HHS Public Access

Author manuscript

Monogr Soc Res Child Dev. Author manuscript; available in PMC 2017 September 20.

Published in final edited form as:

Monogr Soc Res Child Dev. 2017 June ; 82(2): 13–30. doi:10.1111/mono.12296.

More than Just Convenient: The Scientific Merits of Homogeneous Convenience Samples

Justin Jager,
Arizona State University

Diane L. Putnick, and
Eunice Kennedy Shriver National Institute of Child Health and Human Development

Marc H. Bornstein
Eunice Kennedy Shriver National Institute of Child Health and Human Development

Abstract

Despite their disadvantaged generalizability relative to probability samples, non-probability convenience samples are the standard within developmental science, and likely will remain so because probability samples are cost-prohibitive and most available probability samples are ill-suited to examine developmental questions. In lieu of focusing on how to eliminate or sharply reduce reliance on convenience samples within developmental science, here we propose how to augment their advantages when it comes to understanding population effects as well as subpopulation differences. Although all convenience samples have less clear generalizability than probability samples, we argue that homogeneous convenience samples have clearer generalizability relative to conventional convenience samples. Therefore, when researchers are limited to convenience samples, they should consider homogeneous convenience samples as a positive alternative to conventional or heterogeneous) convenience samples. We discuss future directions as well as potential obstacles to expanding the use of homogeneous convenience samples in developmental science.

The roots of sociodemographic differences – including sexual orientation, gender, ethnicity, urbanicity, SES, culture, and nationality – in developmental processes and trends are complex and likely the product of layered interactions among biological, behavioral, and sociocultural factors (Betencourt & Lopez 1993; Crimmins & Saito, 2001; Jager & Davis-Kean, 2011; Phinney, 1996). Nonetheless, they are important to unpack because without a scientific base of knowledge regarding human health and behavior that takes into account the sociodemographic diversity of the population, health care delivery, planning, and policy making would be compromised by inadequate information and potentially misleading generalizations (Betencourt & Lopez, 1993; Mays, Ponce, Washington, & Cochran, 2003).

For this reason, a sizable amount of developmental science research is devoted to understanding developmental processes and trends in specific sociodemographic groups as

Correspondence concerning this article should be addressed to: Justin Jager, T. Denny Sanford School of Social and Family Dynamics, Arizona State University, PO Box 873701, Tempe AZ 85287-3701. Phone: 480-727-9673. Fax: 480-965-6779. justin.jager@asu.edu.

well as differences across two or more sociodemographic groups. Despite their disadvantages in generalizability relative to probability samples, much of this research relies on convenience samples–a fact that does not bode well for the field of developmental science (Bornstein, Jager, & Putnick, 2013). Developmental scientists should rely more on probability samples, for reasons we describe below. Nonetheless, because convenience samples are commonly used, we focus here on how developmental scientists can limit the disadvantages of convenience samples when it comes to understanding population effects as well as subpopulation differences. As we outline below, relative to conventional (or heterogeneous) convenience samples (i.e., samples that are open to all sociodemographic subgroups), homogeneous convenience samples (i.e., samples that are intentionally limited to specific sociodemographic subgroups and therefore homogeneous on one or more sociodemographic factors) should, on average, yield estimates with clearer, albeit narrower, generalizability and, therefore, provide more accurate accounts of population effects and subpopulation differences. On this basis, we argue that when researchers are limited to convenience samples, they should adopt homogeneous convenience samples as a positive alternative to conventional convenience samples.

Before distinguishing between conventional and homogeneous convenience samples, we compare and contrast convenience sampling in general with probability sampling and then, using an illustration, discuss in more depth the key disadvantage of all convenience samples: due to poor generalizability they often yield biased estimates of the target population and its sociodemographic subpopulations. Next, we describe conventional and homogeneous convenience sampling, and explain why, of the two, homogeneous convenience sampling provides clearer generalizability and, therefore, a more accurate account of its target population effects and subpopulation differences. We conclude by discussing future directions as well as potential obstacles to expanding the use of homogeneous convenience samples within developmental science.

Probability Sampling Versus Convenience Sampling

Within developmental science, sampling strategies generally fall into two broad categories: non-probability sampling and probability sampling (Bornstein et al., 2013; Levy & Lemeshow, 2011). Probability sampling strategies are any methods of sampling that utilize some form of random selection, which entails setting up a process or procedure that assures that different members of the target population have equal probabilities of being chosen. Probability sampling strategies include simple random sampling as well as more complex sampling designs such as stratified sampling and cluster sampling (and its variants such as probability proportional to size sampling; see Bornstein et al., 2013; Cochran, 1977; Levy & Lemeshow, 2011). The key advantage of probability sampling strategies is that they all, when carried out properly, should yield an unbiased sample that is representative of the target population. As a result, researchers can safely assume that estimates obtained from probability samples are both unbiased and generalizable. The key disadvantage of probability sampling strategies is that they present a significant challenge to execute. That is, the sizes of probability samples need to be quite large, often coming at great costs in terms of money, time, and effort. Moreover, designing probability samples requires substantial expertise. Indeed, due to the costs and challenges associated with probability samples, many of the prominent probability samples within developmental science are managed by Federal agencies or large research centers with substantial annual budgets, such as Add Health (http://www.cpc.unc.edu/projects/addhealth) and the Early Childhood Longitudinal Program (ECLS; http://nces.ed.gov/ecls). Furthermore, as Davis-Kean and Jager (chapter in this SRCD Monograph) discuss in more detail, most existing probability samples are ill-suited to examine developmental questions.

Non-probability sampling strategies are any methods of sampling that do not utilize some form of random selection. By far the most common non-probability sampling strategy used within developmental science is convenience sampling (for review see Bornstein et al., 2013), which is a sampling strategy where participants are selected in an ad hoc fashion based on their accessibility and/or proximity to the research. One of the most common examples of convenience sampling within developmental science is the use of student volunteers as study participants. The key advantages of convenience sampling are that it is cheap, efficient, and simple to implement. The key disadvantage of convenience sampling is that the sample lacks clear generalizability. Moreover, these advantages and disadvantages apply, albeit in varying degrees, to all types of convenience samples. Therefore, the advantages and disadvantages of convenience sampling are the reverse of probability sampling. Whereas probability samples yield results with clearer generalizability, convenience samples are far less expensive, more efficient, and simpler to execute.

Even though probability sampling is more advantaged in terms of scientific merit (i.e., probability sampling yields samples with clearer generalizability), convenience samples are the norm within developmental science. Bornstein et al. (2013) tallied the use of probability sampling and different types of nonprobability sampling from 2007 to 2011 in five prominent developmental science journals. Among the studies for which the type of sampling strategy could be conclusively determined, 92.5% utilized a convenience sample. Probability sampling accounted for only 5.5% of studies. Thus, from a tally of recent publications in prestigious journals in developmental science, convenience samples were the norm and were over 16 times more likely to be used than probability samples.

Poor Generalizability Leads to Estimate Bias: An Illustration

Because the generalizability of convenience samples is unclear, the estimates derived from convenience samples are often biased (i.e., sample estimates are not reflective of true effects among the target population because the sample poorly represents the target population). This bias extends to estimates of population effects as well as estimates of subpopulation differences. We illustrate these effects by outlining the known population parameters for the association between harsh parenting and externalizing as well as ethnic differences in that association, and then compare known population parameters to estimates obtained from three hypothetical convenience samples.

For the purposes of this illustration, we use the following target population: White and Black youth between the ages of 10 and 19 in the United States. Based on data from the United States Census and research on the association between harsh parenting and externalizing in this target population, we know the population parameters of the association between harsh parenting and externalizing with some confidence; they are listed in the first row of Table 1. Specifically, based on the 2010 United States Census (2012), the White-Black breakdown is roughly 80%/20%. Based on studies utilizing national probability samples of children and adolescents, the effect size (Cohen's *d*) for the White-Black difference in the income-to-needs ratio, a common indicator of socioeconomic status (SES), is around 1.0 (Davis-Kean & Sexton, 2009; Geronimus, Bound, Keene, & Hicken, 2007). For ease of interpretation SES is centered around the population mean and has a *SD* of 1.0. Consequently, given the White-Black population breakdown of 80%/20% and the effect size of 1.0 for the White-Black difference in income-to-needs ratio, mean SES for the White population = 0.2 and mean SES for the Black population = −0.8 (i.e., the difference between the White and Black means equals 1.0 and the weighted average of the White and Black means is 0). Based on extant research, among the total population (i.e., White and Black combined) harsh parenting is positively correlated with externalizing ($\rho \approx .25$; Bailey, Hill, Oesterle, & Hawkins, 2009; Burnette, Oshri, Lax, Richards, & Ragbeer, 2012; Rothbaum & Weisz, 1994). Finally, although some research suggests no ethnic difference in association between harsh parenting and externalizing (Berlin et al., 2009; Gershoff, Lansford, Sexton, Davis-Kean, & Sameroff, 2012; McLoyd & Smith, 2002), we focus on the substantial amount of research that suggests the association is higher for White children ($\rho^{w} \approx .30$) than for Black children ($\rho^{b} \approx .10$; Deater-Deckard, Dodge, Bates, & Pettit, 1996; Gunnoe & Mariner, 1997; Lansford, Deater-Deckard, Dodge, Bates, & Pettit, 2004). For the purposes of illustration, imagine that for both ethnic groups the association between harsh parenting and externalizing increases as levels of SES increase, but does so more for White adolescents ($\rho^{w} = .280 + .100*\text{SES}$)[1] than for Black adolescents ($\rho^{b} = .128 + .035*\text{SES}$)[2]. Although hypothetical, such an interaction is plausible because ethnicity and SES often interact with one another to inform psychosocial outcomes (Desimone, 1999; Kessler & Neighbors, 1986). The association between harsh parenting and externalizing across levels of SES (±2.0 *SD*) is graphed for each ethnic group in Figure 1; as the level of SES increases, the ethnic difference in the association between harsh parenting and externalizing also increases.

To devise hypothetical convenience samples, we use school-based samples as a heuristic. We do so because in developmental science a common sampling frame (e.g., the set of participants from which a sample is drawn) for convenience samples is a particular school or a particular school district, and schools and school districts can vary dramatically from one another in terms of ethnic and socioeconomic distribution. The demographic characteristics and sample estimates for each of three hypothetical convenience samples are listed in Table 1. In Sample A ("High White, High SES"), both White adolescents and those of higher SES are overrepresented, and the ethnic difference in SES is 40% smaller relative to the target population. In Sample B ("High Black, Low SES"), Black adolescents and those of lower SES are overrepresented, and the ethnic difference in SES is 70% smaller relative to the target population. Finally, in Sample C (High SES White, Low SES Black), each of the two ethnic groups is properly represented; however, among White adolescents those of higher

[1]When mean SES for the White adolescent population ($M^{w} = 0.2$) is applied to this equation, $\rho^{w} = .30$ (i.e., .30 = .280 + .100*0.2), which is the population parameter for White adolescents.

[2]When mean SES for the Black adolescent population ($M^{b} = -0.8$) is applied to this equation, $\rho^{b} = .10$ (i.e., .10 = .128 + .038*−0.8), which is the population parameter for Black adolescents.

SES are overrepresented, whereas among Black adolescents those of lower SES are overrepresented. As a result, in Sample C the ethnic difference in SES is three times larger than it is in the target population.

When considered individually, all three convenience samples yield misleading or biased estimates of the overall population effect (i.e., $r \neq \rho$) and of the subpopulation difference (i.e., $[r^{w} - r^{b}] \neq [\rho^{w} - \rho^{b}]$). Specifically, because White adolescents and adolescents of higher SES are overrepresented in Sample A and ρ is higher among White adolescents and those of higher SES, ρ is overestimated in Sample A ($r = .39$). Also, because $[\rho^{w} - \rho^{b}]$ is larger among adolescents of higher SES (Figure 1), and adolescents of higher SES are overrepresented in Sample A, $[\rho^{w} - \rho^{b}]$ is also overestimated in Sample A ($[r^{w} - r^{b}] = .25$). The estimates for Sample B are also biased; whereas the estimates for Sample A are too high, the estimates for Sample B are too low. Because Black adolescents and adolescents of lower SES are overrepresented in Sample B and ρ is lower among Black adolescents and adolescents of lower SES, ρ is underestimated in Sample B ($r = .11$). Also, because $[\rho^{w} - \rho^{b}]$ is smaller among those of lower SES (Figure 1), and those of lower SES are overrepresented in Sample B, $[\rho^{w} - \rho^{b}]$ is also underestimated in Sample B ($[r^{w} - r^{b}] = .10$). Finally, because ρ is particularly high among high-SES White adolescents, and high-SES adolescents are overrepresented in the White subsample of Sample C, both ρ and ρ^{w} are overestimated in Sample C ($r = .36$; $r^{w} = .43$). Whereas ρ^{w} is overestimated in Sample C, ρ^{b} ($r^{b} = .07$) is underestimated in Sample C (because high-SES adolescents are underrepresented among the Black subsample), and as a result $[\rho^{w} - \rho^{b}]$ is overestimated in Sample C ($[r^{w} - r^{b}] = .36$).

When considered collectively, how would researchers integrate the findings from these three convenience samples? Their estimates of population effects and subpopulation differences are inconsistent; therefore, they cannot all be correct estimates of the same target population. But is one sample’s set of estimates more valid (or less invalid) than the others? Because we know the true population parameters for our example, here we are able to judge the validity of each sample’s estimates. But for most scientific investigations using convenience samples to study a given developmental process, the true population parameters of the target population are not known. (If the true population parameters were known, then there would be no reason to undertake the study in the first place.) Therefore, when attempting to integrate inconsistent findings across a set of studies using convenience samples, investigations typically do not have the known population parameters to use as a benchmark. Although this example involved a set of hypothetical studies, substantial variation in the sociodemographic composition of convenience samples is all too common across studies examining a given developmental characteristic in an equivalent target population. Importantly, these variations make it difficult to determine whether inconsistencies across studies represent true population differences or instead are artifacts of differences in sample composition. Put succinctly, science is supposed to be cumulative; however, the use of convenience samples can translate into across-study inconsistencies that are difficult to integrate and, therefore, build upon.

Conventional Versus Homogeneous Convenience Sampling

Despite their disadvantaged generalizability, convenience samples are the standard within developmental science, and likely will remain so because probability samples are cost-prohibitive and most probability samples are ill-suited to examine developmental questions. Instead of focusing on how to reduce the use of convenience samples within developmental science, we focus here on how to limit their disadvantages when it comes to understanding population effects as well as subpopulation differences. Although all convenience samples have less clear generalizability than probability samples, not all convenience samples are the same, and some convenience samples have clearer generalizability than others. We argue that homogeneous convenience samples have clearer generalizability relative to conventional convenience samples. In developmental science, homogeneous convenience samples are far less common than conventional convenience samples. Therefore, we believe that one way to minimize the disadvantages of convenience samples is through the strategic use of homogenous convenience samples in place of conventional convenience samples. Below we describe in more detail what we mean by conventional and homogeneous convenience samples, and then we describe why, of the two, homogeneous convenience sampling has clearer generalizability. Next, we describe their advantages and disadvantages when it comes to estimating population effects as well as subpopulation differences.

Conventional convenience samples

The sampling frame for conventional convenience samples is not intentionally constrained based on sociodemographic background (i.e., participants of all sociodemographic backgrounds are eligible for participation). For example, aside from the fact that they were limited to two ethnic groups for the sake of simplicity, the three hypothetical convenience samples listed in Table 1 are conventional convenience samples. For these samples, the sampling frame was truly ad hoc (regardless of sociodemographics, all were welcome to participate provided they volunteered). We refer to these types of convenience samples as “conventional” because they are by far the most common type of convenience sample in developmental science; however, these types of convenience samples can also be conceptualized as *heterogeneous* convenience samples because, by design, the expectation is heterogeneity (i.e., diversity) in all sociodemographic factors As part of their tally of the types of sampling strategies within developmental science, Bornstein et al. (2013) found that among the studies that utilized a convenience sample, 89% were conventional convenience samples.

Homogeneous convenience samples

In contrast to conventional convenience sampling, the sampling frame for homogeneous convenience sampling is intentionally constrained with respect to sociodemographic background. In homogeneous convenience sampling researchers undertake to study (and therefore sample) a population that is homogeneous with respect to one or more sociodemographic factors (e.g., the overall population is composed of just Blacks or Whites). Thus, the target population (not just the sample studied) is a specific sociodemographic subgroup. For example, for a sample that is homogeneous with respect to ethnic group, the sampling frame is limited to, say, just Black Americans, and only Black

Americans are sampled. Homogeneous samples can differ in their degree of sociodemographic homogeneity. For example, the target population and its matching sample could be limited to one sociodemographic factor such as ethnicity (e.g., Black Americans); two sociodemographic factors such as SES and ethnicity (e.g., affluent Black Americans); three sociodemographic factors, such as gender, SES, and ethnicity (e.g., female, affluent, Black Americans), and so forth. The greater the number of homogeneous sociodemographic factors, the more homogeneous the sample and the narrower the sampling frame. Although relatively rare, homogeneous samples are used in developmental science, often to examine underrepresented sociodemographic groups (e.g., ethnic or sexual minorities). As part of their tally of the types of sampling strategies in developmental science, Bornstein et al. (2013) found that among the studies that utilized a convenience sample, only 8.6% were homogeneous convenience samples.

Homogeneous convenience samples offer narrower but clearer generalizability

The key advantage of homogeneous convenience samples, relative to conventional convenience samples, is their clearer generalizability. Because the sampling frame of homogeneous convenience samples is more homogeneous than the sampling frame for conventional convenience samples, researchers can be more confident with respect to generalizability. Why does a more homogeneous sampling frame translate into clearer generalizability? Logic dictates that the more homogeneous a population, the easier (more probable) it is to generate a representative sample, even when using convenience sampling. Therefore, by intentionally constraining the sampling frame to reduce the amount of sociodemographic heterogeneity, the chance of bias in sampling, as it relates to sociodemographic characteristics of the target population, is reduced (although not all together eliminated).

Imagine two different convenience samples that seek to examine the same developmental process. Each convenience sample consists of 500 families, and both samples are taken from the same large Midwestern city. The first is a conventional convenience sample and, because it does not limit its sampling frame with respect to any sociodemographic factors, contains at least some amount of heterogeneity on many sociodemographic factors. The second is a homogeneous convenience sample and, because it limits its sampling frame with respect to ethnicity (only samples Black families), SES (only samples middle-class families), and national origin (only samples families within which both birth parents were born in the United States), it contains no heterogeneity on these sociodemographic factors. Now imagine that the findings differed between the two samples, which would not be surprising given the stark sociodemographic differences between the two samples. Which sample's findings would have clearer generalizability? In our view, the findings from the homogeneous convenience sample would have the clearer generalizability. That is, we could be more confident that the findings from the homogeneous convenience sample generalize to middle-class, native-born, Black families than we could be that the findings from the conventional convenience sample generalize to all families (regardless of ethnicity, class, or national origin). This is because, in comparison to the conventional convenience sample, the homogeneous convenience sample should, on average, have a sociodemographic distribution that more closely reflects the sociodemographic distribution of its target population, and therefore, its estimates of its target population should, on average, be more accurate, precise, and valid.

The key disadvantage of homogeneous convenience samples, relative to conventional convenience samples, is their narrower generalizability. Although homogeneous convenience samples have clearer generalizability, their findings also generalize to a more circumscribed population. Returning to the example above, although the findings from the homogeneous convenience sample of middle-class, native-born, Black families have clearer generalizability than do the findings from the conventional convenience sample of all families, the findings from the homogeneous convenience sample, at best, only generalize to middle-class, native-born, Black families. Therefore, the findings from the homogenous convenience sample reveal very little if anything about families that are not middle-class, native-born, and Black. Another disadvantage of homogeneous convenience samples is that, if they are samples of underrepresented sociodemographic groups, they can be more costly and time consuming relative to conventional convenience samples. For example, more effort is involved in recruiting 350 Gay/Lesbian adolescents of Hispanic descent from lower-class families than is involved in recruiting 350 adolescents of any sexual orientation, ethnicity, or social class. Finally, as is the case for all types of sampling, clearly and accurately defining one's target population is essential for homogeneous convenience sampling to be effective. After all, to maximize the alignment (and therefore generalizability) between a sample and its target population, researchers must have a firm and detailed understanding of their target population.

Although the generalizability of homogeneous convenience samples is clearer, if narrower, relative to conventional convenience samples, we emphasize that both homogeneous convenience samples and conventional convenience samples have poor generalizability relative to probability samples. On a hypothetical continuum of generalizability, probably samples are at one end and conventional convenience samples are at the other end. Homogeneous convenience samples fall somewhere in between, although likely closer to conventional convenience samples. However, the more homogeneous they are (i.e., the more sociodemographic factors that are homogeneous), the closer they fall *in terms of generalizability* to probability samples. Again though, the more homogeneous they are, the narrower their generalizability. Therefore, with respect to the estimation of population effects, homogeneous convenience samples should, on average, provide more accurate population estimates, albeit of a more circumscribed population. We now turn to the implications of these arguments for the estimation of subpopulation differences.

Homogeneous samples and subpopulation differences

If a homogeneous convenience sample is homogeneous with respect to the sociodemographic factor of interest, then homogeneous convenience samples are ill-suited for directly examining sociodemographic differences (e.g., a convenience sample homogeneous with respect to ethnicity is not equipped to examine ethnic differences). However, relative to conventional convenience samples, homogeneous convenience samples are better-suited to address sociodemographic differences when aggregating across a series of studies. Consider the six conventional convenience samples listed in Table 2a. Like the three convenience samples in Table 1, each of the six samples in Table 2a is heterogeneous with respect to both ethnicity and SES and all six samples have varying sociodemographic representation with respect to ethnicity and SES. As was the case for the three convenience samples in Table 1, likely the six conventional convenience samples in Table 2a would yield conflicting findings. These conflicting findings would be difficult to explain because there is ethnic and SES heterogeneity *both* within and across the samples. The reasoning is similar to that of a 2*3 ANOVA, with ethnicity (Black and White) and SES (low, middle, high) as the two factors. In this example, the ANOVA would have six cells just as Table 2a has six cells. Within an ANOVA, for the between-factor variance to be separated from the within-factor variance each cell must be homogeneous with respect to both factors (i.e., all variance in factors is across-cell). The problem with conventional convenience samples is that there is heterogeneity for both sociodemographic factors (i.e., ethnicity and SES) within each of the six samples as well as heterogeneity for both factors across the six samples. As a result, it is difficult to parse exactly how, if at all, ethnicity and SES heterogeneity contribute to between-sample differences in findings.

Next, consider the six homogeneous convenience samples listed in Table 2b. Each of the six homogeneous convenience samples is homogeneous with respect to both ethnicity and SES; however, they vary as to which ethnic group and which category of SES is homogenous. Like the six conventional convenience samples, the six homogeneous convenience samples would likely yield conflicting findings, but unlike the six conventional convenience samples, for the six homogeneous convenience samples any between-sample differences in findings could be reasonably attributed to ethnic and/or SES heterogeneity. Because all ethnic and SES heterogeneity is between-sample for the homogeneous convenience samples, between-sample differences in findings can be more clearly attributed to ethnic and SES heterogeneity, or at least they can be with greater confidence relative to the conventional convenience samples. Thus, when considered individually each of the six homogeneous convenience samples has narrower but clearer generalizability than each of the six conventional convenience samples, and when considered collectively the homogeneous convenience samples also provide a more accurate and encompassing account of sociodemographic differences than do conventional convenience samples.

If a homogeneous convenience sample is heterogeneous with respect to the sociodemographic factor of interest, then it is well suited for directly examining sociodemographic differences. For example, a homogeneous convenience sample that is homogeneous with respect to SES, but heterogeneous with respect to ethnicity, is well-suited to examine ethnic differences because more than one ethnic group can be compared, while holding SES constant. Moreover, the key advantage (i.e., clearer generalizability) and disadvantage (narrower generalizability) of homogeneous convenience samples relative to conventional convenience samples would also apply to the examination of sociodemographic differences. However, like conventional convenience samples, homogeneous convenience samples may lack sufficient power to detect group differences, leading to Type II errors. Imagine a homogeneous convenience sample ($N = 200$) that matches the first data column in Table 2b such that it is heterogeneous in ethnicity (includes White and Black adolescents) and homogeneous with respect to SES (includes only low-SES adolescents or those with SES < -1.0). Because Black adolescents represent only 20% of the population, only 40

Black participants would be expected in the sample. For example, based on α = .05, in an ANOVA design with two groups (or an independent samples *t*-test) this homogeneous convenience sample yields sufficient power (≥ .80) to detect Black-White differences provided the effect size (*d*) is ≥ .50 (where power is determined by $\delta = d * \sqrt{(n_1 n_2)/(n_1 + n_2)}$; Faul, Erdfelder, Lang, & Buchner, 2007).

An alternative to a homogeneous convenience sample would be a quota homogeneous convenience sample. In quota sampling, another form of non-probability sampling, fixed numbers of participants from different sociodemographic groups are recruited, typically using convenience sampling (e.g., separate samples of White and Black adolescents, each of which are convenience samples of 100 individuals). Returning to the example above, imagine a homogeneous convenience sample (N = 200) that matches the first data column in Table 2b such that it is heterogeneous with respect to ethnicity (includes both White and Black adolescents), homogeneous with respect to SES (includes only low-SES adolescents or those with SES < −1.0), but includes equal numbers of White and Black adolescents. Based on α = .05, in an ANOVA design with two groups this quota homogeneous convenience sample yields sufficient power (≥ .80) to detect Black-White differences provided the effect size (*d*) is ≥ .40, which translates into a 20% reduction in the size of the minimally detectable effect relative to the homogeneous convenience sample.

Looking Forward

Despite their disadvantaged generalizability relative to probability samples, convenience samples are the standard within developmental science, and likely will remain so because probability samples are cost-prohibitive and most available probability samples are ill-suited to examine developmental questions. The advantaged generalizability of probability samples is both important and well-documented within the sampling literature, but it obscures the fact that, in terms of generalizability, some convenience samples are less disadvantaged than others. Therefore, in addition to comparing and contrasting the merits of probability samples with convenience samples, we believe that the field should devote more attention to comparing and contrasting the merits of different types of convenience samples. After all, given the prevalence of convenience samples within developmental science, it behooves developmental scientists to minimize disadvantages when it comes to generalizability. With respect to generalizability, we believe that homogenous convenience samples as well as quota homogeneous convenience samples have key advantages over conventional conveniences samples and should be used more. To be clear, we are not advocating for the increased use of convenience samples within developmental science, as we believe the opposite and advocate that the use of probability samples within developmental science should increase. However, when researchers are limited to convenience samples, we advise adopting homogeneous convenience samples as a positive alternative to conventional convenience samples. Additionally, we are not advocating for the elimination of conventional convenience samples. Instead, we recommend that the current ratio of conventional to homogeneous convenience samples within developmental science, which is about 11 to 1 (Bornstein et al., 2013), is not optimal for our science and should be brought into balance. That said, given existing paradigms within developmental science, we see at least two obstacles to the increased adoption of homogeneous convenience samples.

The first may be a concern on the part of researchers that it could be difficult to obtain protocol approval by internal review boards (IRBs) or to secure external funding for homogeneous convenience samples. Many funding agencies require (or strongly recommend) inclusion of all major sociodemographic groups. For example, the National Institutes of Health released guidelines about including women and minorities in clinical research in 1994 (revised in 2001; NIH Office of Extramural Research, 2001) which indicate that all grant applications are evaluated for the inclusion of sociodemographic groups, and if groups are omitted, a strong justification is required. Although one way to avoid the potential ire of IRBs and funding agencies is to collect a heterogeneous convenience sample and then limit one's analyses to a homogeneous subsample (e.g., limit analyses to only European Americans), this approach has two key limitations. Aside from its inefficiency, it is only feasible for sociodemographic subgroups that are well-represented among the target population (and already well-represented within developmental research). For example, one could not collect a heterogeneous convenience sample and then limit the analyses to only Native Americans because, more than likely, the sample size of Native Americans would be far too small to examine on its own. Therefore, instead of collecting conventional or heterogeneous convenience samples and restricting analyses to a homogeneous subsample, we encourage researchers to make principled theoretical and statistical arguments to support their choices of better sampling strategies, even if the strategy proposed is homogeneous with respect to one or more sociodemographic groups. Researchers may be required to provide scientific and practical justification to IRBs and parents/community leaders to explain why certain groups are being excluded from study; some statistical justifications are provided herein. There also are specific steps that granting agencies and journal editors can take to encourage the use of homogeneous convenience samples. For example, granting agencies could set aside funds to support research using homogeneous convenience samples to study underrepresented (and understudied) subpopulations. Additionally, journal editors could organize special issues that are limited to studies that use homogeneous convenience samples to examine a specific substantive topic, say adolescent attachment, but vary as to the sociodemographic group of focus.

The second potential obstacle is developmental scientists' reticence to share data. In many cases researchers hold exclusive rights to their data and tightly restrict access to their data, although this will become less of an obstacle over time, given that NIH now requires that all applications provide data sharing plans. However, for scientific knowledge to accumulate, researchers using homogeneous convenience samples will either have to share their data with other researchers or alternatively work collaboratively with other researchers in a manner that is not currently common. For example, returning to the six homogeneous convenience samples listed in Table 2b, to examine differences across ethnicity or SES directly, two or more of the six research teams would have to work together in one of two ways. At the point of data collection, they could coordinate their efforts so that they use the same measures and procedures and are therefore able to pool and analyze all data once collected. However, this level of coordination prior to data collection often proves challenging given varying priorities, time-tables, and resources across research teams. As another possibility, even if different research teams do not use the exact same measures, they could still integrate their data where possible post-hoc using integrative data analysis (Curran & Hussong, 2009). Additionally, using meta-analysis, a single research team could examine whether a given effect size varies depending on a sample's sociodemographics. However, in many cases this may prove difficult because developmental studies often do not provide detailed information regarding their sample's sociodemographics (see Bornstein et al., 2013).

Our core thesis is that because convenience samples of homogeneous populations (i.e., homogeneous convenience samples) are more likely to be representative than convenience samples of heterogeneous populations (i.e., conventional convenience samples), homogeneous convenience samples should, on average, yield more valid (less unbiased) estimates than conventional convenience samples. However, in terms of providing valid estimates, relative to homogeneous convenience samples, how much more disadvantaged are conventional convenience samples and how much more advantaged are probability samples? Moreover, are there specific conditions under which homogeneous convenience samples perform particularly better than conventional convenience samples (e.g., samples of smaller size or when multiple sociodemographic factors are homogeneous)? It is important for future research to address these issues both theoretically and statistically because the answers found will reveal exactly how much the field has to gain from the increased use of homogenous convenience samples.

Acknowledgments

This research was supported by the Intramural Research Program of the NIH, NICHD.

References

Bailey JA, Hill KG, Oesterle S, Hawkins JD. Parenting practices and problem behavior across three generations: Monitoring, harsh discipline, and drug use in the intergenerational transmission of externalizing behavior. Developmental Psychology. 2009; 45(5):1214–1226. [PubMed: 19702387]

Berlin LJ, Ispa JM, Fine MA, Malone PS, Brooks-Gunn J, Brady-Smith C, Ayoub C, Bai Y. Correlates and consequences of spanking and verbal punishment for low-income White, African American, and Mexican American toddlers. Child Development. 2009; 80(5):1403–1420. [PubMed: 19765008]

Bettencourt H, Lopez SR. The study of culture, ethnicity, and race in American psychology. American Psychologist. 1993; 48:629–637.

Bornstein MH, Jager J, Putnick DL. Sampling in developmental science: Situations, shortcomings, solutions, and standards. Developmental Review. 2013; 33:357–370. [PubMed: 25580049]

Burnette ML, Oshri A, Lax R, Richards D, Ragbeer SN. Pathways from harsh parenting to adolescent antisocial behavior: A multidomain test of gender moderation. Development and Psychopathology. 2012; 24:857–870. [PubMed: 22781859]

Curran PJ, Hussong AM. Integrative data analysis: The simultaneous analysis of multiple data sets. Psychological Methods. 2009; 14(2):81–100. [PubMed: 19485623]

Cochran, WG. Sampling Techniques. 3. New York: Wiley; 1977.

Crimmins EM, Saito Y. Trends in health life expectancy in the United States, 1970–1990: Gender, racial, and educational differences. Social Science and Medicine. 2001; 52:1629–1641. [PubMed: 11327137]

Davis-Kean PE, Sexton HR. Race differences in parental influences on child achievement: Multiple pathways to success. Merrill-Palmer Quarterly. 2009; 55(3):285–318.

Deater-Deckard K, Dodge KA, Bates JE, Pettit GS. Physical punishment among African American and European American mothers: Links to children's externalizing behaviors. Developmental Psychology. 1996; 32:1065–1072.

Desimone L. Linking parent involvement with student achievement: Do race and income matter? The Journal of Educational Research. 1999; 93:11–30.

Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. Behavior Research Methods. 2007; 39:175–191. [PubMed: 17695343]

Geronimus AT, Bound J, Keene D, Hicken M. Black-white differences in age trajectories of hypertension prevalence among adult men and women, 1999–2002. Ethnicity and Disease. 2007; 17:40–48. [PubMed: 17274208]

Gershoff ET, Lansford JE, Sexton HR, Davis-Kean P, Sameroff AJ. Longitudinal links between spanking and children's externalizing behaviors in a national sample of White, Black, Hispanic, and Asian American Families. Child Development. 2012; 83(3):838–843. [PubMed: 22304526]

Gunnoe ML, Mariner CL. Toward a developmental-contextual model of the effects of parental spanking on children's aggression. Archives of Pediatric and Adolescent Medicine. 1997; 151:768–775.

Jager J, Davis-Kean PE. Same-sex sexuality and adolescent psychological well-being: The influence of sexual orientation, early reports of same-sex attraction, and gender. Self and Identity. 2011; 10(4): 417–444. [PubMed: 22505839]

Kessler RC, Neighbors HW. A new perspective on the relationships among race, social class, and psychological distress. Journal of Health and Social Behavior. 1986; 27:107–115. [PubMed: 3734380]

Lansford JE, Deater-Deckard K, Dodge KA, Bates J, Pettit GS. Ethnic differences in the link between physical discipline and later adolescent externalizing behaviors. Journal of Child Psychology and Psychiatry. 2004; 45:801–812. [PubMed: 15056311]

Levy, PS., Lemeshow, S. Sampling of Populations: Methods and Applications. 4. New York: Wiley; 2011.

Mays VM, Ponce NA, Washington DL, Cochran SD. Classification of race and ethnicity: Implications for Public Health. Annual Review of Public Health. 2003; 24:83–110.

McLoyd VC, Smith J. Physical discipline and behavior problems in African American, European American, and Hispanic Children: Emotional Support as a moderator. Journal of Marriage and Family. 2002; 64:40–53.

National Institutes of Health. Office of Extramural Research. NIH Policy and Guidelines on The Inclusion of Women and Minorities as Subjects in Clinical Research – Amended, October, 2001. 2001. Available online: http://grants.nih.gov/grants/funding/women_min/guidelines_amended_10_2001.htm

Phinney JS. When we talk about American ethnic groups, what do we mean? American Psychologist. 1996; 51:918–927.

Rothbaum F, Weisz JR. Parental caregiving and child externalizing behavior in nonclinical samples: A meta-analysis. Psychological Bulletin. 1994; 116(1):55–74. [PubMed: 8078975]

U.S. Census Bureau. Current Population Survey, Annual Social and Economic Supplement, 2011. Washington, DC: 2012.

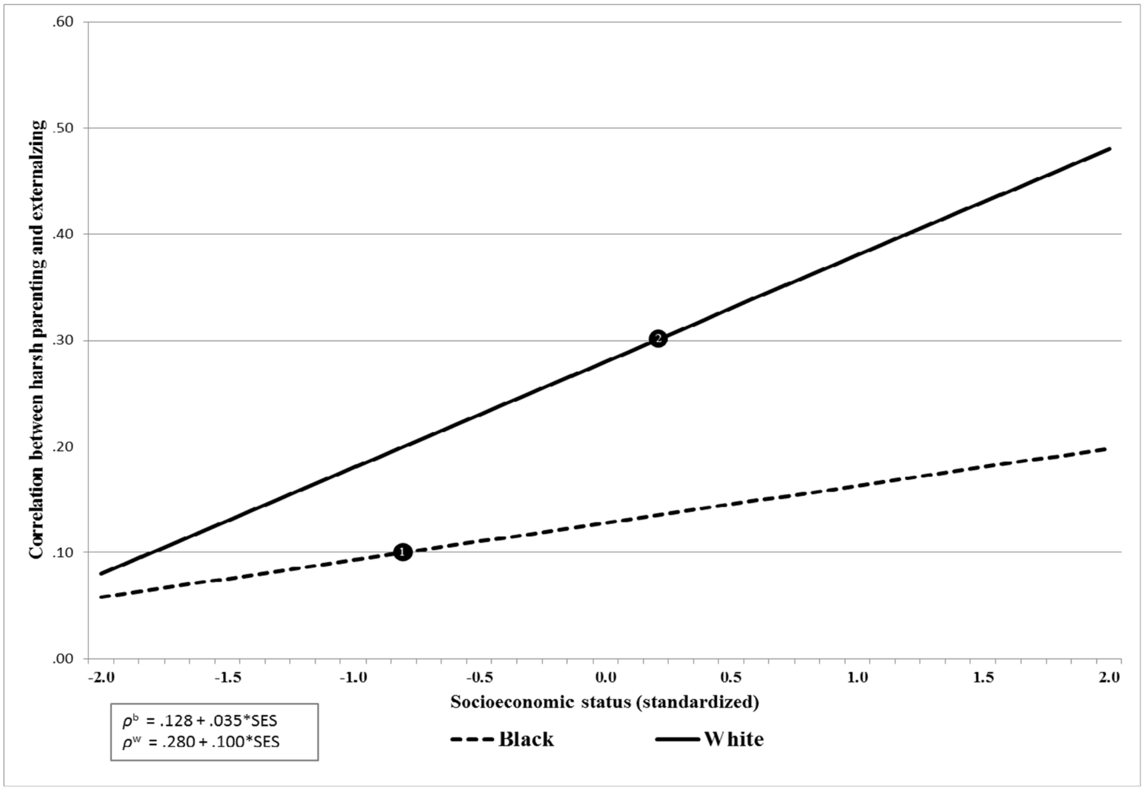

Figure 1.

Correlation between harsh parenting and externalizing, by ethnicity and SES. ❶ = population parameter (i.e., ρ^b = .10) for the Black adolescent population, for which mean SES = −0.8 (i.e., .10 = .128 + .035*−0.8). ❷ = population parameter (i.e., ρ^w = .30) for the White adolescent population, for which mean SES = 0.2 (i.e., .30 = .280 + .100*0.2)

Table 1

Population parameters and sample estimates for the association between harsh parenting and externalizing among White and Black adolescents

	Demographic characteristics						Association between harsh parenting and externalizing							
	Representation		Socioeconomic status				Population parameters (ρ)				Sample estimates (r)			
	White	Black	Total	White	Black	Ethnicity d	ρ	ρ^w	ρ^b	$\rho^w - \rho^b$	r	r^w	r^b	$r^w - r^b$
Target population	80%	20%	.00	.20	−.80	1.0	.25	.30	.10	.20				
Sample A: High White, High SES	95%	5%	1.20	1.23	.63	.6					.39	.40	.15	.25
Sample B: High Black, Low SES	30%	70%	−1.20	−.99	−1.29	.3					.11	.18	.08	.10
Sample C: High SES White, Low SES Black	80%	20%	.90	1.50	−1.50	3.0					.36	.43	.07	.36

Note: Assuming $N = 500$,

^{w}White;

^{b}Black

Table 2

Sociodemographic characteristics of hypothetical convenience samples, by SES, ethnicity, and type of convenience sample

	SES		
	Low	**Medium**	**High**
Ethnicity			
		a) Conventional convenience samples	
Black	30% White/70% Black *M* SES = −1.5 (−2.5|2.5)	25% White/75% Black *M* SES = 0.0 (−2.5|2.5)	15% White/85% Black *M* SES = 1.5 (−2.5|2.5)
White	90% White/10% Black *M* SES = −1.5 (−2.5|2.5)	92% White/8% Black *M* SES = 0.0 (−2.5|2.5)	88% White/12% Black *M* SES = 1.5 (−2.5|2.5)
		b) Homogeneous convenience samples	
Black	0% White/100% Black *M* SES = −1.5 (<−1.0)	0% White/100% Black *M* SES = 0 (≥ −1.0 and ≤ 1.0)	0% White/100% Black *M* SES = 1.5 (> 1.0)
White	100% White/0% Black M SES = −1.5 (<−1.0)	100% White/0% Black *M* SES = 0 (≥ −1.0 and ≤ 1.0)	100% White/0% Black *M* SES = 1.5 (> 1.0)

Note: For each conventional convenience sample, participants from all levels of SES are represented to some extent. For example, among the Low SES conventional convenience samples, although the average SES is −1.5, there are some participants with Medium (≥ −1.0 and ≤ 1.0) and High SES (> 1.0) as well. For each homogeneous convenience sample, participants from only a single level of SES are represented. For example, among the Low SES homogeneous convenience samples, the samples are limited to those with Low SES (<−1.0).