. According to the HUI(3), two years of life for someone who is deaf produces fewer QALYs [quality of life years] than one year of life for someone in full health. In contrast, many in the deaf community deny that deafness is a disability at all. This assertion is not sour grapes: many in the deaf community decline the partial restoration of hearing made possible by a cochlear implant, thereby showing an effective preference for deafness over partial hearing" (Hausman 2015, 90–91).

The concept and literature on "transitional justice" deals directly with normative issues of "crime," "wrongdoing," and the like. The

During Conflict Justice data set focuses exclusively on "justice" as crimes or wrongdoings related to the events of a particular conflict. The data set includes information on six forms of addressing wrongdoing: trials, truth commissions and commissions of inquiry, reparations, amnesties, purges, and exiles (Loyle and Binningsbø 2018, 445). It is not at all obvious that amnesty is justice, even though it is included in data sets on transitional justice. Typically, amnesty means that justice is not being served. So to include amnesty in the concept of transitional justice is to make a clear—and in my opinion completely invalid—normative decision. The United Nations takes the same stance in refusing to include amnesty in its list of standard and approved transitional justice mechanisms (United Nations Secretary General 2010).

The philosophy of social science literature has dealt with the connect between normative and empirical issues within conceptualization using the concept of "mixed hypotheses." Alexandrova defines them in this way: "A hypothesis is mixed if and only if (1) It is an empirical claim about a putative causal or statistical relation. (2) At least one of the variables in this claim is defined in a way that presupposes a moral, prudential, or political value judgement about the nature of this variable" (Alexandrova 2017, 82). Tiberius signals the importance of "mixed" claims:

> Without identifying them as mixed claims, philosophers have noted normative content in concepts of efficiency, rape, spousal abuse, unemployment, divorce, inflation, aggression, health and specific diseases, and, of course, well-being. My notion of a mixed claim captures these examples. What has not been done is to settle whether mixed claims should be part of science and if so what rules they should obey. (Tiberius 2004)

Most of the concepts discussed in this volume are mixed in this philosophical sense. The normative ball is often present, if most often not in the air but rather left on the ground.

One should be suspicious of any claim that says a social science concept has no normative underpinnings or dimensions. For example,

Rotberg argues that his influential indicator of governance is not very normative:

> My work on governance (here and earlier), in contrast, has been predicated on governance—and hence good governance—being defined as the performance of national, provincial, municipal, and other governments in delivering specific articulated political goods. The designation of such political goods is much more descriptive than normative; the political goods in question are those that citizens everywhere—in every culture, in every political jurisdiction—expect their governments (and their political leaders) to provide. (Rotberg 2018, 35)

Whether many or few people demand a political good does not make it more or less normative in character. One should be wary of claims about "no normative judgement." Some describe the concept of democracy as a decision-making procedure. This is true, but there is a significant set of normative values about individual liberties, rights, etc. that underlie democracy as a decision-making procedure.

In summary, figure 1.1 argues that semantics should be driven by considerations other than just pure meaning. Because concepts are used to describe the world, to explain the world, and to evaluate the world, these criteria need to feed into the semantics of the concept.

Also in the figure is an arrow from the semantics–ontology ball to the measurement square. The methodology of this arrow constitutes a very large part of this volume. A central methodological question is the connection of all of these aspects of concept construction to numeric measures. Numeric measurement appears in data sets as well as in variables in statistical hypotheses, essentially anything which receives numeric values (this includes dichotomous presence–absence as measurement). In short, connecting conceptualization with numeric variables and the construction of data sets forms a core part of the methodology of concepts.

Once one places concepts in the larger social science research enterprise, the nature of conceptual juggling becomes clearer. Concepts–measurement plays four core roles in the research enterprise:

- They are used in descriptive inference. One often asks questions about the evolution of poverty or peace in the world; the answer requires concepts and corresponding measures.
- One uses concepts as independent and explanatory factors. Thus the features that constitute independent variables must in general play some role in causal mechanisms and causal explanations, otherwise why would one include features that have no causal relevance in the concept?
- Concepts and measures appear as dependent variables. Thus to explain the level, occurrence, or changes in those dependent variables one must be explaining the attributes that constitute the concept.
- Concepts are used in case selection and scope decisions. These are not fundamentally different from independent or dependent variables. One cannot have a complete causal inference research design without making important case selection decisions that apply concepts.

These uses of concepts and measures naturally align with the concept juggling balls. They are important in conceptualization because they play core roles in social science research.

Concept Semantics and Measurement: Aggregation, Scaling, and Transformations

> The questions "What counts as a measurement of (physical quantity) *X*?" and "What is (that physical quantity) *X*?" cannot be answered independently of each other.
>
> BAS VAN FRAASSEN

Unlike most "methodological" treatments, which focus more on the statistics and mathematics of measurement, my treatment starts with meaning, conceptualization, and ontology. Later chapters in the book explore in detail mathematical, operational, and measurement issues, but this chapter and the next two focus on meaning, semantics, and

ontology. The mathematically inclined will already see much of what is coming down the pike in later chapters because the mathematics of concept structure already implies *classes* of mathematical procedures for measurement. I stress "classes" because the mean, OR, AND really form classes of options and not a specific mathematical function.

One of the most important and hard methodological, conceptual, and measurement challenges lies in connecting semantics and ontology with aggregation, scaling, and measurement. Concepts are about definitions, ontology, and semantics. At the same time we want quantitative measures for these concepts. How can one make the connection closely and smoothly so that measurement reflects as much as possible meaning and ontology? Various chapters in this volume address aspects of this challenge, in particular issues of making data–indicators commensurable (i.e., scaling) and aggregating numeric data–indicators, as well as defining attributes. This must all remain as faithful as possible to the meaning and semantics of the concept.

Aggregation becomes critical when it comes to quantification and measurement. Many researchers want a numeric value for a concept for each observation for some population or scope. The UN wants HDI data for all countries for as many years as possible. The same is true for democracy data sets.

Core to concept structure is the problem of aggregation. The continuum of aggregation involves *degrees of substitutability*. To analyze this substitutability continuum I use a combination of fuzzy logic and public goods economics, i.e., production and utility functions. The three aggregation classes below are not a free-floating typology, but rather constitute alternatives along a continuum of substitutability. There are three zones of the substitutability continuum, but it is still a continuum. It is useful to consider the logics that underlie the three zones:

- Weakest link: the minimum is the core aggregating principle, along with functions that produce smaller values than the minimum, no substitutability.
- Mean: the mean is typically used for "indicators of," moderate substitutability.

- Best shot: the maximum or sum of the defining features, extensive substitutability.

The metaphor of weakest link comes from the literature on the production of public goods (e.g., Sandler 1992), which uses the idea that the strength of a chain is the value of the weakest link. In mathematical logic this is the minimum aggregation rule used in fuzzy logic as well as QCA. The best shot comes also from the literature on public goods, and signals that what really matters is the best of the dimensions. Economics examples often involve technology: what matters is not the average technology or product but the best one. In the fuzzy-logic context this is the maximum or OR aggregation function.

For economists the standard way to address aggregation is via utility and production functions. For example, consumption poverty analyses rely on some utility function. Some utility functions are weakest link (including the classic Cobb–Douglas production function), others are mean (e.g., the generalized mean, known to economists as the constant elasticity of substitution function), and others are best shot. They vary in the degree of substitutability between dimensions or indicators. Fuzzy logic is clearer on this because it formalizes the three zones and includes aggregation procedures smaller than the minimum and larger than the maximum.

Chapter 6 explores some of the core issues surrounding concept structure and aggregation choices. One of the tensions is that between the definitions and the mathematical operations used in aggregation. For example, it is quite common for the semantics of the concept to invoke necessary and sufficient conditions while the aggregation operation uses sums and means (see the Polity example in the next chapter). It is also quite possible to use different aggregation schemes for different levels of the basic framework. As discussed at length in the next chapter, the logic of aggregation often varies between levels.

Creating numeric measures in complex concepts means deciding on aggregation rules and the structure of the concept. This is unavoidable. Even in purely conceptual, philosophical, and moral philosophy

discussions one must take—usually implicitly—a position on concept structure (aka aggregation). For example, Kantian moral philosophy typically invokes weakest link procedures while utilitarianism allows for high levels of substitutability.

Chapter 7 on hybrid concepts continues the discussion of chapter 6 on aggregation. It introduces hybrid structures, particularly those where instead of using the same aggregation procedure across a level, one uses a mixture of logical AND and logical OR. Hybrid structures are also those that use different aggregation procedures at different levels. Chapter 10 focuses on a popular hybrid where the secondary level uses logical AND while the data-indicator level uses logical OR. Hybrids seem particularly appropriate for gray zone concepts where often one requires that the gray zone not be at either end of the spectrum, i.e., two necessary conditions, but then uses different aggregation procedures—logical OR or family resemblance—for the middle, gray zone. Hybrid concept structures are a critical part of any conceptual and measurement toolbox.

Concept structure is also key factor in the discussion of intension–extension, where intension is the ontology of the concept and extension is empirical coverage. Chapter 9 analyses how concepts can be looser or stronger in terms of the number of cases that they cover, i.e., cases that score high in terms of their membership in the concept. The intension–extension distinction was imported into political science by Sartori (1970) and then discussed by Collier and his students (e.g., Collier and Mahon 1993). The key fact is that by removing necessary defining features, the empirical coverage of the concept increases. Chapter 9 discusses how concept structure in general, including best shot and weakest link, relates to empirical extension. Family resemblance was the philosophers response to requirements of necessary defining features: conceptualization without any necessary conditions at all. Family resemblance is another way to think about concept structure ranging from weakest link to best shot depending on the family resemblance rule. The relationship between intension and extension is illustrated concretely using the Democracy–Dictatorship data set (Cheibub et al. 2010).

In practice, for a variety of reasons—philosophical, practical, policy, theoretical, and causal—one must aggregate. To do this aggregation requires some important decisions about the common scale. Aggregation across dimensions is required to move up to the next level. For example, one popular option is to convert all measures to the standard deviation scale; latent variable measurement models usually do this.[2] So the question in general is not *whether* to aggregate but *how* to aggregate. One needs a framework for thinking about aggregation decisions. This is a core goal of this volume.

In a more philosophical vein, scholars argue that dimensions are incommensurate. For example, in *Disadvantage*, Wolff and De-Shalit argue vigorously against reducing all the elements of lives to a single numerical scale. They argue that aggregation across the diverse elements of lives is bound to miss things of vital importance. For example, Sen describes "non-commensurability" as "a much-used philosophical concept that seems to arouse anxiety and panic" (2009, 240). In philosophy, incommensurability is used as an argument against aggregation at all (e.g., Chang 1998). In practice a complex concept or definition involves putting together different components. So at some important semantic level one cannot even construct a definition if one is strict about the noncommensurability of defining features.

To aggregate across dimensions requires a common scale. However, it might well be the case that different dimensions, data, or indicators are of different scale types. From psychology one has the classic typology of scales, including nominal, ordinal, interval, and ratio. Very common in research design textbooks is the conventional wisdom (dating to the 1950s) about these scales. Aggregation procedures usually minimally require interval data, and most often require ratio or real numbers. Ideally one needs methods of converting dichotomous, ordinal, and interval data into ratio variables, which are required for most mathematical operations.

Some approaches, e.g., economics, virtually assume ratio or real numbers. This is because to do virtually any complex mathematical

2. It appears to be common among experimenters to evaluate causal effects in terms of standard deviations.

operation one needs ratio variables. One frequently sees nominal variables converted into interval ones with no justification, e.g., adding up a bunch of 0–1 nominal indicators. It is sometimes not clear whether a quantitative measure is interval or ratio. Chapter 5 addresses some of the generally ignored issues and common problems of scaling. One advantage of the fuzzy-logic approach is the provision of a methodology for converting various scale types into real-valued numbers.

Aggregation across different dimensions or indicators is often routinized. However, it cannot be stressed enough how important the semantic issues involved are. Extremely common for example is to standardize via linear transformation of each dimension so that the minimum is 1 and the maximum is 100 (or something equivalent). The core semantic equivalence question lies in what constitutes "high" or "low" for each dimension. To assess this, one must explore the connection between the numeric measures and meaning. To use my canonical example, how much GDP per capita does it take for a country to be wealthy or poor? Creating common scales for dimensions or indicators is a central semantic operation.

Chapter 4 introduces a core methodological device, the "semantic transformation." One can connect semantics, meaning, and definitions with some related quantitative measures or indices. A principal example is how one could connect concepts such as poverty, wealth, and economic development with data such as GDP per capita. Chapter 4 is devoted to exploring the relationship between indicators such as GDP per capita and concepts such as poor country. It provides a methodology for tightening the connection between what we mean in practice by such concepts as poor country and common measures of poverty. The methodology of semantic transformations also provides a framework for thinking about standard default transformations such as dichotomizing and linear relationships between quantitative measures and meaning. The semantic transformation methodology also proves critical in thinking about how to transform ordinal data to ratio which is essential to aggregation operations.

The semantic transformation methodology allows one to compare and contrast different options. A common tension is between concepts that include thresholds versus linear transformations. One option is to

use straight GDP per capita as a measure of poverty. Often one uses logged GDP per capita. Almost all poverty measures, such as MPI and Sen–Nussbaum, use thresholds. Dichotomization is in fact a threshold operation.

There is really an underlying continuum here as well, which might be called "degree of thresholdness." Linear transforms have no threshold. S-curve relationships have varying degrees of thresholdness depending on how sharp the S-curve is. Dichotomous transformations are the extreme threshold where there is a vertical threshold leap. For example, typically poverty measures are linear from zero until the threshold (poverty line) and then flat above that. Headcount measures of poverty use the extreme dichotomization threshold transformation.

The methodological issue is how to link semantics with numeric measures. For social science and social policy to move together it means that this linkage needs to be made explicit and justified. Fuzzy logic provides a framework for linking data-indices and concepts. I encourage researchers to take existing data and use them to the extent possible to incorporate what they mean by the concept in question.[3] I will take GDP per capita data and see how one might semantically transform them to fit better with concepts such as poverty and wealth.

Sets of Concepts, Concept Pairs, Bipolar Concepts, and Typologies

Concepts in this introduction focus on what the next chapter calls the "positive pole," such as human well-being, poverty, and peace. In fact, concepts often come in pairs, one of which is seen as the "negation" or "opposite" of the other. Examples appearing frequently in this volume include democracy–autocracy, peace–war, and poverty–wealth. To give them a name I refer to them as concept pairs or "bipolar concepts." For example, the Polity measure of democracy–autocracy is

3. In the previous edition of this book (a chapter cut for space reasons in this edition) I took the Polity measure of democracy and reworked it to conform better with the concept of democracy of the Polity researchers.

literally "democracy minus autocracy," where there is an autocracy concept as well as a democracy concept. Finally, concepts often come in sets under the rubric of typologies. For example, there are a variety of typologies of autocracy. All of these options raise the question of the relationship *between* connected concepts.

The position defended here is that the negation of a positive pole concept such as democracy is literally "not-democracy." The conceptual claim—which as we will see has huge causal implications—is that autocracy is not the same as not-democracy. One can see this often in semantic practice when people state that "peace is not just the absence of war." A person making this statement is rejecting the idea that peace = not-war.

Bipolar concepts are often fundamentally incomplete. A core example from the next chapter deals with the World Bank's conceptualization of poverty using individual or household consumption data. Poverty is conceived of as low levels of consumption. However, consumption by itself runs from low to high as a concept. One must apply additional criteria to consumption to get to the concept of interest, which is poverty. Many social science concepts have this feature. For example, "education" as years of schooling is typically incomplete. The linkage to theories and hypotheses is critical: one does not have hypotheses about education *tout court* but rather about "educated" or "uneducated." Similarly, one has hypotheses about democracy, not the bipolar democracy–autocracy.

In addition to bipolar concepts where one is the "opposite" of the other, one often sees sets of concepts. Scholars usually refer to these sets as typologies. In chapter 8 I define a typology as a set of concepts, typically more than two.[4]

A set of concepts can—and often does—contain bipolar pairs. For example, Schmitter's (1974) very influential typology focused on the positive pole of corporatism, with its opposite of pluralism, along with the additional two types "monist" (e.g., USSR) and "syndicalist."

4. I do not treat the separate issue of typologies of causal relationships.

Chapter 8 argues that typologies created following standard advice are fraught with problems in multiple ways. The standard advice is to make typologies that are (1) mutually exclusive and (2) exhaustive. Following this advice leads to major methodological, empirical, and causal problems. I use Geddes's (Geddes et al. 2014) influential typology of autocratic regimes as my main example; however, these problems are inherent in standard typology methodology, i.e., anyone following these methodological rules is going to run up against the same problems as Geddes.

While this volume focuses mostly on individual concepts, one cannot ignore the issue of bipolar concepts. In addition, scholars love typologies—sets of concepts—which may or may not include bipolar pairs, and which may or may not be somewhat ordered (i.e., not a nominal set of concepts).

In short, there are sets of concepts that have various kinds of relationships with each other, as bipolar pairs or as members of typologies. My methodology of concepts provides a framework for analyzing these common conceptual situations.

Complex Concepts Producing Complex Theories

J. S. Mill started his *System of logic* with a "book" on concepts because they are used as components of scientific propositions. Chapter 10 looks at how multilevel concepts appear in theories. The basic framework—presented in the next chapter—is about one concept. What happens when you begin to put complex concepts together into a theoretical model that makes causal claims about the world?

Chapter 10 outlines what I call two-level theories. Complex concepts are multidimensional and multilevel. When you begin to put them together into theories, those theories then also become complex. In particular, theories become multilevel: I focus on theories that have two theoretical levels, which not surprisingly I call two-level theories. The multidimensionality feature of these complex theories is less novel and interesting. People are completely used to multivariate analyses.

However, the multilevel nature of theories is much less present in research.[5]

One can construct a large variety of two-level theories with differing concept structures and causal relationships. Chapter 10 focuses on a very common one illustrated by Skocpol's famous *States and social revolutions*. It turns out that many famous scholars have independently arrived at this two-level theoretical structure. I use a variety of examples to explore the key variants on the basic model, including Skocpol's theory of social revolutions, Ostrom's work on common pool resource institutions, Kingdon's influential model of agenda setting, and Hick's study of the causes of the welfare state. With these various examples we will see different ways to build concepts, as well as several kinds of causal and noncausal relationships within and among dimensions and levels.

Conclusion

Concepts are theories about ontology: they are theories about the fundamental constitutive elements of a phenomenon. While many quantitative scholars may find the term "ontological" provocative and many interpretivists may object to my usage, I use the term in a straightforward way to designate the core characteristics of a phenomenon and their interrelationships. For example, we can ask about what constitutes a welfare state. Typically, these are states that provide goods and services like unemployment insurance, medical services, and retirement benefits. To *be* a welfare state *is* to provide these goods and services.

Good concepts pick out the causally relevant factors in phenomena. To use Cartwright's nice terminology (1989), defining features of concepts pick out the causal capabilities of phenomena. For example, a good conceptualization of democracy includes those features that are relevant in causal mechanisms about the impact of democracy.

5. There exists a huge literature on "hierarchical models" (e.g., Gilman and Hill 2007), which is quite different from what I present. The variables–concepts in these models are still treated in a single-level way.

As such, this volume rejects a strong nominalist, Red Queen, *Alice in Wonderland,* view of concepts.[6]

Good concepts accurately describe the world. Many people want to know whether poverty is increasing or decreasing in the world (Alkire et al. 2015; Ravallion 2016), or whether the world is getting more peaceful and less violent (Pinker 2011; 2018). To answer these questions requires good empirically descriptive concepts.

Good concepts often require serious normative justification. One's conceptualization of human well-being must be at least partially normative.

Because concepts serve various purposes, like causal analysis, description, and normative analysis, their ontology can and should vary according to purpose.

In terms of social science, I propose a causal, ontological, and realist view of concepts. It is an ontological view because it focuses on what constitutes a phenomenon. It is causal because it identifies ontological attributes that have causal capacities and that play a key role in causal hypotheses, explanations, and mechanisms. It is realist because it involves a descriptive analysis of the phenomenon. My approach stresses that concept analysis involves ascertaining the constitutive characteristics of a phenomenon that have central causal capacities. These causal powers and their related causal mechanisms play a role in our theories. A purely semantic analysis of concepts, words, and their definitions is never adequate by itself, at least in social science, and probably not in moral and political philosophy either.

Thus this volume embraces a realist philosophy of science like that described by Kitcher:

> Minimal realism holds that there are objects independent of human cognition. Strong realism adds the thesis that, independently of us, these objects are assorted into natural kinds and that there are causal processes in which they participate. The task of

6. It is notable that Lewis Carroll was also a prominent logician, and that informs many great parts of *Alice*.

> science is to expose the causal structure of the world, by delineating the pre-existent natural kinds and uncovering the mechanisms that underlie causal dependencies. (Kitcher 1992, 104)

Concepts are about these "natural kinds" that participate in causal mechanisms and causal explanations.

Most important concepts we use are *multidimensional* and *multilevel* in nature. Many, if not most, of the core social science concepts are complex in this sense. This volume provides a systematic methodology for analyzing and constructing complex concepts along with the appropriate corresponding quantitative measures.

INDEX