Received: 30 April 2019 | Revised: 28 June 2019 | Accepted: 1 July 2019

DOI: 10.1111/soc4.12727

ARTICLE

WILEY

If all you have is a hammer, everything looks like a nail: Operationalization matters

Christina R. Steidl[1] | Regina Werum[2]

[1]Department of Sociology, University of Alabama in Huntsville, U.S.A

[2]Department of Sociology, University of Nebraska - Lincoln

Correspondence
Christina R. Steidl, Department of Sociology, University of Alabama in Huntsville, Huntsville, AL 35899.
Email: christina.steidl@uah.edu

Funding information
National Science Foundation, Grant/Award Numbers: 1727969, 1728044

Abstract

Operationalization has been the focus of less research than many other methodological topics. In this article, we argue that considering operational decisions is particularly critical for those who study stratification, because measures of inequality often involve multiple layers of operationalization: researchers first decide how to assign individuals to social groups (e.g., race), which are then themselves used to construct measures of group-level differences and inequality (e.g., racial segregation). We provide examples of this by drawing on contemporary debates about how to operationalize social groups based on class, race, gender, and religion. Then we discuss three examples (religion, racial segregation, and family type) of second layer operationalization decisions, focusing on the consequences of operational decisions for research findings. We conclude by discussing the broader implications of operational decisions, focusing particularly on issues of power and applications for policy makers.

1 | INTRODUCTION

This article addresses why researchers need to think carefully about how to "operationalize" anything and everything. To operationalize means to develop measures for a concept, and it is typically viewed as a methodological issue. But as we show below, it is also important because operationalization impacts how we see and understand patterns of social inequality. This article addresses operationalization within the broader context of quantitative research on social inequality or stratification. To engage our readers, we begin with a seemingly simple question related to our own research: Is there a gender gap in science, technology, engineering, and mathematics (STEM)?

Answers to this question are central to ongoing public debate about whether a shortage of STEM professionals exists and, if so, about its causes, extent, and dimensions (Anft, 2013; Martin & Carnevale, 2013; NRC, 2006; NSB,

2014; Teitelbaum, 2014). Might some of the disagreements regarding this issue be traced back to discrepancies in operationalization? As researchers studying the gender gap in STEM, we wondered to what extent researchers' operationalization of "STEM" (i.e., which fields were counted as STEM) might have influenced findings related to the gender gap. Comparing the proportion of STEM bachelor's degrees earned by women across four federal agencies' measures of STEM, we found substantial variation: women earn only 36.5% of STEM degrees when we counted those fields defined as STEM by the Department of Education, but women earn 55.9% of STEM degrees when we counted those fields defined as STEM by the National Institutes of Health, which use a broader measure that includes medical fields. These patterns hold even controlling for other factors in multivariate statistical models (Steidl, Werum, Harcey, Absalon, & MillerMacPhee, 2019). In other words, how we operationalized STEM produced either a gender gap that substantially favored men, or one that slightly favored women. Of course, many readers will note that our findings depend not only on how we operationalize STEM, but also on how we operationalize gender (who do we count as a man? And who do we count as a woman?), an issue to which we will return later!

While writing up our findings, we were reminded that our experience is part of a classic phenomenon that comes up across the social sciences—and beyond (e.g., Is Pluto a "planet"? How do we operationalize the concept of "planet"?). Perhaps some of our readers recall similar classroom exercises in college-level methods courses? Yet despite its relevance to both quantitative and qualitative scholars, operationalization has been the focus of less research than many other methodological topics. A quick Google Scholar search of articles published in *Sociological Methodology* and *Sociological Methods & Research* produces only 194 results for "operationalization OR operationalize" from 1980 to 2019, but 1,224 results for "sampling" and 1,436 results just for "regression." Yet, the STEM example above clearly highlights how the way we operationalize concepts impacts how we see and understand patterns of social inequality. We contend that careful consideration of operationalization is critically important for researchers, policy makers, and practitioners.

To make the case for our argument, this article reviews the critical role operationalization plays during multiple phases of social research. Although we focus primarily on quantitative research, many of our points apply to qualitative research as well. We begin by providing a brief synopsis of what operationalization does/does not entail, then illustrate how operationalization impacts research findings, highlighting key areas of research and prominent debates about operationalization within the field of social stratification. We conclude by discussing the broader impacts of operational decisions made by researchers.

2 | DEFINING OPERATIONALIZATION

The term "operationalization" is typically used to signify the process by which social scientists create empirical (data-based) indicators to measure broader (theoretically based) concepts (Adcock & Collier, 2001). Operationalization is not the same as conceptualization, although these two processes often occur together (Carr et al., 2017).

Conceptualization refers to the process of concretely defining an abstract concept (Hox, 1997). Let's imagine that we want to investigate whether playing video games impacts academic performance. This analysis involves two very broad abstract concepts: video games and academic performance. First, we must conceptualize or define these concepts more concretely. What characteristics define a game as a "video game"? Do all games played on a screen or monitor count? If I play solitaire on my phone, is that a video game? Then, what do we mean by academic performance? There are many aspects of academic performance. Are we concerned about grades? Classroom behavior? Performance on standardized tests?

After conceptualization, we can begin operationalizing our concepts. For quantitative researchers, this often involves either constructing survey questions or choosing particular measures in secondary data. Continuing our example above, if we were designing our own survey, we might ask students how many hours per day they spend playing games on a computer or gaming console. Alternately, we could list a variety of *types* of video games (e.g., first person shooter games, massively multiplayer online (MMO) games, puzzle games, and online card games) and ask

students to report which they have played during the past week. With regard to academic performance, we might ask students to report their GPA. If we were using an existing secondary dataset instead of gathering our own data, we would choose those variables that best fit our conceptualization.

In quantitative research, the processes of conceptualization and operationalization occur prior to analysis; in qualitative research, they may occur before, during, and even after data collection (Carr et al., 2017). Qualitative researchers often seek to understand how participants understand particular concepts. For example, in a qualitative interview, a researcher might ask what kind of video games a participant plays. Such a question allows the *participant* to conceptualize the concept of video game, enabling the researcher to better understand the meaning that individuals attach to such abstract concepts. Operationalization in qualitative research also frequently includes the development of a coding schema. For example, a researcher studying violence in video games might code each game according to the following schema: (a) scene(s) resulting in the spilling of blood, (b) scene(s) including a character intentionally causing another character physical harm, (c) scene(s) resulting in the on-screen death of a character, and (d) scene(s) wherein a character uses a weapon to threaten physical harm to a character.

As the examples above suggest, choosing to operationalize a concept in different ways introduces the possibility that findings might differ precisely because the measurement differs. In fact, research on operationalization has shown that how researchers operationalize key concepts greatly influences the results of empirical analyses (Firebaugh, 2008). We saw this quite clearly in our own research with regard to the gender gap in STEM. Lay readers may mistakenly assume that differences in how researchers operationalize—and the differences in empirical findings reported—are a sign of disagreement, or worse yet confusion, among scientists. Yet, variation in operationalization across studies should not be interpreted as problematic. Proper operationalization of a concept depends on the particular research context (Adcock & Collier, 2001; Mahoney, 2004). For example, the particular indicators that we use to measure democracy may depend upon the historical time period in which we are measuring it (Collier, 1999). Operationalization may also require "updating" as the social assumptions underlying concepts change (Agger, 1991).

While considering the impact of different operational decisions matters for all sociologists, it is particularly critical for inequality or stratification researchers, because measures of stratification often involve several layers of operationalization: researchers make operational decisions about how to assign individuals to social groups (e.g., race, class, and gender), which are then themselves used to construct measures of social inequality. For example, sociologists studying racial segregation can choose among a range of indices and scales to operationalize segregation, yet these measures of segregation themselves rely on the prior classification of individuals into racial groups. Below, we provide several key examples of how debates about operationalization are playing out at each of these layers in empirical research on social stratification.

3 | OPERATIONALIZING LAYER ONE: CLASS, RACE, GENDER, AND RELIGION

What we might call the first layer of operationalization in stratification research involves sorting individuals into social groups. But wait, you say—why don't we just allow survey respondents to sort themselves into groups? Even research based on self-reported group affiliations depends substantially on operational decisions made by researchers. Below, we focus on the operationalization of four group affiliations: social class, race, gender, and religion.

Debates about how to operationalize the concept "social class" stretch back to the writings of Marx, Weber, and Durkheim. Even students relatively new to sociology may be familiar with three classical yet incongruent measures of social class. The first, traceable to Marx (1848, 1894) operationalizes class as a relational measure, emphasizing key social cleavages (e.g., employer/employee) and focusing on the social relations of production. Neo-Marxist theorists like Wright (1984) have provided more complex operational schemes rooted in this relational understanding of class that incorporate not only notions of ownership but also power and control (e.g., employer/manager/employee).

The second, traceable to Weber (1922), operationalizes class as a gradational measure, focusing on the economic assets that provide life chances and material standards of living. Following this tradition, contemporary sociologists often measure class in terms of socio-economic status, a composite measure that accounts for income, education, and occupational status. Because of its complexity, socio-economic status does not lend itself to easy categorization (e.g., one might have high income combined with low level of education), and often leads to debates about whether/where class boundaries lie—what is the difference between being working class and middle class? K. Smith (2007) thus suggests that Weber's concept of class may more accurately be operationalized probabilistically, such that class boundaries lie at the point at which social mobility becomes increasingly unlikely.

A third way to conceptualize social class focuses on its subjective salience, how people "locate themselves and others within a structure of inequality" (Durkheim 1893; Grusky & Sørenson, 1998; Weeden & Grusky, 2005; Wright, 2005, pg. 182). This operationalization is even more complex. While it requires some subjective response from individuals, the particular groups with which individuals identify can range considerably—from occupational categories and lifestyles to vague categories like "middle class."

Not surprisingly, how we operationalize social class influences our empirical results. For example, Muntaner et al.'s (2010) meta-analysis of health research found that although the health literature more commonly employs neo-Weberian measures of class (partially due to availability in major survey datasets), the recent incorporation of neo-Marxist measures based on employment status (e.g., manager, supervisor, and employee) allows for the identification of new mechanisms associated with health inequalities.

A substantial body of research has also focused on operationalizing *race/ethnicity* (Campbell & Rogalin, 2006; Connelly, Gayle, & Lambert, 2016; Desmond & Emirbayer, 2009; Saperstein, Kizer, & Penner, 2016; Saperstein & Penner, 2012). Although race may appear more immutable than social class, and thus easier to operationalize, racial and ethnic categories themselves are both contested and contextual (Omi & Winant, 1987). For example, Hitlin, Brown, and Elder (2007, p. 587) found that although Hispanic is considered an ethnic category, "a significant portion of self-reported Hispanics treat that identity as a race, most often by choosing 'other' when asked for their racial classification." Likewise, the decision to allow respondents to select multiple racial identities on the 2010 U.S. Census substantially changed the proportion of Americans who identified as American Indian/Native Alaskan (Lieber, Bhaskar & Porter, 2016). Other research has focused on the order and wording of survey questions related to race/ethnicity (Hirschman, Alba, & Farley, 2000) and techniques for obtaining skin color variables (Flores & Telles, 2012; Villarreal, 2010).

The case of race helps to illustrate the multiple steps of operationalization employed in assigning individuals to social groups. During the design phase, researchers make operational decisions about potential categories to list on a survey, from which respondents self-identify (e.g., determining the number of racial categories, deciding whether/how to allow respondents to indicate multi-racial identities). These decisions depend not only on the desire to accurately describe individuals but also on methodological limitations, like the knowledge that fewer options tend to produce adequate sample sizes for comparison between groups (Durso & Gates, 2013). Another step occurs after data collection and prior to the analysis phase, when answer categories are often combined after the fact, to create categories like "White Hispanic" or "Black Hispanic."

Numerous studies have investigated how operational decisions made by researchers after data collection influence outcomes (Bailey, Loveman, & Muniz, 2013; Campbell & Troyer, 2007; Cheng & Powell, 2011; Hitlin, Brown & Elder, 2007; Mateos, Singleton, & Longley, 2009; Patterson & Dagadu, 2016; Saperstein et al., 2016; Zuberi, 2008). For example, analyzing the effects of 27 different racial/ethnic taxonomies, Harcey and Smith (2017) found substantial variation in outcomes depending on whether they assigned Hispanic individuals to a racial category (Black, White, or other) or included "Hispanic" itself as a category in their racial taxonomy. Likewise, combining racial categories with indicators of nativity (e.g., Asian immigrant and Asian American), Irizarry (2015) finds substantial within-group variation in teacher perceptions—especially among Asians and Latinx.

Recent scholarship has also taken a critical look at how we operationalize *sex and gender*. A primary critique focuses on the binary operationalization of gender (i.e., male/female) in most nationally representative surveys,

which categorically excludes transgender and nonbinary individuals (Bauer, Braimoh, Schema, & Dharma, 2017; Miller & Grollman, 2015; Sumerau, Mathers, Nowakowski, & Cragun, 2016; Westbrook & Saperstein, 2015). Most large national surveys employ a binary measure of gender. For example, the General Social Survey sorts individuals into two sexes, male and female, flagging changes in reported gender as errors to be corrected by researchers (Westbrook & Saperstein, 2015). This limits the operational decisions available to researchers using secondary data, who can only analyze what is available. Put differently, when all you have is a hammer, everything looks like a nail.

Alternatives to the binary operationalization of sex/gender frequently involve the addition of answer categories (e.g., nonbinary, intersex, transgender, and other gender) or the separation of sex from gender identity (e.g., cisgender woman, cisgender man, and transgender woman), but scholars have also called for more nuanced measures (Sumerau, Mathers, Nowakowski, & Cragun, 2016). Bauer et al. (2017) recommend that broad population surveys employ a multi-dimensional operationalization, asking about assigned sex at birth, current gender identity, and current gender expression.

Other research has focused on measuring variation within and across the categories of masculinity and femininity. Most notably, Magliozzi, Saperstein, and Westbrook (2016) developed and tested the operationalization of gender using masculinity and femininity scales and found substantial diversity among both cisgender and transgender respondents. They argue that including more nuanced measures of gender in surveys would help to expose dimensions of gender experience and gender inequality that are concealed by binary measures of sex/gender. This growing body of research drives home the extent to which operational decisions can reinforce dominant social realities, effectively erasing the voices and experiences of non-dominant groups.

Finally, we point to research on *religion.* How do we operationalize religious affiliation? Historically, researchers in the United States have relied upon self-reported denominational affiliations, but changes in the United States religious landscape have prompted some debates about the validity of such self-reports and as well as the usefulness of such a classification system. Contemporary Americans are likely to switch denominations (T. Smith & Kim, 2005), to identify as "non-denominational" (Steensland et al., 2000), or to report having "no religion" (Hout & Fischer, 2002; T. Smith & Kim, 2005). Unfortunately, operationalization of religious affiliation categories on surveys often fails to distinguish between those belonging to "non-denominational" Protestant churches and those without religious affiliations of any kind (i.e., no denomination; Woodberry and C. Smith, 1998). Moreover, using words like fundamentalist, evangelical, and charismatic in surveys has produced varied results, with some researchers finding that respondents distinguish between these groups (C. Smith, 1998) and others suggesting that the terms themselves are confusing for respondents (Smidt, Kellstedt, & Guth, 2009; Woodberry & C. Smith, 1998).

Knowing that Americans are more likely to identify with a particular congregation, not a denomination (Dougherty, Johnson & Polson, 2007; Warner, 1994), some researchers have suggested that survey questions ask respondents to identify their local congregation instead of their denominational affiliation (Dougherty et al., 2007). On one hand, operationalizing religious affiliation at such a localized level would likely increase response validity. On the other hand, it requires extensive operational decision-making at the second layer (as discussed below), when researchers must decide how to sort those reported denominations—or local congregations—into meaningful groups for comparison.

To summarize, operationalization plays a critical role in defining and reinforcing group membership and boundaries. Here, we have reviewed some of the literature on the operationalization of class, race, gender, and religion. Similar discussion has surrounded the operationalization of group boundaries with regard to sexual orientation (Better & Simula, 2015; Compton, Meadow, & Schilt, 2018; Sumerau et al., 2016), immigrant status (Anderson & Blinder, 2015; Simcox, 1997), and other fundamental categories used to identify social groups. We hope that the discussion above has made clear the extent to which race, class, gender, and religion are not simply individual-level attributes. Far from it. As social constructs, categories of race, class, gender, and religion reflect both how individuals see themselves and how society sees them. The operational decisions made by researchers also reflect these social realities and social research constitutes one location through which these social constructs are produced and maintained.

4 | OPERATIONALIZING LAYER TWO: RELIGION, RACIAL SEGREGATION AND FAMILY TYPE

The previous section focused on the first layer of operationalization within stratification research—sorting individuals into groups. This section examines the second layer of operationalization, wherein stratification researchers utilize first-layer measures to operationalize subsequent measures of group difference and inequality. We focus on three examples: religion, racial segregation, and family type.

Our first example picks up where we left off in the previous section, with religion. After gathering data on denominational or congregational affiliation, researchers must aggregate these diverse responses into larger groupings to facilitate comparison, a process that has been the source of much debate within the literature (Dougherty et al., 2007; Steensland et al., 2000; Woodberry & Smith, 1998). In the 1990s, this second layer of operationalization frequently involved organizing self-reported denominations along a fundamentalist-moderate-liberal continuum (T. Smith, 1990). This approach was so popular that the General Social Survey (GSS) included a generated variable that automatically assigned respondents to one of these three categories based on their denominational affiliation. This second layer of operationalization allowed for comparison between large groups (e.g., how do fundamentalists differ from liberals?). However, some researchers remained concerned that the fundamentalist-moderate-liberal continuum conflated political with religious attitudes and failed to accurately capture social differences between groups (Kellstedt & Smidt, 1996).

Steensland et al. (2000) introduced an alternate classification system, best known by is GSS variable name, RELTRAD. Like the fundamentalist-moderate-liberal continuum, RELTRAD sorts Protestants based on their self-reported denomination. However, RETRAD groups individuals based on structural networks among denominations, instead of using denomination as a proxy for beliefs/ideology. Put another way, instead of focusing on "What do Methodists believe?" as a means of determining religious classification, Steensland et al. (2000) asked, "With what national religious organizations (e.g., National Council of Churches) are Methodist churches affiliated?" and "Which other denominations are affiliated with these same organizations?" This new system resulted in six religious categories: mainline Protestant, evangelical Protestant, Black Protestant, Roman Catholic, Jewish, and Other (e.g., Mormon, Jehovah's Witness, Muslim, Hindu, and Unitarian). Using this operational system, RELTRAD, improves our ability to predict other social attitudes (political views, economic attitudes, etc.) and employs less confusing vocabulary (Steensland et al., 2000; Woodberry, Park, Kellstedt, Regnerus, & Steensland, 2012). By the late 2000s, the RELTRAD system had already become the standard method for coding GSS and other survey data related to religious affiliation (Dougherty et al. 2007; Woodberry et al., 2012). This example demonstrates how quickly norms of operationalization can shift, and how ongoing research on methodology can help scholars reconstruct their measures to better fit the changing landscape of the social world.

Our second example examines research on *racial segregation*. Again, measures of racial segregation rest on second-layer operational decisions, shaped by first-layer operational decisions about how to assign racial/ethnic group affiliations. Vigorous debate has arisen around how to conceptualize and operationalize racial segregation (James & Taeuber, 1985; Massey & Denton, 1998; Reardon & Firebaugh, 2002; Reardon & O'Sullivan, 2004), but two very common operational practices illustrate our point.

One way to conceptualize segregation is as the uneven *distribution* of groups across an area—e.g., to what extent do schools within a district or neighborhoods within a city appear racially representative of the larger district/city? This conceptualization is particularly useful when we try to explain structural differences in resources available to groups based on their location (e.g., housing, medical care, and food). Most researchers operationalizing segregation this way have used a dissimilarity index, which indicates the proportion of one group that would have to be reallocated in order to achieve proportional representation across subunits.

An alternate way to conceptualize segregation is in terms of group *composition*, capturing notions of diversity and inter-group interactions. Measuring segregation in this way is more relevant to explaining segregation as it

relates to interpersonal attitudes and group dynamics. In this case, segregation is best operationalized using indices of exposure or isolation (Condron, Tope, Steidl, & Freeman, 2013; Reardon & Firebaugh, 2002). The exposure index, for example, indicates the proportion of one group present in the context of a typical member of the other group—e.g., if we take a typical Black student, what proportion of their classmates are White?

Neither of these measures is intrinsically better or worse (in methodological terms they are equally valid and reliable). Yet operational decisions made by researchers can lead to very different measures of the same concept, and by extension to different research conclusions. Freeman and Steidl (2016, pp. 176–77) illustrate this point using a hypothetical school district: "Imagine that in [this] district, 100% of students are White (so, necessarily, each school is attended by 100% White students). This district has a dissimilarity index of 0 … [and] the exposure index for this district is 0." In other words, a racially homogenous school in a racially homogenous district would be considered completely integrated, when segregation is operationalized using a distributional measure, but considered completely segregated, when segregation is operationalized using a compositional measure. This example highlights how the same abstract concept (segregation), when conceptualized differently (distribution vs. composition) and operationalized differently (dissimilarity index vs. exposure index), can lead to very different research findings. Perhaps unsurprisingly then, how researchers operationalize segregation has been shown to influence both the direction and magnitude of racial stratification outcomes, ranging from preterm birth rates to achievement test scores, and from school suspension rates to neighborhoods/housing (Britton & Shin, 2013; Condron et al., 2013; Freeman & Steidl, 2016; Massey & Denton, 1998).

Our third example of the impact of operationalization on stratification research revolves around the measurement of *family type*, particularly those measures of family type explicitly tied to gender and sexual orientation. As debates about same-sex marriage swept the United States in the late 2000s, a growing body of research focused on how to identify or capture same-sex couples using large survey data sets (Black, Gates, Sanders, & Taylor, 2007; Cheng & Powell, 2015; Gates & Steinberger, 2009; Powell, Bolzendahl, Geist & Steelman, 2010). While most large surveys include questions about the respondent's marital status, and some include questions about the respondent's sexuality, few include questions about the sex or gender identity of the respondent's spouse. Decisions about how to operationalize family type are thus what we call second-layer operationalization, dependent upon the first-layer operationalization of questions related to sexual orientation, marital status, etc.

Regnerus's (2012) study comparing social and psychological outcomes of children who grew up in households with different family types brought the operationalization of family type to the forefront of sociological and public awareness. Regnerus reported that adult children whose parents had same-sex relationships tended to have worse outcomes than adult children who grew up in intact biological families. However, attempts to replicate these findings produced conflicting results causing other scholars to call into question aspects of Regnerus's methods (Cheng & Powell, 2015; Perrin, Cohen, & Caren, 2013; Rosenfeld, 2015). Cheng and Powell (2015) document precisely how Regnerus's operationalization of same-sex relationships and "intact biological families" influenced his findings. They note that, although the study purports to examine the effect of being raised in different family types, Regnerus counted respondents who were raised by their biological parents, but whose parents separated after they reached adulthood, as not being raised in an intact biological family. Instead, he created a separate category: divorced later or had joint custody. Cheng and Powell (2015) critiqued this operationalization: because those parents who later divorced may have had less stable marriages, the decision to not count these individuals as having grown up in intact biological families artificially inflated the positive outcomes for those remaining in the "intact biological families" category. Moreover, Cheng and Powell questioned Regnerus's ability to accurately identify same-sex families, which he operationalized using questions about parental romantic relationships that were not asked of respondents who reported growing up in intact biological families. In other words, first-layer operational decisions about who should be asked about parental romantic relationships resulted in inconsistencies in operationalizing same-sex families at the second layer. Correcting for these and other issues with operationalization, Cheng and Powell reanalyzed the data, and found that "the results are either inconclusive or suggestive that adult children raised by same-sex two-parent families show a comparable adult profile to their peers raised by two-biological-parent families" (2015, p. 624).

This debate again illustrates the extent to which operationalization can influence the stratification or inequality patterns our research detects. Moreover, close examination and replication encouraged by such debates constitutes optimal social science practice.

5 | DISCUSSION AND CONCLUSION

In this article, we have demonstrated the complexities associated with operationalization and reviewed several of the key debates related to operationalization within the literature on social stratification. In particular, we hope that readers have a better understanding of two things: the way that multiple layers of operationalization build upon each other over the course of a research project, and the potential for operational decisions to shape research findings. We have also noted that operationalization has received far less attention than other methodological issues. Too frequently, operational decisions are presented descriptively instead of critically—as merely a technical issue. In other words, most journals expect that researchers explain *how* they measured concepts, but not *why* they chose to measure them as such. We contend that careful consideration of operationalization is critically important for researchers, policy makers, and practitioners. In this final section then, we address how methodological decisions related to operationalization can impact the "real world"—examining issues of power and policy.

Operationalization has always been entwined with power dynamics and structures of oppression (Gamson & Moon, 2004; Omi & Winant, 1987; Rosino, 2016). Operational decisions allow researchers to create, reinforce, or challenge the social construction and meaning of social group membership. Classical theorists like Du Bois (1940), de Beauvoir (1952), and Said (1978) have documented how categories of knowledge constructed by and for the dominant group have been employed to reinforce that group's privilege and position. So how can we engage in responsible operationalization? To some extent, feminist scholarship has led the way in addressing issues of power as they relate to research methodology, in particular calling for the practice of reflexivity (Harding, 1992; Olesen, 1994): "Partiality, and not universality, is the condition of being heard; individuals and groups forwarding knowledge claims without owning their position are deemed less credible than those who do" (Collins, 1990, p. 290). Moreover, as scholars from previously under-represented groups join in the production of knowledge, they have challenged existing operational practices. Social realities themselves are not necessarily changing (there have been transgender people for a long time), but only recently have social scientists begun to take these realities into account.

Critical race scholars have also played an important role, drawing attention to how methodological practices (including the operationalization of race), arising out of a history of social control and domination, (re)produce whiteness as normative and thereby produce the racialized "other" as deviant (Zuberi & Bonilla-Silva, 2008). Thus, our scientific understanding of metabolic activity, garnered from years of historic research on primarily White subjects, creates a racially biased understanding of health that normalizes White bodies and constructs non-White bodies as aberrant (Hatch, 2016). Moreover, when we conceptualize and measure differences in outcomes based on "race" instead of based on "exposure to racism," our findings perpetuate essentialist notions of race (Hatch, 2016; Holland, 2008) and can lead to poorly designed interventions and policies. For example, Hatch (2016) argues that when we identify race as a risk factor for diabetes, instead of racial differences in access to healthy food, we will likely allocate more resources to educating Black Americans about the risks of diabetes than to building grocery stores in underserved communities.

Operationalization is also tied to social policy development and outcomes. To further illustrate, we turn again to the example of school segregation. Bussing practices became common starting in the 1960s, approximately a decade after the *Brown v. Board of Education* SCOTUS decision that declared racial segregation of schools unconstitutional, often as part of court-ordered desegregation plans (Hermalin & Farley, 1973; Olzak, Shanahan, & West, 1994). In this case, the policy goal—school desegregation—was pursued with the underlying assumption that changing the *distribution* of students across schools within a district would change their educational and social experiences.

Evaluating the effectiveness of bussing became difficult as White families fled schools and districts (Frankenberg, 2013; Logan, Zhang, & Oakley, 2017; Owens, 2016; Quillian, 2014; Saporito & Sohoni, 2006) and neighborhoods organized to change district boundaries in ways that made the district appear integrated even as schools within districts remained segregated or became more so (Breslow, Wexler, & Collins, 2014; Orfield, Bachmeier, James, & Eitle, 1997; Saporito & Sohoni, 2006). Moreover, school choice, k-12 privatization, and homeschooling have also been linked to racial demographics (Bankston & Caldas, 2000; Boutcher, Kronberg, & Werum, 2018; Davis, 2009; Logan et al., 2017; Renzulli & Roscigno, 2005). Many of these practices produced racially homogenous schools within racially homogenous districts that arguably look integrated, depending on how we measure segregation.

Moreover, sociologists of education have shown that segregation within schools may rival segregation between schools and districts. Schools have become racially segregated both in academic settings (e.g., tracking and AP classes; see Davis, 2009; Kubitschek & Hallinan, 1996) and extracurricular settings (e.g., cafeteria, clubs, and detention; see Freeman & Steidl, 2016; Moody, 2001; Tyson, 2011). This suggests the need for studying segregation both in terms of distribution (differential access to resources and opportunities within the school) and composition (fewer opportunities for students to interact with one another on an everyday basis).

The example of school desegregation highlights the importance of employing multiple methods of operationalization in designing policy, as well as evaluating policy effectiveness. We encourage readers to consider additional ways in which policies have been designed, implemented and evaluated based on empirical research, particularly by decisions related to operationalization. For instance, the abovementioned debate regarding the ideal of an "intact biological family" (see Cheng & Powell, 2015; Regnerus, 2012) clearly reverberates through ongoing public discourse and even recent federal family policy changes related to the U.S. Supreme Court's recognizing of same-sex marriages in 2015 (*Obergefell v. Hodges*).

To summarize, operational decisions have real consequences. They shape our common sense understanding of the world, justifying existing structural conditions and influencing the distribution of resources. Seemingly small operational decisions can significantly influence research findings, and thus shape policy interventions and policy impacts, intended and otherwise. We also have an obligation to remind readers that policy implementation, in turn, raises new questions requiring researchers to generate innovative ways of conceptualizing and operationalizing phenomena of interest. Ideally, this continual feedback loop is what ultimately keeps pushing forward both scientific inquiries and meaningful policy reforms. That operational decisions are too frequently addressed without justification for the sake of space in academic publishing has perhaps allowed for the development of a culture in which decisions related to operationalization are given less attention by readers and reviewers than decisions related to statistical techniques and other methodological issues. Assumptions and decisions made during the operationalization process should be subject to the same types of critical discussion as other steps in the research process.

So how can we improve awareness around issues of operationalization within the field of sociology? Ideally, researchers should operationalize their concepts in multiple ways. When multiple measures of the same concept produce similar results, we have more confidence in our conclusions and call them "robust." Realistically, multiple measures are not always possible (especially when limited data is available), and advocating for such simplistic requirements reduces our ability to study difficult subjects. We recommend that researchers provide some discussion of *why* they have made particular operational decisions, instead of simply describing *how* they operationalized concepts, and that researchers carefully consider how their operationalization may reproduce or challenge oppressive structures.

ACKNOWLEDGEMENTS

This project was supported in part by a collaborative grant from the National Science Foundation (1728044 and 1727969). The authors would like to thank Sela Harcey, Jacob Absalon, and Alice MillerMacPhee for their work as graduate assistants on the STEM and veterans project.

ORCID

Christina R. Steidl https://orcid.org/0000-0003-2143-517X

REFERENCES

Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, *95*, 529–546. https://doi.org/10.1017/S0003055401003100

Agger, B. (1991). Critical theory, poststructuralism, postmodernism: Their sociological relevance. *Annual Review of Sociology*, *17*, 105–131. https://doi.org/10.1146/annurev.so.17.080191.000541

Anderson, B., & Blinder, S. (2015). *Briefing: Who counts as a migrant? Definitions and their consequences (4th Rev)*. Oxford: The Migration Observatory at the University of Oxford.

Anft, M. (2013). The STEM crisis: Reality or myth? The Chronicle of Higher Education, November 11.

Bailey, S. R., Loveman, M., & Muniz, J. O. (2013). Measures of 'race' and the analysis of racial inequality in Brazil. *Social Science Research*, *42*, 106–119. https://doi.org/10.1016/j.ssresearch.2012.06.006

Bankston, C. L., & Caldas, S. J. (2000). White enrollment in nonpublic schools, public school racial composition, and student performance. *The Sociological Quarterly*, *41*, 539–550. https://doi.org/10.1111/j.1533-8525.2000.tb00072.x

Bauer, G. R., Braimoh, J., Schema, A. I., & Dharma, C. (2017). Transgender-inclusive measures of sex/gender for population surveys: Mixed-methods evaluation and recommendations. *PLoS ONE*, *12*(5), e0178043. https://doi.org/10.1371/journal.pone.0178043

Better, A., & Simula, B. (2015). How and for whom does gender matter? Rethinking the concept of sexual orientation. *Sexualities*, *18*, 665–680. https://doi.org/10.1177/1363460714561716

Black, D., Gates, G., Sanders, S. & Taylor, L. (2007). The measurement of same-sex unmarried partner couples in the 2000 U. S. Census. CCPR Working Paper Series CCPR-023-07.

Boutcher, J., Kronberg, A., & Werum, R. (2018). Getting on the 'radar screen': Homeschooling litigation as agenda setting, 1972-2007. *Mobilization*, *23*, 159–180. https://doi.org/10.17813/1086-671X-23-2-159

Breslow, J., Wexler, E., & Collins, R. (2014). Separate and unequal: The return of school segregation in eight charts. PBS Frontline, July 15, 2014.

Britton, M. L., & Shin, H. (2013). Metropolitan residential segregation and very preterm birth among African American and Mexican-origin women. *Social Science & Medicine*, *98*, 37–45. https://doi.org/10.1016/j.socscimed.2013.08.039

Campbell, M. E., & Rogalin, C. L. (2006). Categorical imperatives: The interaction of Latino and racial identification. *Social Science Quarterly*, *87*, 1030–1052. https://doi.org/10.1111/j.1540-6237.2006.00414.x

Campbell, M. E., & Troyer, L. (2007). The implications of racial misclassification by observers. *American Sociological Review*, *72*, 750–765. https://doi.org/10.1177/000312240707200505

Carr, D., Boyle, E. H., Cornwell, B., Correll, S., Crosnoe, R., Freese, J., & Waters, M. C. (2017). *The art and science of social research*. New York: W. W Norton & Company, Inc.

Cheng, S., & Powell, B. (2011). Misclassification by whom? A comment on Campbell and Troyer. *American Sociological Review*, *76*, 347–355. https://doi.org/10.1177/0003122411401249

Cheng, S., & Powell, B. (2015). Measurement, methods, and divergent patterns: Reassessing the effects of same-sex parents. *Social Science Research*, *52*, 615–626. https://doi.org/10.1016/j.ssresearch.2015.04.005

Collier, R. B. (1999). *Paths toward democracy: The working class and elites in Western Europe and South America*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511625626

Collins, P. H. (1990)/2008). *Black feminist thought*. New York: Routledge Classics.

Compton, D., Meadow, T., & Schilt, K. (2018). *Other, please specify: Queer methods in sociology*. Oakland, CA: University of California Press.

Condron, D. J., Tope, D., Steidl, C. R., & Freeman, K. J. (2013). Racial segregation and the black/white achievement gap, 1992 to 2009. *The Sociological Quarterly*, *54*, 130–157. https://doi.org/10.1111/tsq.12010

Connelly, R., Gayle, V. & Lambert, P. S. (2016). Ethnicity and ethnic group measures in social survey data. Methodological Innovations. https://doi.org/10.1177/2F2059799116642885

Davis, T. (2009). Laissez faire education policy: Organization and equity in school choice. Dissertation, Emory University.

de Beauvoir, S. (1952)/1989). *The second sex*. New York: Vintage Books.

Desmond, M., & Emirbayer, M. (2009). What is racial domination? *Du Bois Review: Social Science Research on Race*, *6*, 335–355. https://doi.org/10.1017/S1742058X09990166

Doughtery, K. D., Johnson, B. R., & Polson, E. C. (2007). Recovering the lost: Remeasuring U.S. religious affiliation. *Journal for the Scientific Study of Religion*, *46*, 483–499.

Du Bois, W. E. B. (1940)/2014). *The dusk of dawn*. New York: Oxford University Press.

Durkheim, E. (1893)/1996). *The division of labor in society*. New York: Free Press.

Durso, L. E., & Gates, G. J. (2013). Best practices: Collecting and analyzing data on sexual minorities. In A. Baumle (Ed.), *International handbook on the demography of sexuality, international handbooks of population, vol 5* (pp. 21–42). Dordrecht: Springer.

Firebaugh, G. (2008). *Seven rules for sociological research*. Princeton, NJ: Princeton University Press. https://doi.org/10.1515/9780691190433

Flores, R., & Telles, E. (2012). Social stratification in Mexico: Disentangling color, ethnicity, and class. *American Sociological Review*, *77*, 486–494. https://doi.org/10.1177/0003122412444720

Frankenberg, E. (2013). The role of residential segregation in contemporary school segregation. *Education and Urban Society*, *45*, 548–570. https://doi.org/10.1177/0013124513486288

Freeman, K. J., & Steidl, C. (2016). Distribution, composition and exclusion: How school segregation impacts racist disciplinary patterns. *Race and Social Problems*, *8*, 171–185. https://doi.org/10.1007/s12552-016-9174-9

Gamson, J., & Moon, D. (2004). The sociology of sexualities: Queer and beyond. *Annual Review of Sociology*, *30*, 47–64. https://doi.org/10.1146/annurev.soc.30.012703.110522

Gates, G. J. & Steinberger, M. D. (2009). Same-sex unmarried partner couples in the American Community Survey: The role of misreporting, miscoding and misallocation. Presented at the Annual Meeting of the Population Association of America. Detroit, MI.

Grusky, D. B., & Sørenson, J. B. (1998). Can class analysis be salvaged? *American Journal of Sociology*, *103*, 1187–1234. https://doi.org/10.1086/231351

Harcey, S. R. & Smith, J. A. (2017). The consequences of classification: How varying schemes of racial/ethnic classification impact sociological outcomes. Paper presented at the American Sociological Association Methodology Section midyear meeting, April 26, Chicago, IL.

Harding, S. (1992). *Whose science? Whose knowledge?* Milton Keynes: Open University Press.

Hatch, A. R. (2016). *Blood sugar: Racial pharmacology and food justice in Black America*. Minneapolis, MN: University of Minnesota Press. https://doi.org/10.5749/minnesota/9780816696178.001.0001

Hermalin, A. I., & Farley, R. (1973). The potential for residential integration in cities and suburbs: Implications for the busing controversy. *American Sociological Review*, *38*, 595–610. https://doi.org/10.2307/2094410

Hirschman, C., Alba, R., & Farley, R. (2000). The meaning and measurement of race in the U.S. Census: Glimpses into the future. *Demography*, *37*, 381–393. https://doi.org/10.2307/2648049

Hitlin, S., Brown, J. S., & Edler, G. J. Jr. (2007). Measuring Latinos: Racial vs. ethnic classification and self-understandings. *Social Forces*, *86*, 587–611. https://doi.org/10.1093/sf/86.2.587

Holland, P. W. (2008). Causation and race. In T. Zuberi, & E. Bonilla-Silva (Eds.), *White logic, white methods* (pp. 93–110). Lanham, Maryland: Rowman & Littlefield.

Hout, M., & Fischer, C. S. (2002). Why more Americans have no religious preference: Politics and generations. *American Sociological Review*, *67*, 165–190. https://doi.org/10.2307/3088891

Hox, J. J. (1997). From theoretical concept to survey question. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 47–69). New York: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118490013.ch2

Irizarry, Y. (2015). Utilizing multidimensional measures of race in education research: The case of teacher perceptions. *Sociology of Race and Ethnicity*, *1*, 564–583. https://doi.org/10.1177/2332649215580350

James, D. R., & Taeuber, K. E. (1985). Measures of segregation. *Sociological Methodology*, *15*, 1–32. https://doi.org/10.2307/270845

Kellstedt, L. A., & Smidt, C. E. (1996). Measuring fundamentalism: An analysis of different operational strategies. In J. C. Green, J. L. Guth, C. E. Smidt, & L. A. Kellstedt (Eds.), *Religion and the culture wars: Dispatches from the front* (pp. 193–218). Lanham, MD: Rowman & Littlefield.

Kubitschek, W. N., & Hallinan, M. T. (1996). Race, gender, and inequality in track assignments. *Research in Sociology of Education and Socialization*, *11*, 121–146.

Liebler, C., Bhaskar, R., & Porter, S. (2016). Joining, leaving, and staying in the American Indian/Alaska Native race category between 2000 and 2010. *Demography*, *53*, 507–540. https://doi.org/10.1007/s13524-016-0461-2

Logan, J. R., Zhang, W., & Oakley, D. (2017). Court orders, white flight, and school district segregation, 1970-2010. *Social Forces*, *95*, 1049–1075.

Magliozzi, D., Saperstein, A., & Westbrook, L. (2016). Scaling up: Representing gender diversity in survey research. *Socius*, *2*, 1–11.

Mahoney, J. (2004). Comparative-historical methodology. *Annual Review of Sociology*, *30*, 81–101. https://doi.org/10.1146/annurev.soc.30.012703.110507

Martin, M. & Carnevale, A. (2013). Is the STEM education crisis a myth? NPR Tell Me More November 25. http://www.npr.org/templates/story/story.php?storyId=247166532

Marx, K. (1848)/1978). The communist manifesto. In R. C. Tucker (Ed.), *The Marx-Engels reader* (2nd ed.) (pp. 469–500). New York: Norton W. W. & Company.

Marx, K. (1894)/1959). Capital, volume 3, chapter 9. In M. Eastman (Ed.), *Capital, the communist manifesto, and other writings* (pp. 54–59). New York: The Modern Library.

Massey, D. S., & Denton, N. A. (1998). Hypersegregation in U.S. Metropolitan Areas: Black and Hispanic Segregation along Five Dimensions. *Demography*, *26*, 373–391. https://doi.org/10.2307/2061599

Mateos, P., Singleton, A., & Longley, P. (2009). Uncertainty in the analysis of ethnicity classifications: Issues of extent and aggregation of ethnic groups. *Journal of Ethnic and Migration Studies*, *35*, 1437–1460. https://doi.org/10.1080/13691830903125919

Miller, L. R., & Grollman, E. A. (2015). The social costs of gender non-conformity for transgender adults: Implications for discrimination and health. *Sociological Forum*, *30*, 809–831. https://doi.org/10.1111/socf.12193

Moody, J. (2001). Race, school integration, and friendship segregation in America. *American Journal of Sociology*, *107*, 679–716. https://doi.org/10.1086/338954

Muntaner, C., Borrell, C., Vanroelen, C., Chung, H., Benach, J., Kim, I. H., & Ng, E. (2010). Employment relations, social class and health: A review and analysis of conceptual and measurement alternatives. *Social Science & Medicine*, *71*, 2130–2140. https://doi.org/10.1016/j.socscimed.2010.09.038

National Research Council (2006). *To recruit and advance: Women students and faculty in science and engineering*. Washington, D.C.: National Academies Press.

National Science Board (2014). *Science and engineering indicators 2014*. Arlington, VA: National Science Foundation. http://www.nsf.gov/statistics/seind14/

Olesen, V. (1994). Feminisms and models of qualitative research. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research*. London: Sage.

Olzak, S., Shanahan, S., & West, E. (1994). School desegregation, interracial exposure, and antibusing activity in contemporary urban America. *American Journal of Sociology*, *100*, 196–241. https://doi.org/10.1086/230503

Omi, M., & Winant, H. (1987). *Racial formation in the United States: From the 1960s to the 1990s*. New York: Routledge.

Orfield, G., Bachmeier, M. D., James, D. R., & Eitle, T. (1997). Deepening segregation in American public schools: A special report from the Harvard Project on School Desegregation. *Equity & Excellence in Education*, *30*(2), 5–24. https://doi.org/10.1080/1066568970300202

Owens, A. (2016). Inequality in children's contexts: Income segregation of households without children. *American Sociological Review*, *81*, 549–574. https://doi.org/10.1177/0003122416642430

Patterson, E. J., & Dagadu, H. E. (2016). Fractures in the color line: Consequences of constructions of race and ethnicity on measures of imprisonment. *Sociology of Race and Ethnicity*, *2*, 10–26. https://doi.org/10.1177/2332649215587762

Perrin, A. J., Cohen, P. N., & Caren, N. (2013). Responding to the Regnerus study: Are children of parents who had same-sex relationships disadvantaged? A scientific evaluation of the no-differences hypothesis. *Journal of Gay & Lesbian Mental Health*, *17*, 327–336. https://doi.org/10.1080/19359705.2013.772553

Powell, B., Bolzendalh, C., Geist, C., & Steelman, L. C. (2010). *Counted out: Same-sex relations and Americans' definitions of family*. New York: Russell Sage Foundation.

Quillian, L. (2014). Does segregation create winners and losers? Residential segregation and inequality in educational attainment. *Social Problems*, *61*, 402–426.

Reardon, S., & Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological Methodology*, *32*, 33–67. https://doi.org/10.1111/1467-9531.00110

Reardon, S., & O'Sullivan, D. (2004). Measures of spatial segregation. *Sociological Methodology*, *34*, 121–162. https://doi.org/10.1111/j.0081-1750.2004.00150.x

Regnerus, M. (2012). How different are the adult children of parents who have same-sex relationships? Findings from the New Family Structures Study. *Social Science Research*, *41*, 752–777. https://doi.org/10.1016/j.ssresearch.2012.03.009

Renzulli, L. A., & Roscigno, V. J. (2005). Charter school policy, implementation, and diffusion across the United States. *Sociology of Education*, *78*, 344–365. https://doi.org/10.1177/003804070507800404

Rosenfeld, M. J. (2015). Revisiting the data from the New Family Structure study: Taking family instability into account. *Sociological Science*, *2*, 478–501. https://doi.org/10.15195/v2.a23

Rosino, M. L. (2016). Boundaries and barriers: Racialized dynamics of political power. *Sociology Compass*, *10*, 939–951. https://doi.org/10.1111/soc4.12412

Said, E. W. (1978). *Orientalism*. New York: Pantheon Books.

Saperstein, A., Kizer, J. M., & Penner, A. M. (2016). Making the most of multiple measures: Disentangling the effects of different dimensions of race in survey research. *American Behavioral Scientist*, *60*, 519–537. https://doi.org/10.1177/0002764215613399

Saperstein, A., & Penner, A. M. (2012). Racial fluidity and inequality in the United States. *American Journal of Sociology*, *118*, 676–727. https://doi.org/10.1086/667722

Saporito, S., & Sohoni, D. (2006). Coloring outside the lines: Racial segregation in public schools and their attendance boundaries. *Sociology of Education*, *79*, 81–105. https://doi.org/10.1177/003804070607900201

Simcox, D. (1997). Major predictors of immigration restriction: Operationalizing "nativism.". *Population and Environment*, *19*, 129–143. https://doi.org/10.1023/A:1024619331015

Smidt, C. E., Kellstedt, L. A., & Guth, J. L. (Eds.) (2009). *The Oxford handbook of religion and American politics*. New York: Oxford University Press.

Smith, C. S. (1998). *American evangelicalism: Embattled and thriving*. Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226229225.001.0001

Smith, K. (2007). Operationalizing Max Weber's probability concept of class situation: The concept of social class. *British Journal of Sociology*, *58*, 87–104. https://doi.org/10.1111/j.1468-4446.2007.00140.x

Smith, T. W. (1990). Classifying protestant denominations. *Review of Religious Research*, *31*, 225–245. https://doi.org/10.2307/3511614

Smith, T. W., & Kim, S. (2005). The vanishing protestant majority. *Journal for the Scientific Study of Religion*, *44*, 211–223. https://doi.org/10.1111/j.1468-5906.2005.00277.x

Steensland, B., Park, J. Z., Regnerus, M. D., Robinson, L. D., Wilcox, W. B., & Woodberry, R. D. (2000). The measure of American religion: Toward improving the state of the art. *Social Forces*, *79*, 291–318. https://doi.org/10.1093/sf/79.1.291

Steidl, C. R., Werum, R., Harcey, S., Absalon, J. & MillerMacPhee, A. (2019). Defining STEM, redefining the gender gap: The role of military service in shaping STEM degree trends. Manuscript in preparation.

Sumerau, J. E., Mathers, L. A. B., Nowakowski, A. C. H., & Cragun, R. T. (2016). Helping quantitative sociology come out of the closet. *Sexualities*, *20*, 644–656. https://doi.org/10.1177%2F1363460716666755

Teitelbaum, M. (2014). The myth of the science and engineering shortage. The Atlantic, March 19.

Tyson, K. (2011). *Integration interrupted: Tracking, Black students, and acting White after Brown*. Chicago: University of Chicago Press. https://doi.org/10.1093/acprof:oso/9780199736447.001.0001

Villarreal, A. (2010). Stratification by skin color in contemporary Mexico. *American Sociological Review*, *75*, 652–678. https://doi.org/10.1177/0003122410378232

Warner, R. S. (1994). The place of the congregation in the contemporary religious configuration. In J. P. Wind, & J. W. Lewis (Eds.), *American Congregations* (Vol. 2). New perspectives in the study of congregations. (pp. 54–99). Chicago: University of Chicago Press.

Weber, M. (1922)/1946). Class, status, party. In H. H. Girth, & C. W. Mills (Eds.), *From Max Weber: Essays in sociology* (pp. 180–195). New York: Oxford University.

Weeden, K. A., & Grusky, D. B. (2005). The case for a new class map. *American Journal of Sociology*, *111*, 141–212. https://doi.org/10.1086/428815

Westbrook, L., & Saperstein, A. (2015). New categories are not enough: Rethinking the measurement of sex and gender in social surveys. *Gender & Society*, *29*, 534–560. https://doi.org/10.1177/0891243215584758

Woodberry, R. D., Park, J. Z., Kellstedt, L. A., Regnerus, M. D., & Steensland, B. (2012). The measure of American religious traditions: Theoretical and measurement considerations. *Social Forces*, *91*, 65–73. https://doi.org/10.1093/sf/sos121

Woodberry, R. D., & Smith, C. S. (1998). Fundamentalism et al: Conservative protestants in America. *Annual Review of Sociology*, *24*, 25–56. https://doi.org/10.1146/annurev.soc.24.1.25

Wright, E. O. (1984). A general framework for the analysis of class structure. *Politics and Society*, *13*, 383–423. https://doi.org/10.1177/003232928401300402

Wright, E. O. (2005). *Approaches to class analysis*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511488900

Zuberi, T. (2008). Deracializing social statistics: Problems in the quantification of race. In T. Zuberi, & E. Bonilla-Silva (Eds.), *White Logic, White Methods* (pp. 127–134). Lanham, Maryland: Rowman & Littlefield.

Zuberi, T., & Bonilla-Silva, E. (2008). *White Logic, White Methods*. Lanham, Maryland: Rowman & Littlefield.

AUTHOR BIOGRAPHIES

Christina R. Steidl is an Associate Professor in the Department of Sociology at the University of Alabama in Huntsville. Her research focuses primarily on racial and gender inequality in education and the workplace and has been published in the *American Sociological Review*, *The Sociological Quarterly*, *Symbolic Interaction*, *Critical Sociology*, *Race and Social Problems*, and *Gender, Work & Organization*. Her current project, in collaboration with

Regina Werum, investigates how military service is related to educational and career trajectories and may provide an alternate pathway into STEM for members of under-represented groups.

Regina E. Werum is a Professor in the Department of Sociology at the University of Nebraska-Lincoln. Her research revolves around the way that social policy shapes and reflects social inequalities, especially in the area of education. Mostly quantitative, her projects frequently involve comparative-international and comparative-historical data. Her research has appeared in top Sociology and specialty journals such as the *Journal of Family Issues*, *American Journal of Education*, *Social Stratification and Mobility*, *Social Forces*, *Sociology of Education*, *Mobilization*, and *Social Science History*. Prior to joining the Department of Sociology at UNL, she served as Associate Vice Chancellor for Research at the Office of Research and Economic Development at UNL and as a Program Director (Sociology) for the National Science Foundation.

How to cite this article: Steidl CR, Werum R. If all you have is a hammer, everything looks like a nail: Operationalization matters. *Sociology Compass*. 2019;13:e12727. https://doi.org/10.1111/soc4.12727