

KEVIN ROY *University of Maryland College Park*

ANISA ZVONKOVIC *Virginia Tech**

ABBIE GOLDBERG *Clark University***

ELIZABETH SHARP *Texas Tech University****

RALPH LAROSSA *Georgia State University*****

Sampling Richness and Qualitative Integrity: Challenges for Research With Families

Sampling is one of the most difficult and contentious aspects of qualitative research design. There are few guidelines for sampling decisions or for understanding saturation in qualitative family research. The authors frame the problematic of data quality in the selection of units of analysis and observation and consider how to enhance sample richness. They outline considerations for data quantity and sample size as well as case- and variable-based approaches. With multiple examples from recent and classic studies to illustrate the consequences of sampling decisions, they explore links between

saturation and validity. Finally, they encourage researchers to craft a coherent statement on qualitative integrity to demonstrate how their sampling decisions are rooted in epistemology, theory, and richness and quality of data.

Sampling is one of the most difficult and contentious aspects of qualitative research design (Small, 2009). In family research, qualitative methods can be attuned to a nuanced understanding of processes in families, to the construction of meaning-making for family members, and to exploration of family life in diverse contexts (Daly, 2007; Ganong & Coleman, 2013). But the basic merit of qualitative family research is often judged in the scholarly community simply with reference to the question “How large is the *N*?”—with the assumption that a small number of cases can doom a study to irrelevancy (LaRossa, Goldberg, Roy, Sharp, & Zvonkovic, 2014).

The problems of judging the merit of qualitative research on the basis of sample size highlight larger concerns with how qualitative research is understood and used in the field of family science because sampling is a decision not only about sample size but also about the integrity of the project’s goals, composition and depth of data, and fit with theory. Researchers

Department of Family Science, 1142 SPH Building, School of Public Health, University of Maryland, College Park, MD 20782 (kroy@umd.edu).

*Department of Human Development, Virginia Tech, 366A Wallace Hall, Blacksburg, VA 24061.

**Department of Psychology, Clark University, Worcester, MA 01610-1477.

***Department of Human Development and Family Studies, Texas Tech University, P.O. Box 41230, Lubbock, TX 79409-1230.

****Department of Sociology, Georgia State University, Atlanta, GA 30303-5020.

Key Words: ethnography, grounded theory, qualitative research.

in many other disciplines, such as nursing and public health, have addressed the complications of sampling in qualitative research (e.g., Sandelowski, 1995; Trotter, 2012). To our knowledge, there are no contemporary extended discussions of sampling for qualitative family research, although the importance of these issues was identified in the 2012 *Journal of Marriage and Family* special section "Exchange on Qualitative Research" (see Lareau, 2012; LaRossa, 2012a, 2012b; Matthews, 2012; and Roy, 2012).

We first frame the problem of quality in the selection of units of analysis and observation in the study of families and consider the challenges to enhancing the richness of samples in qualitative family research. Second, we outline considerations for the quantity of family units, or the size of a sample, as well as case- and variable-based approaches. Furthermore, we consider the importance of saturation as a goal of qualitative family research. Finally, we encourage researchers to craft a coherent statement on qualitative integrity to demonstrate how their sampling decisions are rooted in epistemology, theory, and richness and quality of data.

Throughout this article, we cite examples to illustrate the conditions and consequences of sampling decisions, and we present a hypothetical study to model some of the challenges in sampling and saturation in qualitative family research. We do not argue about the worth of a qualitative approach; neither do we compare qualitative methods with quantitative methods. We recognize that different scholars take different views on the "truth" derived from qualitative methods, if there is one or any. Our intent is to address an audience of both qualitative and quantitative researchers, including those who use or who review the use of multiple methodological approaches.

DATA QUALITY AND SELECTION OF UNITS OF OBSERVATION AND ANALYSIS

A distinction between the units of observation and the units of analysis has been made in the methodological literature, and in qualitative family research such distinctions are important as well. Sedgwick (2014) explained the difference this way:

The unit of observation, sometimes referred to as the unit of measurement, is defined . . . as the "who" or "what" for which data are . . . collected.

The unit of analysis is defined . . . as the "who" or "what" for which information is analyzed and conclusions are made.

Thus, for example, in a study that relied on individual interviews with family members to assess the nature and scope of supportive interactions within the families, the unit of observation would be the individual members who were interviewed, whereas the unit of analysis would be the supportive interactions referred to in the interviewees' narratives.

Ruano, Bruce, and McDermott (1969) pointed out that a goal of family research is to find a way to closely relate units of observation and units of analyses. When these units are incongruent, researchers may make invalid conclusions. For example, if researchers want to study family interaction by drawing on data from individuals, the fit seems to be loose at best. To provide a richer sample, we might suggest three other options: interviews with two, three, or four family members simultaneously (i.e., couple or joint interviews, also known as *family group interviews*); ethnographic observations of family members interacting in public or private; or archival studies of family correspondence (e.g., back-and-forth letters). In these designs, the unit of observation and the unit of analysis are more in sync.

Sampling Richness

With appropriate decisions about units of observation and analysis, researchers can improve the quality of data, enhancing richness and depth. As individuals change in relationships over time, we could study dyadic or triadic relationships as clustered units in marital or cohabiting couples or parents and children (McCall & Simmons, 1991; Ruano et al., 1969; Sedgwick, 2014; Sprey, 2013). Families are not confined to households or conjugal relationships, so units of analysis also may be defined as networks instead of individuals within families. Hansen's (2004) research on four care networks used observations and interviews to show how parents, friends, and extended kin combine to form supportive interrelationships to monitor and nurture children. Similarly, networks of families stretch across national boundaries, and family ethnographers have examined how these networks change over time in relation to new policies, needs of sending communities,

and capacities of family members who leave (Bledsoe & Sow, 2011; Dreby, 2010).

Scholars also study family routines or celebrations as events. We could study the range of certain experiences that occur within any one family, or we could study the range of certain ideologies or beliefs that emerge within any one family member's experiences over many years—each experience or belief as a unit of observation nested within a family or family member as a unit of analysis, or an episode of parent–child interaction might be a unit of observation. For example, in an archival study of more than 200 advice-seeking letters from fathers and mothers to parent educator Angelo Patri, LaRossa and Reitzes (1993) analyzed 1,000 mentions of parent–child interaction, and, in a classic example of family ethnography, Engels (1845, cited in Marcus, 1974) presented a case study of industrial conditions in a single city—Manchester, England—while drawing on data from multiple families (units of observation) living in poverty. Stated simply, a single sampling decision can lead to another related, undiscovered, yet clarifying sample of the same family.

Such flexibility in choosing units also opens up opportunities for qualitative studies to examine and even resolve contradictions in data. As we suggest, it is relatively straightforward to seek confirmation—and discover unexpected responses—when researchers work with multiple members of the same family. Other studies have exposed contradictions by comparing interview data with cultural objects (e.g., popular magazine articles) and personal documents (LaRossa, 2011; Meyerowitz, 1994). Classic studies of participant observation (Lareau, 2003) and mixed-methods historical projects (Elder, 1999) have used longitudinal designs to accomplish similar tasks. Contradictions that emerge at one point in time may be resolved as lives change and qualitative researchers adapt their sampling to capture such change.

Picking out contradictions in data and searching for insight into these new complexities can lead to data that grow in both depth and breadth. Families may be sampled at different points to understand changing membership, given that it varies by generation or age. Information can be obtained from family members (via interview or observation) multiple times over many months or years, with each data point being a separate unit of observation to inform the researcher

about the family. In a study of conflict and power in marriage, LaRossa (1977) interviewed 16 couples four times over the course of a pregnancy, and, in a transition-to-parenthood study, 20 couples were interviewed three times over the first 9 months postpartum (LaRossa & LaRossa, 1981). These sampling decisions expanded the number of observations, but the primary purpose was to provide richer, deeper insight into change over time. In each of these two studies, instead of adding more couples the choice was made to keep the sample size (number of couples) small to better engage the complexities of the data. Depth and extent are related, but depth hinges on questions asked and varied angles of vision that are sought.

In short, knowing how many individuals or families are observed or interviewed—the N —does not by itself tell us about the quality of the data obtained. There are many circumstances where a small- N study yields a voluminous set of high-quality information, because of the depth of information, gathered via multiple informants on the unit of analysis, repeated data collection, and so on.

Holism

An assumption that often shapes sampling criteria is that some family phenomena can be understood only by examining how they result from individuals and their interaction. For more than a century, this assumption has informed the debate over *methodological individualism* versus *methodological holism* that continues to haunt much of social science (Udehn, 2002). The distinction between methodological individualism and holism occurs in all research methods, and in this section we focus on the unique processes of emergent study design in qualitative family research. In particular, researchers may confront challenges in matching units of observation in families with units of analysis.

If we return to some of the roots of family research, we recall that Burgess (1926) noted that a family is “a unity of interacting personalities” and made the case for a holistic approach. How does such a vision stack up with the methods used to conduct contemporary family research? Even more than 40 years ago, more attention was paid to the individual than to any other analytic unit of analysis in family research (Ruano et al., 1969). Methodological individualism remains a controversial topic among

methodologists in family science as well. Family researchers have been criticized for being overly individualistic in methods and subject matter (Hagestad & Dannefer, 2001; Sprey, 2013).

Daly (2007) suggested that researchers consider how the phenomenon of interest helps determine the unit of observation. Are individuals the best unit of observation with which to examine the experiences of a family? Are families more than the sum of their parts? Are individual experiences/perceptions, relational dynamics, or family processes more central to inquiry? It would be beneficial for family scholars to consider units of observation that are in sync and holistically oriented. If a researcher were hoping to write about family networks, he or she could study 40 individuals, but there would be a disconnect between the unit of observation and unit of analysis. But if the researcher actually studied five family networks using ethnographic observations, the connection between the unit of observation and unit of analysis would be stronger. The best choice would depend on the goals of the study.

Hess and Handel (1959) challenged the received wisdom about sampling families when they presented the whole-family methodology in their book, *Family Worlds: A Psychosocial Approach to Family Life*. In a study of 33 families, they noted the nested nature of families as units of observation and obtained data from each member of the family (Sedgwick, 2014). This approach would seem to capture the imperative to study families as they are, not as we would like them to be (Daly, 2003). As Handel (1996) argued,

Either there is little recognition that families are groups whose members have individual perspectives on their family membership, or that approach is disregarded because it is not amenable to quantification or statistical analysis. To the extent that the latter perspective prevails, research in the field is being driven by a preference for certain methods rather than by the questions that merit attention. (p. 342)

Families are by nature extremely difficult to “count,” and the study of family processes does not presume that any one way of counting should be privileged. It may also be the case that a standardized a priori determination of which family actors are relevant to a study is a risk, because interviews or observations with more relevant actors might be overlooked. It is interesting that

Hess and Handel’s (1959) choice of individual interviews, and not conjoint interviews, may not have best reflected a true whole-family methodology. Dyadic-level interviews, tapping into dyadic-level processes, would have led to a very different perspective (Reczek, 2014). In this way, a holistic approach contributes to a strength of qualitative methods: the capacity to allow for and prioritize emergence of “family” in data collection and analysis.

The culture that emerges from processes of interaction, from the meaning-making of different family members, is at the core of a science of families. And this emergent nature of family relationships over time makes it a challenge to link units of observation with units of analysis. In a continuing example that we use throughout this article, consider a hypothetical study titled “Moving In: The Process and Influence of Cohabitation on Families.” In this example, the unit of analysis is a decision to move in together. Are we talking about each individual’s decision to move in together (which may not have been made at the same moment), or are we talking about the couple’s collective decision to move in together (the moment they collectively decided and the interactional dynamics that led to that collective decision)? The process of moving in shapes a number of family relationships, each of which could be selected as a unit of observation. Potential windows onto the impact of moving in include joint purchases of furniture or accessories; interviews with partner members, separately or conjoint; shifting to family plans for insurance or technology use; or ethnographic observation of family holiday routines with siblings and parent–child dyads.

Specific methodological approaches might dictate units of observation and analysis for study. Maybe we rely on tried-and-true assumptions: To add complexity to our data collection, we reduce the N ; to focus on broad, basic patterns in the data, we increase the N (Daly, 2007). If, in the “Moving In” study, we were interested in such richness, we might expand our units of observation to include all of the household members of cohabiters before they moved in together, and we could include another generation upward to uncover grandparental attitudes and senses of whether the cohabiting partner of their grandchild is considered family. We could also include friendship groups of the cohabiting couple as units of observation to ascertain their attitudes and experiences with

cohabitation and with the couple. If we were interested in broader patterns we would limit units of observation to focus on a larger total number of cohabiting dyads, which could reveal differences and commonalities between partners who have previously cohabited with others and those who never have or couples for whom income is a major strain and possible driver of the cohabitation as compared to couples without income pressure.

DATA QUANTITY, SIZE, AND SAMPLING LOGIC(S)

When scholars need to decide what counts as scientifically based or evidence-based research, it is relatively easy to fall back on positivist assumptions to guide qualitative research methodology (St. Pierre, 2013). Steinmetz (2005) noted that logical positivism/empiricism appears to be the epistemological unconsciousness of the social sciences. For example, one assumption about sampling criteria might be that the larger the sample (of interviews and/or participants), the more scientific a study becomes (Kvale, 1996). However, it is important to note that there are both positivist and interpretivist approaches to qualitative research (Lin, 1998).

There is no “one way” of doing qualitative family research; in fact, there are vigorous debates among qualitative researchers about how to conduct research as well as how distinct qualitative methods are from quantitative methods. Regardless of their approach, researchers should be explicit about decisions to use guidelines to inform sampling. Imitation of “large *N*” assumptions from positivist and/or quantitative designs glosses over important distinctions in methodology as well as epistemology. Furthermore, when quantitative researchers review and evaluate qualitative work, imitation of methodological assumptions may result in truncation of insights gained from the strengths of qualitative tools and the dismissal of innovative research.

Types of Sampling

Choices for sampling designs for qualitative family research are plentiful, but often the distinction among schemes is unclear or underspecified. If the goal is to understand broad variation within a specific phenomenon, a large heterogeneous sample may be appropriate. If, on the other hand, the goal is to understand the

intricate dynamics of reality construction in certain family groups, a small homogeneous sample may be a better choice. Both large-sample and small-scale qualitative studies benefit from specification of a range of sampling types.

Patton (1990) and Sandelowski (1995) have asserted that all types of sampling in qualitative research can likely be arranged under a broad umbrella of purposeful sampling. Coyne (1997) presented a useful table of a full range of purposeful sampling, reflective of the literature, which uses the terms *purposeful*, *selective*, and *theoretical* interchangeably. Emmel (2013) noted three distinct traditions for sampling, including (a) Glaser and Strauss/Strauss and Corbin (theoretical), (b) Patton (purposive), and (c) Mason (analytic induction). Miles and Huberman (1994) also summarized 16 different strategies for purposive sampling. All of these schemes are based on the assumption that a qualitative researcher will select an individual or family or setting specifically because it is an “information-rich” case (Patton, 1990).

The most important distinction among sampling schemes may be if a researcher intentionally selects a design-based sample prior to data collection (selective) or leaves sample selection to be shaped primarily by the data collection process (theoretical or purposive). Selective sampling serves specific goals that distinguish it from theoretical sampling. Sandelowski (1995) suggested that when researchers make a decision prior to beginning a study about how to sample subjects, their selective sampling scheme may be based on achieving maximum demographic variation (of race, gender, class, etc.) or phenomenal variation (of a sought-after experience or process). A sampling decision during the data collection process may be based on pursuing one or more theoretical leads and seeing the need to increase the variation in the sample to best pursue those leads. Either way, to some extent, variables need to “earn” their way into sampling decisions (Morse, 1998).

The realities of family research might encourage us to consider how both selective and theoretical sampling co-occur in the same study. To some extent, both types of sampling overlap and inform each other in every qualitative research design. It is difficult, if not impossible, to purposively select a sample without at least some implicit theoretical justification. A study may begin with a particular group, such as cohabiting

couples in the hypothetical "Moving In" project. In the course of analyses, researchers may develop concepts or dimensions that make them think the sample should be expanded in some way. They could include couples they had not considered previously, such as repeat cohabiters, to emphasize some of the emerging variation that they find.

Sandelowski and colleagues (Sandelowski, Holditch-Davis, & Harris, 1992) examined couple fertility and indicated how couples were selectively sampled prior to data collection on the basis of medical and social criteria. In later stages of the study, they moved toward a theoretical sampling scheme, adding new couples to the study on the basis of emergent findings and coding. Goldberg (2009, 2010) studied lesbian, gay, and heterosexual couples' transition to adoptive parenthood. On the basis of prior research suggesting the significance of racial match or difference between parents and children, she sought to sample couples who adopted racially (i.e., a child of the same race) or transracially (i.e., a child of a different race) in each of the three groups (lesbian, gay, heterosexual). As the study progressed, geographic location also emerged as an important source of variability in participants' experiences (e.g., couples living in rural areas encountered more barriers in adopting), and therefore an effort was made to sample participants in varying geographic locations.

The "Enough" Point

Family researchers may take first steps toward scientific rigor by specifying which sampling scheme they used and why. But even with a given scheme, what about the nagging question of appropriate sample size? Informal guidelines surface in many articles, perhaps based on practical issues about convenience, resources, and time (Luborsky & Rubinstein, 1995). And both qualitative researchers and reviewers, as well as those more quantitatively oriented, can fall prey to reifying sample size based on having reached the "enough point" (Daly, 2007; Johnson, 2002). Lack of consensus leads to a default suggestion to base sampling size decisions on "the qualitative goal of describing the nature and contents of cultural, social, and personal values and experiences within specific conditions or circumstances, rather than of determining

incidence and prevalence" (Luborsky & Rubinstein, 1995, p. 106).

Despite caution, there has been no shortage of conjecture or recommendations for sample size among some qualitative methodologists. Guidelines for sample size, as we indicated, usually rely on one informant per family and one interview per informant. Creswell (1998, 2002) has urged qualitative researchers to study one cultural group in an ethnography, narrative studies for one person in narrative research, three to five cases in a case study, up to 10 people in a phenomenological study, or 15 to 30 people in a grounded theory study. Morse (1994) offered a guideline for six people in a phenomenological study, 30 to 50 interviews and/or observations for ethnographies and grounded theory research, and 100 to 200 units of observation for qualitative ethological studies. For selective samples, Luborsky and Rubinstein (1995) found 12 to 26 people in each study cells to be "just about right for most authors," and fewer than 10 cases appropriate in medical case studies. Bernard (2000) argued for 36 interviews for most ethnographic studies, and Bertaux (1981) noted that no fewer than 15 persons were acceptable for life history studies. McCracken (1988) felt that going beyond eight long, in-depth interviews was a potential waste of resources. In conducting interviews with cultural members, Spradley (1980) preferred repeated interviews with 25 to 30 participants. And Kvale (1996) observed that current studies usually settled on 15 ± 10 interviews, "due to a combination of time and resources available for the investigation, and of the law of diminishing returns" (p. 102).

It is clear that even qualitative methodologists can quickly offer guidelines for sample size without necessarily specifying how much data are being collected or what the quality of the data is. But strict proscriptive guidelines for sample size can be misleading and unintentionally interfere with goal of nuance and complexity (LaRossa, 2012a; Small, 2009). Taking the sample size of the number of observations out of context of the study design and the richness of the sample, it is difficult to make the case for sampling strategy or sample size. For example, Hess and Handel's (1959) *Family Worlds* focused on five families to describe a set of core family processes. Lewis (1959) did the same in his book, *Five Families: Mexican Case Studies in the Culture of Poverty*, which examined the social and cultural dynamics of

poverty. More recently, Lareau (2003) in her book, *Unequal Childhoods: Class, Race, and Family Life*, detailed the lives of 12 families to illuminate the complex connection between social class and children's life experiences. Both Lewis and Lareau used ethnographic techniques and paid close attention to contextual processes in communities, whereas Hess and Handel relied on interviews with individual family members. In none of these instances can we fully appreciate the contributions that a study makes simply by knowing what its sample size was.

What may be active in all of these sampling decisions is a well-known relationship: that, assuming a limit on resources, the amount of data collected from each unit of observation tends to inversely vary with the number of units of observation in the study as a whole (Morse, 1998). To address multiple levels of analysis and distinctly different research questions, some recent large-scale studies of families provide exceptions to this rule. Multiple public and private agencies funded the "Welfare, Children, and Families: A Three City Study" ethnography (<http://web.jhu.edu/threecitystudy>), in which more than 10,000 documents were collected for 256 low-income mothers and focal children over 5 years of field work. This effort may be one of the largest ethnographies ever completed (with more than 90 ethnographers in five sites), only through the investment of extensive resources.

The depth and breadth of the "Three City Study" data set challenged researchers with the question "Where do we begin?" to define a question and a sample. Burton and colleagues (Burton, Cherlin, Winn, Estacion, & Holder-Taylor, 2009), for example, used participant observation along with individual interviews over 3 years to develop a typology of suspended, compartmentalized, misplaced, and integrated trust in the intimate relationships of low-income mothers. Of mothers in 256 families, only 28 had insufficient data to describe their relationship trajectories. By examining how each of these four patterns varied by presence of abuse in relationships, Burton et al. ended up with subsamples of manageable size to allow for thorough examination. Similarly, Roy and Burton (2007) examined the 256 families to find a purposive sample of 149 mothers who recruited 299 nonresidential fathers or father figures to support the needs of their biological children through the process of *kinscription*. This sample provided broad descriptive patterns for the study. They

honed in on subsamples of mothers who provided more extensive discussion of the search for and monitoring of conventional fathering; navigation of ambiguous intimate relationships; and, in the absence of biological fathers, mothers' recruitment of father figures and paternal kin.

Another large-scale study is the "Rural Families Speak Project" (Bauer & Dolan, 2011; Braun & Anderson, 2005), a 10-state qualitative research project funded through the U.S. Department of Agriculture as a collaboration for a multiuniversity extension system. Researchers gathered qualitative data on a sample of women in poverty in rural areas of the United States. Sampling criteria for women were established (e.g., 200% of the poverty line, women with children in their homes), but difficulties arose when researchers tried to pool data because of the variation between states and the sheer amount of data gathered. Researchers found it challenging to maintain data quality standards, to balance diverse skills of interviewers across states, and to analyze data out of context. Investigators for individual analyses solved these dilemmas by choosing smaller subsamples of the larger pool of interview transcripts, based on specific reference to health or employment problems, for example. Like the "Three City Study," researchers from "Rural Families Speak" could then be attentive to state and community contexts within a broader analyses of health or employment.

Too Small or Too Large

The upshot of this wide array of sampling choices for qualitative family research should not be "anything goes." Studies can have too large of a sample, and they can have too small of a sample to be of use (Sandelowski, 1997). Bigger is not better, but neither is smaller; the N by itself is not sufficient to judge the quality of a study or the depth of analyses to be carried out. Researchers must be attuned to the trade-offs that such decisions demand and the reasons why we might tailor our sample size according to the goals and design of a specific study.

What if we have too large of a sample for a qualitative family study? If we asked a family researcher what she would do with a \$50 million grant with unlimited focus, a qualitative study of 400 families might seem appealing—it would solve the complaints about too small an

N—but would raise a number of critical problems. It can be extremely difficult to conduct detailed analyses of information-rich cases with an overstuffed sample of families. A large-*N* sample may prioritize heterogeneity, which is commendable in many respects but which makes identification, contrast, and comparison of common emergent themes quite a challenge.

A large heterogeneous sample may lead to wasted data and thus is likely to make it more difficult to reach theoretical saturation (Morse, 1998). This type of sample would lead to too much data on common experiences, sapping valuable time and resources on data replication. The large sample may make it difficult to examine data in all of their complexity, limiting ability to probe data collection, develop emergent questions, or contextualize quotes (Lareau, 2012). It would also provide too little data from families at the tails of the distribution (which may prove most insightful). Moreover, qualitative analyses are less focused on frequency and more focused on quality of experiences. According to Morse (1998), with just the right amount of data, a qualitative family researcher

must know what it is and be able to describe the phenomenon and explain all of its quirks and nuances. Because the phenomena in which we are interested do not usually follow demographic trends, we almost certainly will have too much data about some particular event or experience, and gaps and holes in our data about other events. (p. 734)

In some circumstances a large sample can be effective at allowing researchers to discover new themes and even unanticipated experiences. Goldberg (2010) relied on a large heterogeneous sample to identify multiple subsamples, including couples who adopted via foster care (see Goldberg, Moyer, Kinkler, & Richardson, 2012) and couples who adopted using open adoption (Goldberg, Kinkler, Richardson, & Downing, 2011). Roy (Marsiglio & Roy, 2012) integrated four smaller samples of 40 low-income fathers to create a data set of 160 life history interviews, which allowed him to examine thematic variation across a broad range of men in three distinct birth cohorts and distinct cultural contexts as well. In these designs, researchers were clearly developing multiple analyses over many years, and larger samples allowed them to shape and reshape their studies relative to the questions they asked.

It is less likely that researchers will be critiqued for erring on the side of choosing too large of a sample than too small of a sample. When family researchers lean toward large samples as a ruler of good research, large *N*s can become prescriptive. For example, of the 58 qualitative articles published in the *Journal of Marriage and Family* between 2000 and 2010, 25 had 45 or more participants (nine had at least 115 participants; Sharp, Zvonkovic, Humble, & Radina, 2014). If the quality of a qualitative study is assumed to be dependent on a large sample, then the standards used in quantitative designs may be relied on to assess the data (e.g., is the sample statistically representative?). Alternatively, if we reject the large size in research circumstances where it has no justifiable purpose, then we free up artificial or nonsensical demands for large sample sizes and invite pursuit of other paradigms, which promotes more innovative science overall.

What if we have too small of a sample for a qualitative family study? A small sample size may work well for selection of critical cases of families but may not yield enough data to show desired variation of an important family process or phenomenon. Researchers using grounded theory methods (GTM) will not be able to identify consistent themes, which would threaten the credibility of such a study. Insightful critiques of small sample size may suggest that these studies do not offer repetition of themes, which proves critical for some types of theoretical saturation (Lincoln & Guba, 1985). It may also make it more difficult to use constant comparative methods that are associated with theoretical saturation in GTM (Glaser & Strauss, 1967).

Qualitative family researchers are increasingly piecing together innovative samples in order to fit analyses to data. Bernardi (2011) sampled 116 participants in family networks, but those networks included 49 focal individuals and three elected network partners ($n = 67$). This provided a large sample size as well as two smaller subsamples within which to conduct related analyses. In a study of Salvadoran immigrants, Abrego (2009) interviewed 47 parents and 83 of their relatives, which also allowed for a “large” sample size of 130. Similar to her work in *Unequal Childhoods* (2003), Lareau teamed with Weininger to examine home ownership and school choice among upper middle-class families (Weininger & Lareau, 2009). They used a mixed-method approach with different samples,

including 137 interviews with 12 in-depth observations (or family ethnographies). These studies provide models to which we return when we again discuss the example of the “Moving In” study. The use of 10 cohabiters or 115 cohabiters in and of itself does not lead to quality science. In fact, it might be the innovative use of 10 cohabitating couples, nested in broader analyses of 30 families, which also include 15 interviews with parents of cohabiters and 15 interviews with siblings of cohabiters, that can offer new insight into how cohabitation decisions are shaped not only by individual preference but also by parental expectations, competition between siblings, or marital socialization processes. These multiple sampling decisions are then linked to saturation and to theory about family processes.

Right-sizing a qualitative family study may start on the surface with a decision about sample size, but underlying this decision is an assumption that runs counter to received standards of “good” social science. Qualitative analysis brings insight into context, meaning, and processes only through analyses that are “generically about maximizing understanding of [a case] in all of its diversity” (Sandelowski, 1995, p. 180). It encourages us to consider the values and trade-offs in a commonly used variable-oriented approach as well as a more seldom-used case-oriented approach (Ragin & Becker, 1992).

Case- and Variable-Based Logics

Although it is not a necessary condition, at its most basic level, the bulk of social science relies on a degree of comparison (Abend, Petre, & Sauder, 2013). The sampling decisions we make in qualitative family research also reflect comparisons. Like qualitative researchers, family demographers, psychologists, or economists are thirsty for in-depth studies of families (Small, 2009). However, they remain unsure about the contributions of ethnographic or interview-based studies with small *Ns*. As we have noted, some qualitative research takes a positivist approach, whereas other research takes an interpretivist approach (Lin, 1998). Yet even within these diverse approaches it is more common for qualitative research to make only limited claims to statistical representativeness.

We also almost always look at dimensions of family experience and their connections to one another and almost always address variables

and relationships (Abend et al., 2013; LaRossa, 2012a, 2012b; Ragin, 1987). Glaser and Strauss (1967) addressed maximizing and minimizing differences in theoretical sampling by attention to dimensions (or variables). What is at issue is expanding our repertoire of approaches and tailoring them to the strengths of qualitative methods.

A variable-based logic to family inquiry is commonly used in quantitative methodology. This logic in quantitative research begins with a predetermined sample size (usually individuals) in which a sample is drawn from a population in such a way that each individual has an equal chance of selection (Yin, 2002). Furthermore, the same questions and equivalent data are collected from each individual. If the logic works properly, analytical findings based on these methodological decisions—sample selection and size—should generalize to the population and be statistically representative. Qualitative family research can draw heavily on a variable-based approach as well. The logic of GTM, LaRossa (2005) pointed out, is variable based in that the methods require the comparison of cases in order to both develop dimensions (or variables) and hypothesize about the relationship among those dimensions.

A case-based logic is different and perhaps uniquely challenging for social scientists because the term *case* is used in so many different ways (Ragin & Becker, 1992). Each case offers slightly different data due to slightly different sets of questions or measures; one case leads to findings that inform the selection of the subsequent case. Case study logic may differ from a variable-based approach in that it privileges sampling based on what may be idiosyncratic characteristics. Some scholars argue that a case-oriented approach is best for very small samples. In his book *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*, Ragin (1987) noted that the approach is tailored for

when the number of relevant cases is small. The comparison of two to four positive cases with the same number of negative cases is manageable. As the number of cases and the number of relevant causal conditions increase, it becomes more difficult to use a case-based approach. (p. 49)

Ragin and Zaret (1983) also argued that case-oriented and variable-oriented approaches

are complementary in part because cases should be “examined as wholes—as combinations of characteristics” (p. 16).

A case-based approach is rare in qualitative family research in large part because, as we have noted, there are few case studies of whole families. One way to expand the repertoire of qualitative family research is through consideration of how case-based logic can offer unique and valuable insights. In sociology, Small (2009) argued, the use of case-based approaches in urban ethnographic research allows qualitative methods to gain legitimacy and to engage in a broader multimethod dialogue with survey-based demographic studies. Similarly, in family science, a case-based study would examine cases—possibly couples, families, extended kin networks, or neighborhoods and communities—as “wholes.” This approach would focus less on comparison of aspects of each case and more on how different family or network configurations can yield insights into how these entities operate as something akin to a system. Appreciation and encouragement of this approach would open the door for more network analyses (Hansen, 2004), family ethnography (Lareau, 2003), and life course historical case studies (Handel, 2003) to diversify and expand the reach of qualitative family research.

Case- and variable-based approaches can also be used in combination. In a study of 20 couples in transition to first- or second-time parenthood, LaRossa and LaRossa (1981) used a variable-based approach when they hypothesized about the relationship between the helplessness of infants and the continuity of care that parents provided and about consequences of different levels of care (primary, secondary, tertiary). LaRossa and LaRossa also presented case studies of four of the couples to delineate the process by which, over the course of 9 months after the baby’s arrival, the culture and conduct of families changed, with a tendency for couples to shift toward a more traditional division of labor, especially with respect to infant care.

In a recent family study related to work travel (Swenson & Zvonkovic, in press), each family member was, in essence, a case with a distinctive vantage point, and his or her case was analytically elaborated while also using elements of his or her family situation as samples for variable-oriented analyses and contrasts (e.g., each family member’s perceptions of work travel as a case; families in which Dad travels, Mom

travels, both travel, as variable-oriented samples of families). Marsiglio and Roy (2012) used interviews from several qualitative studies of fathers to consider possible family policy initiatives, but they also chose one or two case studies of individual fathers to reflect a more comprehensive and complex understanding of dimensions of fathering, such as personal meaning, kin work, or nurturant bonds with children.

In our hypothetical “Moving In” study it is clear that each approach would yield different but complementary insights. A study of 30 couples—or even 60 individuals within those couples—might tend toward a variable-based approach, with adequate data to examine variation in decisions that led to cohabitation or in one partner’s reluctance to share expenses and his or her partner’s response to such reluctance. It would be difficult to capture family dynamics beyond those of individual- or couple-level experiences. A study of five families, each of which undergoes the buildup and aftermath of a cohabitation decision, might tend toward a case-based approach. Considered more holistically, each family would be framed as a distinct network and set of dynamics, perhaps representing a type or trajectory. It would be difficult to find enough variation in decision that led to cohabitation to reach saturation (as in the sample of 30 couples), but this study would offer deep description of interaction of extended family members, sibling reaction, and conflict between each partners’ parents. A small number of families might also make participant observation more feasible. All of these dimensions may be critical to a family-based understanding of cohabitation but would be challenging to offer in a variable-based study.

If qualitative family researchers develop a rigorous logic of case selection to justify their methods and do not rely as heavily on the language and fundamental concepts of statistical inference, sample size might not be as problematic for reviewers, consumers, and the broader intellectual community of family science. As a consequence, family science journals might publish articles with Methods sections that denote, for example, “a set of 12 (case) families” as well as “a sample with an *N* of 45 parents (in different families).” Also, despite recriminations and disjunctions across both approaches (Ragin, 1999), family scientists might understand the value of each approach. In both, the process of methods—of how data collection

uniquely informs data analysis, which then informs collection again—is given priority over an often awkward and misguided fit of statistical inference guideline for qualitative methods.

SATURATION IN QUALITATIVE FAMILY RESEARCH

In fields like health science research, saturation has become the gold standard by which sample sizes are determined and by which theoretical advances are demonstrated (Guest, Bunce, & Johnson, 2006). Saturation may hold promise in family science, but as a methodological process it is often poorly described. If saturation is a “key to excellent qualitative work” (Morse, 1995), we must also guard against using saturation as the only goal of qualitative work (Bowen, 2008; O’Reilly & Parker, 2013).

Data saturation, or *informational redundancy*, occurs “when researchers sense they have seen or heard something so repeatedly that they can anticipate it” (Sandelowski, 2008, p. 875). *Theoretical saturation* occurs when researchers are satisfied “that the properties and dimensions of the concepts and conceptual relationships selected to render the target event are fully described and that they have captured its complexity and variation” (Sandelowski, 2008, p. 875). These different types of saturation are different moments in qualitative inquiry. Data saturation is relevant during a design phase and as design shifts to accommodate emergent findings. Theoretical saturation is especially important for the purposes of analyses. Of course, many qualitative studies are based on an iterative process of data collection and analysis, which suggests that data and theoretical saturation are closely related. In this way, richness of data leads to clear theoretical development—and ultimately to integrity of qualitative methods.

Data Saturation

At which point do we have the requisite amount of data to begin to thoroughly describe a family phenomenon or to distinguish repetition of themes? In family research, again, decisions about units of analyses become critical. Are 20 individuals likely to lead to discussion of common themes in the same way that 10 couples might? If families can be sampled in many different ways, repeated themes or codes will vary with the focus of a study and the research design to carry it out.

Often it is just assumed that studies with large samples will provide more than enough data to address a research question. However, if a researcher develops a question that would be better answered with a small sample, perhaps a longitudinal design with fewer cases is jettisoned for a study with a larger number of cases. If a survey mentality persists as the study is developed, the researcher designs an interview schedule of an excessive amount of questions, which cannot be realistically covered in, say, a 1-hour or 90-minute interview and which is likely to result in very “thin” data. Despite the large sample, the researcher does not really have more than enough data to answer the questions of that study.

If a bias towards large-*N* studies exists, it is possible that family scientists may give more credence to such research, which raises a host of questions and gives rise to some unintended consequences. Privileging (intentionally or not) large samples may discourage studies with smaller sample sizes that are better tailored to fit research considerations. For example, a large amount of repetitive data may impede analyses, pulling researchers away from in-depth consideration of specific cases and toward counts as a way to validate findings (Lareau, 2012). Family research may be particularly positioned to offer rich details of daily experience, but such data may be ignored in pursuit of more interviews to bump up sample size.

Saturation may also reflect the quality or depth of data, which can appear to contradict the advantage of having larger sample sizes. Sharp (Sharp & Ispa, 2006; Sharp, Ispa, & SoRelle, 2014) recollected attending a conference presentation with plenary speakers who discussed findings from a large study of mothers living in poverty. With more than 200 women their sample, the speakers were confident about patterns they had identified with descriptive data from the large number of participants. In contrast, Sharp and her colleagues spent several years in the field, working on research with nine single mothers living in poverty. Not only did their findings map onto primarily descriptive findings from the larger study but also many of the nuanced understandings of low-income mothering/partnering appeared to be absent from the study with 200+ participants. A simple statement that saturation has been “reached” should not substitute for hard work of demonstrating

how saturation was achieved—and often the proof is in the pudding (i.e., the findings).

The trade-offs are complex. For example, in the hypothetical “Moving In” study researchers may opt for a very large sample of 100 couples with hopes that they will capture a great deal of information of this process across many different family units. In so doing, they would give priority to dimensions that they can gather in common across these families, such as family structure. They might develop a typology of communication strategies for couples. However, it may be more difficult to find data that capture the distinctly constructed meanings behind cohabitation, the daily routines that differ from couple to couple, or cultural and local contexts that shape expectations for cohabitation.

In an interview-based study of fishing families, Zvonkovic, McGraw, and Manoogian-O’Dell (2000) found that only two focus groups, especially if those had been the groups with the smallest number of participants, would not have saturated some of the theoretical concepts. Although concepts such as wives’ independence from husbands were prominent in each group, concepts such as links to community schedules and the ways workplace policies of the women affected families would not have been clear. The emergence of these themes occurred in part because the ages of the women and the presence and ages of their children were not part of the purposive sampling of the study participants at that point. Furthermore, group dynamics were different in each focus group; in particular, the previous relationships of women in the focus groups appeared to affect dynamics in such a way that they allowed for more personal discussion to occur earlier on in the focus group session.

Theoretical Saturation

First defined as a fundamentally inductive process (Glaser & Strauss, 1967), grounded theory methodology has been revised to reflect more of a dialectical process of induction and deduction (Charmaz, 2006; Emerson, Fretz, & Shaw, 1997; LaRossa, 2005). The distinction between data and codes and between theoretical and selective sampling ensures that a researcher develops codes as symbols in the course of reading and analyzing data. The codes then guide further data collection. As Glaser and Strauss (1967) stated,

“The analyst jointly collects, codes, and analyzes his data and decides which data to collect next and where to find them in order to develop his theory as it emerges” (p. 36). Glaser and Strauss then identified saturation in grounded theory research as the moment at which

no additional data are being found, whereby the researcher can develop properties of the category. As he sees similar instances over and over again, the researcher becomes empirically confident that a category is saturated. . . . [W]hen one category is saturated, nothing remains but to go on to new groups for data on other categories, an attempt to saturate these categories also. (p. 65)

Daly (2007) restated this early definition by asking researchers the following question:

After repeated conversations with participants, have [you] reached the point where [you] understand their experience as fully as possible? Have [you] sampled a sufficient range of individuals to allow [you] to understand both the nature of their patterned experience and the variation that exists in their experience? (p. 178)

How might a theory about families actually emerge during a process of saturation? In his examination of the use of GTM in family science, LaRossa (2005) made the following observation:

When a researcher got to a point where the addition of another indicator to those already grouped under a concept did not appear to generate significantly new insights about that concept, then, in GTM terms, the concept is *theoretically saturated*. A theoretically saturated concept essentially is a well-grounded concept. (p. 841)

He stressed not just the embeddedness of a concept within a network of related linkages and how dense such a network of linkages might be, but how each new linkage transforms that concept.

Instead of relying on the number of times a concept emerges to convey its importance, theoretical saturation rests on close examination of all of the contexts and related themes that are somehow related to it. It is the exhaustion of unique dimensions that flesh out, clarify, transform, or dimensionalize that leads to a fully saturated concept. In a study of 41 young adult men disconnected from school and work, Roy, Messina, Smith, and Waters (2014) examined

the processes of adultification of these men when they were boys or adolescents. They discovered that half of the young men believed that they grew up as “men of the house” in the absence of their own fathers or other men in the household. The researchers then explored these interviews, to link them to developmentally off-time provision of authority and independence that could both hinder and bolster a sense of successful adulthood in later life.

Saturation may emerge in stages and in often unexpected ways. Lee and Zvonkovic (2014) used conjoint interviews with couples who voluntarily chose not to have children. Diverging from past research based on interviews with couples, they closely linked the empirical unit of analyses (couple) with the theoretical unit of analysis (the decision-making process). The researchers continued to recruit more couples who varied in timing of the decision and the pathways to decisions, even if one of them had entered into the relationship unsure about having children. In another example, Kinkler and Goldberg (2014) aimed to study lesbian, gay, and heterosexual individuals who had become single parents by choice through adoption. They agreed that saturation seemed to have been reached by the time they reached 28 participants. However, unexpectedly, they were contacted by several additional gay male single parents by choice, and their interviews generated some entirely new themes that caused them to revisit their earlier conclusion that saturation had in fact been reached.

Ryan and Bernard (2003) indicated that saturation also depends on investigator experience and fatigue and the number of analysts reviewing data. In other words, theoretical saturation is in many ways dependent on the skills of a researcher or team of researchers. As a qualitative family researcher begins to develop theoretical constructs in his or her own work, new analyses that contribute to this framework may require smaller sample sizes or lend themselves to case study designs. Taking into account informational redundancy, the process of theoretical saturation, and the idiosyncrasies of their research team, Guest et al. (2006) critically explored 60 in-depth interviews with women in West Africa. They found that saturation occurred after only 12 of the interviews. However, as they pointed out, they may have found common themes because of their reliance on semistructured interviews, in contrast to open-ended interviews, which might

require more cases to reach saturation. The more homogeneous a sample, the closer saturation may seem to be, given that themes are more likely to be repeated in very specific cases with very specific questions.

Saturation and Validity

Saturation must finally be placed within a broader consideration of validity in qualitative research. Again, there is a wide range of perspectives on validity in qualitative research. Some researchers critique and even reject the goal of achieving validity (Wolcott, 1990). Lincoln and Guba (1985) transformed the concept of validity into a question of securing the credibility of qualitative data. Consideration of credibility mirrors some of Lin's (1998) discussion of how interpretivist approaches seek accuracy, faithfulness, and richness of data.

In many ways, the process of saturation as we have presented it relies on many of the same prompts that lead to a sense of validity (or credibility). In particular, our emphasis on sampling decisions that lead to richness and quality of data is very similar to using strategies that would promote data saturation and provide for credibility in particular. We would expect that other strategies that promote credibility, such as prolonged exposure in the field, repeated observations, member checks, or peer debriefing, are also related to richness and saturation. Also, saturation is about the relation of data saturation and theoretical saturation, which reflects how a deep and precise relation between data and theory drives any claims of credibility.

It may seem surprising that so few cases can lead to theoretically significant findings that have something to say to other, larger-*N* studies. We can understand cases outside of a specific qualitative study when we construct a theoretical insight or framework that is portable, moving from families in one specific context out toward other families in a range of other contexts. We do not argue that it is impossible to generalize with a qualitative approach, but we encourage scholars to open up consideration of generalizability by close attention to sampling decisions. Qualitative family analyses can make limited claims for statistical representativeness but unlimited claims for their general theoretical significance (see Richardson, 1988; also Small, 2009).

The risk for family science is uncritical emphasis of rigid methodological guidelines at the expense of rigorous examination of findings. Sandelowski (1997) acknowledged that qualitative studies cannot be easily and quickly summarized in part because of the unique theoretical offerings of each set of findings:

Knowledge accumulation is less about knowing more than about having more perspectives from which to know. . . . [T]he knowledge development proceeds not by addition, but rather by reformulation of idea. The goal of qualitative inquiry is not the mere accumulation of information, but rather the transformation of understanding. (p. 128)

A PROPOSAL FOR A STATEMENT ON QUALITATIVE INTEGRITY

What remains is the need for a concise, alternative standard by which to evaluate sampling in qualitative family research. When researchers write about their sampling decisions or their coding processes, they fall into a dense soup of qualitative terminology that spans disciplines. It seems easier, especially when we communicate with colleagues who use statistical analyses, to borrow the trappings of preference for large samples that would make everyone content because they at least seem more generalizable or representative. But there is no power analysis for qualitative family research (Luborsky & Rubinstein, 1995; Onwuegbuzie & Leech, 2007).

In place of that, we argue that researchers should integrate a discussion of *qualitative integrity* in the Method sections of their articles. Specifically, qualitative integrity would address a comprehensive list of three issues raised here: (a) the closeness of fit between units of observation and analysis, (b) the achievement of sample richness related to quality and sample size determination, and (c) the demonstration of data saturation and the thorough development of conceptual properties and dimensions that reflect theoretical saturation. A statement on qualitative integrity would follow alongside subsections on data collection techniques, sample description, data analyses frameworks, and trustworthiness. It would be tied closely to broad research questions, theoretical assumptions, and epistemological and ontological stances of the researcher. In the broadest sense, then, integrity would be reflected through efforts to

show methodological congruence (Burns, 1989; Richards & Morse, 2012).

Qualitative integrity is based in part on theoretical grounding (Luborsky & Rubinstein, 1995). For many analyses of family life, researchers would consider how their sampling decisions fit within one version or another of a grounded theory framework. Beyond this, however, researchers would present the possible range of studies that used similar designs or possibly informed their choice of sampling and saturation process. In these studies, how were units of observation nested within each other, how did units of observation relate to units of analysis, and how was theoretical saturation achieved? Perhaps there are other studies that offer a contrast, those in which samples were too large or too small to adequately address the research questions of the study. A careful examination of theoretical grounding will also provide an audit trail of decisions and trade-offs that will demystify qualitative family work.

The process of writing about qualitative integrity will encourage researchers to critically examine their own assumptions and decisions about sampling as well. If qualitative family research can provide insight into how families construct meaning, we need to be clear about how our own decisions about units of observation and sample richness may singlehandedly transform such meaning. For example, family researchers may use their own concepts to derive a sample of "no-parent families," and the cases that they study are rooted in the experiences of grandparents, *compadres*, or "other mothers" as they raise complex networks of children. Although the term *no-parent families* may reflect terms used in nationally representative surveys households, it clashes with the very essence of a "case" in such a study. Likewise, developing a sample of families based on another measure of structure, such as father absence, may lead researchers to overlook the dynamics of nonresidential father contact or involvement. When searching for theoretical saturation regarding themes of father absence, they may be encouraged to forgo the insights and descriptions that families have of their own lives in order to validate methodological decisions on sampling.

A commitment to articulate qualitative integrity regarding sampling is a way for researchers to continue to engage and debate these important issues. Instead of offering a

generic take on qualitative methods, this rationale will situate researchers in a rich body of previous empirical literature and theory, grant access to a long tradition of qualitative methodology, and make explicit how seemingly simple decisions about sampling are at the core of innovative research that contributes to our understanding of families.

CONCLUSION

Trotter (2012) wrote that a lingering challenge for questions about sampling is the linguistic/definitional heritage of qualitative research. Each discipline that has contributed to qualitative research—for example, sociology, anthropology, education, public health, and nursing—uses different terminology, for a range of designs. In a field such as family science, which draws advantages from multidisciplinary and methodological pluralism, the lack of a consensus on qualitative terminology is a formidable challenge. What often results are rather generic qualitative studies, which muddle or slur theoretical perspectives or methodological assumptions regarding sampling (Caelli, Ray, & Mill, 2003; Coyne, 1997). Such studies lead to claims of deficient inquiry, in which small-*N* studies are said to yield preliminary findings that add color to the “real” social science of large-*N* studies. Alternatively, they lead to deficient practice, in which qualitative research appears to be easy or natural, requiring little training and skills to practice (Sandelowski, 1997).

What is needed in qualitative family research, we argue, is a deep consideration of rigor. However, rigor does not need to arrive in the form of rules and procedures that preestablish a target sample size for a qualitative study. Rigor for authors, reviewers, and consumers of qualitative research might be best conveyed as fidelity to the spirit of qualitative work (Sandelowski, 1993). We have presented the construct of sample richness from which qualitative integrity can be derived as a heuristic that partially addresses these concerns and encourages more dialogue about the critical and complicated issue of sampling in qualitative family research.

NOTE

The authors thank Stacy Gorman for her research assistance with this article.

REFERENCES

- Abend, G., Petre, C., & Sauder, M. (2013). Styles of causal thought: An empirical investigation. *American Journal of Sociology*, *119*, 602–654.
- Abrego, L. (2009). Economic well-being in Salvadoran transnational families: How gender affects remittance practices. *Journal of Marriage and Family*, *71*, 1070–1085. doi:10.1111/j.1741-3737.2009.00653.x
- Bauer, J., & Dolan, E. (Eds.). (2011). *Rural families and work*. Springer: New York.
- Bernard, H. R. (2000). *Social research methods*. Thousand Oaks, CA: Sage.
- Bernardi, L. (2011). A mixed-methods social networks study design for research on transnational families. *Journal of Marriage and Family*, *73*, 788–803. doi:10.1111/j.1741-3737.2011.00845.x
- Bertaux, D. (1981). From the life history approach to the transformation of sociological practice. In D. Bertaux (Ed.), *Biography and society: The life history approach in the social sciences* (pp. 29–45). Beverly Hills, CA: Sage.
- Bledsoe, C., & Sow, P. (2011). Back to Africa: Second chances for the children of West African immigrants. *Journal of Marriage and Family*, *73*, 747–762. doi:10.1111/j.1741-3737.2011.00843.x
- Bowen, G. (2008). Naturalistic inquiry and the saturation concept: A research note. *Qualitative Research*, *8*, 137–152.
- Braun, B., & Anderson, E. (2005). A multisite, mixed methods study of rural low-income families. In V. Bengtson, A. Acock, K. Allen, P. Dilworth-Anderson, & D. Klein (Eds.), *Sourcebook of family theory and research* (pp. 484–487). Thousand Oaks, CA: Sage.
- Burgess, E. (1926). The family as a unity of interacting personalities. *Family*, *7*, 3–9.
- Burns, N. (1989). Standards for qualitative research. *Nursing Science Quarterly*, *2*, 44–52.
- Burton, L. M., Cherlin, A., Winn, D., Estacion, A., & Holder-Taylor, C. (2009). The role of trust in low-income mothers' intimate unions. *Journal of Marriage and Family*, *71*, 1107–1124. doi:10.1111/j.1741-3737.2009.00658.x
- Caelli, K., Ray, L., & Mill, J. (2003). “Clear as mud”: Toward greater clarity in generic qualitative research. *International Journal of Qualitative Methods*, *2*(2). Retrieved from <http://ejournals.library.ualberta.ca/index.php/IJQM/article/view/4521/3651>
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.
- Coyne, I. (1997). Sampling in qualitative research: Purposeful and theoretical sampling, merging or clear boundaries? *Journal of Advanced Nursing*, *26*, 623–630.

- Creswell, J. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.
- Creswell, J. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson Education.
- Daly, K. (2003). Family theory versus theories families live by. *Journal of Marriage and Family*, 65, 771–784. doi:10.1111/j.1741-3737.2003.00771.x
- Daly, K. (2007). *Qualitative methods for family studies and human development*. Thousand Oaks, CA: Sage.
- Dreby, J. (2010). *Divided by borders: Mexican migrants and their children*. Berkeley: University of California Press.
- Elder, G. H., Jr. (1999). *Children of the Great Depression: Social change in life experience* (25th anniversary ed.). Boulder, CO: Westview Press.
- Emerson, R., Fretz, R., & Shaw, L. (1997). *Writing ethnographic fieldnotes*. Chicago: University of Chicago Press.
- Emmel, N. (2013). *Sampling and choosing cases in qualitative research: A realist approach*. Thousand Oaks, CA: Sage.
- Ganong, L., & Coleman, M. (2013). Qualitative research on family relationships. *Journal of Social and Personal Relationships*, 31, 451–459.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine.
- Goldberg, A. E. (2009). Lesbian and heterosexual preadoptive couples' openness to transracial adoption. *American Journal of Orthopsychiatry*, 79, 103–117.
- Goldberg, A. E. (2010). The transition to adoptive parenthood. In T. W. Miller (Ed.), *Handbook of stressful transitions across the life span* (pp. 165–184). New York: Springer.
- Goldberg, A. E., Kinkler, L. A., Richardson, H. B., & Downing, J. B. (2011). Perceptions and experiences of open adoption among lesbian, gay, and heterosexual couples: A qualitative study. *Journal of Marriage and Family*, 73, 502–518. doi:10.1111/j.1741-3737.2010.00821.x
- Goldberg, A. E., Moyer, A. M., & Kinkler, L. A., & Richardson, H. B. (2012). "When you're sitting on the fence, hope's the hardest part": Experiences and challenges of lesbian, gay and heterosexual couples adopting through the child welfare system. *Adoption Quarterly*, 15, 1–28.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18, 59–82.
- Hagestad, G., & Dannefer, D. (2001). Concepts and theories of aging: Beyond microfication in social science approaches. In R. Binstock & L. George (Eds.), *Handbook of aging and the social sciences* (5th ed., pp. 217–237). New York: Academic Press.
- Handel, G. (1996). Family worlds and qualitative family research: Emergence and prospects of whole-family methodology. *Marriage & Family Review*, 24, 335–348.
- Handel, G. (2003). *Making a life in Yorkville: Experience and meaning in the life course narrative of an urban working-class man*. New York: Aldine.
- Hansen, K. (2004). *Not-so-nuclear families: Class, gender, and networks of care*. New Brunswick, NJ: Rutgers University Press.
- Hess, R., & Handel, G. (1959). *Family worlds: A psychosocial approach to family life*. Chicago: University of Chicago Press.
- Johnson, J. M. (2002). In-depth interviewing. In J. F. Gubrium & J. A. Holstein (Eds.), *Handbook of interview research: Context and method* (pp. 103–119). Thousand Oaks: Sage.
- Kinkler, L. A., & Goldberg, A. E. (2014). *Lesbian, gay, and heterosexual single adoptive parents by choice: A qualitative study*. Manuscript in preparation.
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage.
- Lareau, A. (2003). *Unequal childhoods: Class, race, and family life*. Berkeley: University of California Press.
- Lareau, A. (2012). Using the terms hypothesis and variable for qualitative work: A critical reflection. *Journal of Marriage and Family*, 74, 671–677. doi:10.1111/j.1741-3737.2012.00980.x
- LaRossa, R. (1977). *Conflict and power in marriage: Expecting the first child*. Beverly Hills, CA: Sage.
- LaRossa, R. (2005). Grounded theory methods and qualitative family research. *Journal of Marriage and Family*, 67, 837–857. doi:10.1111/j.1741-3737.2005.00179.x
- LaRossa, R. (2011). *Of war and men: World War II in the lives of fathers and their families*. Chicago: University of Chicago Press.
- LaRossa, R. (2012a). Thinking about the nature and scope of qualitative research. *Journal of Marriage and Family*, 74, 678–687. doi:10.1111/j.1741-3737.2012.00979.x
- LaRossa, R. (2012b). Writing and reviewing manuscripts in the multidimensional world of qualitative research. *Journal of Marriage and Family*, 74, 643–659. doi:10.1111/j.1741-3737.2012.00978.x
- LaRossa, R., Goldberg, A., Roy, K., Sharp, E., & Zvonkovic. (2014). *Qualitative Research Commission Report*. Retrieved from <https://www.ncfr.org/jmf/qualitative-research-and-journal-marriage-and-family>

- LaRossa, R., & LaRossa, M. M. (1981). *Transition to parenthood: How infants change families*. Beverly Hills, CA: Sage.
- LaRossa, R., & Reitzes, D. C. (1993). Continuity and change in middle class fatherhood, 1925–1939: The culture–conduct connection. *Journal of Marriage and the Family*, *55*, 455–468.
- Lee, K., & Zvonkovic, A. (2014). Journeys to remain childless: A grounded theory examination of decision-making processes among voluntarily childless couples. *Journal of Social and Personal Relationships*, *31*, 535–553.
- Lewis, O. (1975). *Five families: Mexican case studies in the culture of poverty*. New York: Basic Books.
- Lin, A. (1998). Bridging positivist and interpretivist approaches to qualitative methods. *Policy Studies Journal*, *26*, 162–180.
- Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Luborsky, M., & Rubinstein, R. (1995). Sampling in qualitative research: Rationale, issues, and methods. *Research on Aging*, *17*, 89–113.
- Marcus, S. (1974). *Engels, Manchester, and the working class*. New York: Random House.
- Marsiglio, W., & Roy, K. (2012). *Nurturing dads: Social initiatives for contemporary fatherhood*. New York: Russell Sage Foundation.
- Matthews, S. M. (2012). Enhancing the qualitative-research culture in family studies. *Journal of Marriage and Family*, *74*, 666–670. doi:10.1111/j.1741-3737.2012.00982.x
- McCall, G. J., & Simmons, J. L. (1991). Levels of analysis: The individual, the dyad, and the larger social group. In B. Montgomery & S. Duck (Eds.), *Studying interpersonal interaction* (pp. 56–81). New York: Guilford Press.
- McCracken, G. (1988). *The long interview*. Newbury Park, CA: Sage.
- Meyerowitz, J. (1994). *Not June Cleaver: Women and gender in postwar America, 1945–1960*. Philadelphia: Temple University Press.
- Miles, M., & Huberman, A. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Morse, J. (1994). Designing funded qualitative research. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 220–235). Thousand Oaks, CA: Sage.
- Morse, J. (1995). The significance of saturation. *Qualitative Health Research*, *5*, 147–149.
- Morse, J. (1998). What’s wrong with random selection? *Qualitative Health Research*, *8*, 733–735.
- Onwuegbuzie, A., & Leech, N. (2007). A call for qualitative power analyses. *Quality & Quantity*, *41*, 105–121.
- O’Reilly, M., & Parker, N. (2013). “Unsatisfactory saturation”: A critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research*, *13*, 190–197.
- Patton, M. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.
- Ragin, C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
- Ragin, C. (1999). The distinctiveness of case-oriented research. *Health Services Research*, *34*, 1137–1151.
- Ragin, C., & Becker, H. (1992). *What is a case? Exploring the foundations of social inquiry*. New York: Cambridge University Press.
- Ragin, C., & Zaret, D. (1983). Theory and method in comparative research: Two strategies. *Social Forces*, *61*, 731–751.
- Reczek, C. (2014). Conducting a multi family member interview study. *Family Process*, *53*, 318–335.
- Richards, M., & Morse, J. (2012). *README FIRST for a user’s guide to qualitative methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Richardson, L. (1988). Secrecy and status: The social construction of forbidden relationships. *American Sociological Review*, *53*, 209–219.
- Roy, K. (2012). In search of a culture: Navigation of dimensions of qualitative research. *Journal of Marriage and Family*, *74*, 660–665. doi:10.1111/j.1741-3737.2012.00981.x
- Roy, K., & Burton, L. (2007). Mothering through recruitment: Kinscription of non-residential fathers and father figures in low-income families. *Family Relations*, *56*, 24–39.
- Roy, K., Messina, L., Smith, J., & Waters, D. (2014). Growing up as “man of the house”: Adulthood and transition into adulthood for young men in economically disadvantaged families. *New Directions for Child and Adolescent Development*, *143*, 55–72.
- Ruano, B., Bruce, J., & McDermott, M. (1969). Pilgrim’s Progress II: Recent trends and prospects in family research. *Journal of Marriage and the Family*, *31*, 688–698.
- Ryan, G., & Bernard, H. (2003). Techniques to identify themes. *Field Methods*, *1*, 85–109.
- St. Pierre, E. A. (2013). The appearance of data. *Cultural Studies ↔ Critical Methodologies*, *13*, 223–227.
- Sandelowski, M. (1993). Rigor or rigor mortis: The problem of rigor in qualitative research revisited. *Advances in Nursing Science*, *2*, 1–8.
- Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing & Health*, *18*, 179–183.
- Sandelowski, M. (1997). “To be of use”: Enhancing the utility of qualitative research. *Nursing Outlook*, *45*, 125–132.
- Sandelowski, M. (2008). Theoretical saturation. In L. Given (Ed.), *The Sage encyclopedia of qualitative research methods* (pp. 875–876). Thousand Oaks, CA: Sage.

- Sandelowski, M., Holditch-Davis, D., & Harris, B. (1992). Using qualitative and quantitative methods: The transition to parenthood of infertile couples. In J. Gilgun, K. Daly, & G. Handel (Eds.), *Qualitative methods in family research* (pp. 301–323). Newbury Park, CA: Sage.
- Sedgwick, P. (2014). Unit of observation versus unit of analysis. *British Medical Journal*, *348*, g3840. doi:10.1136/bmj.g3840
- Sharp, E. A., & Ispa, J. M. (2006). Matters of the heart: Romantic relationships. In J. M. Ispa, K. R. Thornburg, & M. A. Fine (Eds.), *Keepin' on: The everyday struggles of families in poverty* (pp. 133–156). Baltimore: Brookes.
- Sharp, E. A., Ispa, J., & SoRelle, D. (2014). *Love and trouble: Paradoxical romantic ideas and experiences of young, single, African American mothers living in poverty*. Manuscript in preparation.
- Sharp, E., Zvonkovic, A., Humble, A. & Radina, M. E. (2014). Cultivating the family studies terrain: A synthesis of qualitative conceptual articles. *Journal of Family Theory & Review*, *6*, 139–168.
- Small, M. (2009). How many cases do I need? On science and the logic of case selection in field-based research. *Ethnography*, *10*, 5–38.
- Spradley, J. (1980). *Participant observation*. New York: Holt, Rinehart & Winston.
- Sprey, J. (2013). Extending the range of questioning in family science through ideas from the exact sciences. *Journal of Family Theory & Review*, *5*, 4–19.
- Steinmetz, G. (2005). The epistemological unconscious of U.S. sociology and the transition to post-Fordism: The case of historical sociology. In J. Adams, E. S. Clemens, & A. S. Orloff (Eds.), *Remaking modernity: Politics, history, sociology* (pp. 109–157). Durham, NC: Duke University Press.
- Swenson, A. R., & Zvonkovic, A. M. (in press). Navigating mothering: Frequent work travel as a vehicle for navigating independence in families. *Sex Roles*.
- Trotter, R. (2012). Qualitative research sample design and sample size: Resolving and unresolved issues and inferential imperatives. *Preventative Medicine*, *55*, 398–400.
- Udehn, L. (2002). The changing face of methodological individualism. *Annual Review of Sociology*, *28*, 479–507.
- Weininger, E. B., & Lareau, A. (2009). Paradoxical pathways: An ethnographic extension of Kohn's findings on class and childrearing. *Journal of Marriage and Family*, *71*, 680–695.
- Wolcott, H. F. (1990). On seeking—and rejecting—validity in qualitative research. In E. W. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate* (pp. 121–152). New York: Teachers College Press.
- Yin, R. (2002). *Case study research*. Thousand Oaks, CA: Sage.
- Zvonkovic, A. M., McGraw, L. A., & Manoogian-O'Dell, M. (2000). A multi-method research project on commercial fishing families: Multiple windows on resilient women and families. In S. Hanna & M. Hall-Arbor (Eds.), *Resilience in fishing communities, businesses, and families* (pp. 83–96). Corvallis: Oregon State University Press.