

ELIZABETH T. GERSHOFF *University of Texas*

ANDREW GROGAN-KAYLOR *University of Michigan*

Race as a Moderator of Associations Between Spanking and Child Outcomes

The cultural normativeness perspective argues that parenting practices such as spanking are more beneficial for children when they occur in cultural groups within which they are normative. Research on this issue in the United States has focused on race as a marker of culture, and findings have been mixed. The present study presents meta-analyses of five studies that reported effect sizes separately for White (n=11,814) and Black (n=3,065) American children (5 to 14 years of age). Mean weighted effect sizes for both groups indicated statistically significant associations with detrimental outcomes; they were not statistically significantly different from one another. Contrary to the cultural normativeness perspective, these results demonstrate that spanking is similarly associated with detrimental outcomes for White and Black children in the United States.

Parenting beliefs, values, and goals are known to vary as a function of environmental and cultural contexts (Bronfenbrenner & Morris, 2006; Dasen & Mishra, 2000; Rogoff, 2003). As a result, parenting practices vary across cultures because cultures differ in which parenting practices they believe will promote those values and goals (Mistry, Chaudhuri, & Diez, 2003;

Super & Harkness, 1986). Such cross-cultural variation in a variety of parenting practices has been observed in several recent multinational comparisons (Bornstein et al., 2012; Gershoff et al., 2010; Lansford et al., 2014; Runyan et al., 2010).

In the United States, the notion of cultural differences in parenting has largely focused on one indicator of culture, namely the race or ethnicity of the parents, and on one indicator of parenting, namely physical punishment. Research on cultural differences in parenting has been driven largely by interest in the cultural normativeness perspective (Deater-Deckard & Dodge, 1997). According to this perspective, parenting practices have more beneficial (or fewer detrimental) effects on children when they occur in cultural groups within which they are normative, and they have more detrimental (and fewer beneficial) effects when they are non-normative (Deater-Deckard & Dodge, 1997). The normativeness perspective hypothesizes that if physical punishment is administered in a cultural context in which spanking is considered normative and acceptable, then the child who is spanked will be more likely to accept and comply with the parents' disciplinary message, thus reducing negative behavior over time. By contrast, if physical punishment is administered in a context in which spanking is less normative and more aberrant, then the child will likely reject the parents' disciplinary message, and the discipline will be ineffective in promoting appropriate behavior and may instead elicit negative reactive behavior. This hypothesis grew

Department of Human Development and Family Sciences,
University of Texas, 108 East Dean Keeton Street, Austin,
TX 78712 (liz.gershoff@austin.utexas.edu).

Key Words: Cultural normativeness, race, spanking.

in part from observations that Black parents tend to both endorse (Flynn, 1998; Mosby, Rawls, Meehan, Mays, & Pettinari, 1999) and use (Grogan-Kaylor & Otis, 2007; Slade & Wissow, 2004) physical punishment more often than parents of other racial or ethnic groups, and in part from findings in some studies that harsh physical punishment was not linked with child externalizing behavior among Black families but was linked with higher child externalizing behaviors among White families (Deater-Deckard, Dodge, Bates, & Pettit, 1996). The cultural normativeness perspective thus proposed that physical punishment has no impact, or a less detrimental impact, on children in Black families, who tend to accept the practice as a normal part of being a child in a Black family than on children in White families, in which physical punishment is less normative (Deater-Deckard & Dodge, 1997).

There is debate about whether race is a true marker of culture. According to some theoretical perspectives, ethnic groups are linked by a shared culture and social history, whereas race groups are linked by physical similarities and shared geographic origin (Hall, Yip, & Zárate, 2016). Yet in practice, race and ethnicity are often conflated, particularly for Black Americans, for whom a shared race means a shared history of slavery and continued pervasive discrimination based on skin color and geographic origin (Umaña-Taylor et al., 2014). Thus, while a person's race in and of itself does not connote a particular culture, in the United States black skin connotes shared experiences that may sometimes function as a culture that can shape parenting practices, including discipline. That said, using race to define culture is limiting, as ascribed race likely does not do justice to the diversity of cultural beliefs and practices within a race group. There is substantial variation in attitudes about and use of physical punishment within the Black community (Kelley, Power, & Wimbush, 1992). To establish current racial differences in support of spanking for the present study, we examined racial differences in attitudes about spanking in the 1986–2014 administrations of the General Social Survey, a recurring national survey of adults in the United States (Smith, Marsden, Hout, & Kim, 2016). We found that Black respondents were more likely to have favorable attitudes toward spanking than White respondents (85% vs. 73%), but also that there was considerable variation within each racial group. Indeed, race explained

less than 2% of the variance in attitudes toward spanking in a regression model. These findings indicate that race is not a strong determinant of spanking attitudes and that the majority of both Blacks and Whites in the United States are supportive of spanking.

In examining race as a marker of culture, several studies have found patterns consistent with cultural normativeness theory. Spanking has been found to be associated with more externalizing in children from White families and with less externalizing among children from Black families in some studies (Gunnoe & Mariner, 1997; McLeod, Kruttschnitt, & Dornfeld, 1994). Three additional studies have reported this same pattern of results extending prospectively from 5 to 13 years of age (Deater-Deckard et al., 1996; Lansford, Deater-Deckard, Dodge, Bates, & Pettit, 2004; Lansford, Wager, Bates, Dodge, & Pettit, 2012), but each used the same data set that formed the basis of Deater-Deckard and Dodge's (1997) theory, and thus should be viewed as extensions of the same finding, not replications of it.

However, not all studies of race and physical punishment have been consistent with the cultural normativeness perspective. Indeed, a series of methodologically strong studies has failed to replicate the finding that race moderates the links between physical punishment and child outcomes. Using longitudinal data and controlling for children's early levels of behavior problems, these studies have found that physical punishment was linked with increases in children's problematic behavior over time within both White and Black families (Lau, Litrownik, Newton, Black, & Everson, 2006; Pardini, Fite, & Burke, 2008). Additional longitudinal and nationally representative studies that controlled for early problem behavior have also failed to find differences in the extent to which physical punishment predicts increases in problem behavior when Black, Latino, and White families are compared (Berlin et al., 2009; Grogan-Kaylor, 2005; McLoyd & Smith, 2002).

Despite this accumulation of evidence contradicting the cultural normativeness perspective on physical punishment's effects on children, cultural normativeness remains a prevalent argument regarding race and physical punishment. Horn, Joseph, and Cheng (2004) conducted a systematic review (but not meta-analysis) of seven studies examining the link between physical punishment and child outcomes in

Black families and concluded that physical punishment appeared to be associated with positive or neutral outcomes in longitudinal studies. These authors used a broad definition of physical punishment and thus included studies that assessed the use of objects to punish children and physical restraint. In the present study, we conducted a more rigorous test of whether race moderates the associations between physical punishment and child outcomes by restricting our analysis to studies of spanking and by conducting a statistical meta-analysis of studies that reported effect sizes for physical punishment separately by racial group. If physical punishment is linked with positive or neutral outcomes for Black children but with negative outcomes for White children, the cultural normativeness perspective would be supported. If physical punishment is instead linked with negative outcomes for children from both Black and White families, or for only children from Black families, the cultural normativeness argument would be rejected.

The goal of the present study was to use meta-analysis to ascertain the level of support for the notion that cultural normativeness, as indicated solely by a family's race, moderates links between spanking and children's outcomes. We acknowledge the limitation of using race as a marker of normativeness of physical punishment given the substantial variation in attitudes about and use of physical punishment within the Black community (Kelley et al., 1992). There are unfortunately too few studies that have directly assessed community norms about physical punishment to conduct a meta-analysis of those studies (e.g., Gershoff et al., 2010; Lansford et al., 2005), and we are aware of no study that has compared normativeness of physical punishment across race or ethnic groups in the United States. This study thus provides a narrow test of the cultural normativeness perspective by focusing solely on the issue of whether any links between parents' use of spanking and children's behaviors are moderated by the race of the family.

METHOD

Sample of Studies

Five studies met all criteria for inclusion in the meta-analyses; that is, they included (a) measures of parents' use of spanking and (b)

bivariate statistics per recommendations for meta-analyses (Borenstein, Hedges, Higgins, & Rothstein, 2009) and separately for Black and White subsamples of families, and (c) they were published in a peer-reviewed journal to ensure some degree of study quality. Because there were so few studies, each is described briefly here, with key characteristics of the studies summarized in Table 1.

The study by Christie-Mizell, Pryor, and Grossman (2008) used data from the National Longitudinal Study of Youth, Child sample (NLSY-C), a national survey that oversampled racial minorities. The full sample consisted of 1,139 White ("European American") children and 713 Black ("African American") 6- to 14-year-old children and their mothers. Mothers reported how often they spanked their children in the previous week in the 1992 interview. The child outcome, depressive symptoms, was measured 2 years later (1994) with mother ratings of five items from the Behavior Problems Index (Peterson & Zill, 1986). The effect size from Christie-Mizell et al. was thus longitudinal, but measures of spanking and the child outcome were provided by the same rater (mother).

Deater-Deckard et al. (1996) reported data from 466 White ("European American") and 100 Black ("African American") families living in three cities (Nashville, TN; Knoxville, TN; or Bloomington, IN), although the sample for the correlation used in these analyses was 372 White and 88 Black families. Mothers' use of "physical discipline" was a composite of three measures at children's entry to kindergarten: an interviewer's rating of mothers' responses to open-ended discipline items; mothers' responses to hypothetical vignettes (whether they voluntarily said they would use any form of physical punishment, "including spanking and hitting" [p. 1067]); and mothers' responses to the aggression subscale of the Conflict Tactics Scale (Straus, 1979), which included "threatening or actually throwing something at the child, hitting or spanking the child with the hand, or hitting or spanking the child with an object" (p. 1067). The outcome variable was a composite of children's externalizing behaviors as rated by teachers and by peers. Teacher ratings of children's externalizing behaviors were obtained in kindergarten, first grade, second grade, and third grade using the Child Behavior Checklist Teacher's Report Form (Achenbach, 1991) and were averaged across the four waves. Peer

Table 1. *Characteristics of Studies Included in the Meta-analyses*

| Study | Design | Different raters of spanking and outcome? | Measure of physical punishment | Child outcome | Association between race and frequency of physical punishment | Age of children at spanking measure | Location of sample |
|-------------------------------|--|---|---|--|---|-------------------------------------|---|
| Christie-Mizell et al. (2007) | Longitudinal | No | Spanking | Depressive symptoms | $mean_w = .34$; $mean_b = .85$; $p < .001$ | 6-14 years | National (NLSY-C) |
| Deater-Deckard et al. (1996) | Longitudinal | Yes | Composite that includes spanking and hitting with an object | Externalizing behavior problems | $r_b = .22, p < .001^a$ | Entering kindergarten | Nashville, TN, Knoxville, TN, Bloomington, IN |
| Foshee et al. (2005) | Longitudinal (dating violence); Cross-sectional (physical abuse) | Yes | Spanking or hitting | (a) Dating violence; (b) victim of physical abuse | $mean_w = .44$; $mean_b = .47$; ns | $M = 13.8$ years | Johnston County, NC |
| Gershoff et al. (2012) | Longitudinal | Yes | Spanking | Externalizing behavior problems | $mean_w = 1.14$; $mean_b = 1.59$; $p < .01$ | $M = 6.2$ years | National (ECLS-K) |
| McLeod et al. (1996) | Cross-sectional | No | Spanking | Antisocial behavior | $mean_w = .68$; $mean_b = .88$; $p < .01$ | $M = 9.13$ years | National (NLSY-C) |

^w White subsample.

^b Black subsample.

^a White was reference group for the dichotomous race variable.

sociometric ratings of a child's aggression and conflict with the teacher were obtained at the same four waves and averaged. The teacher and peer ratings were then combined to create a single externalizing composite. The effect size from Deater-Deckard et al. included longitudinal data in the outcome composite variable and used separate raters for spanking (mothers) and the child outcome (teachers and peers). However, the measure of physical punishment in this study included use of objects to hit or spank a child, which is a potentially abusive practice. Thus, the meta-analysis was run once including this study and a second time excluding it.

Foshee, Ennett, Bauman, Benefield, and Suchindran (2005) reported data from a study of 958 White and 188 Black adolescents in a rural county in North Carolina. When they were in eighth or ninth grade, the adolescents reported how often their mother spanked or hit them when punishing them, with response options ranging from *never* (0) to *very often* (3). They also responded to the question, "How often has an adult ever hit you with the purpose of hitting you?" Response options were *never* (1), *1 to 3 times* (2), *4 to 9 times* (3), and *10 or more times* (4); this item was used as an indicator of physical abuse, an outcome variable. Nineteen months later, the adolescents reported how often they had engaged in violence against a person with whom they had a date; dating violence included scratched, slapped, kicked, bit, burned, hit with a fist, beat up, and assaulted with a knife or gun. Dating violence was the second outcome variable included in the meta-analyses from this study. Both outcomes are rated by the same person (the adolescent), and one is longitudinal (dating violence) while the other is cross-sectional (victim of physical abuse).

The study by Gershoff, Lansford, Sexton, Davis-Kean, and Sameroff (2012) used data from 7,057 White and 1,352 Black families in the Early Childhood Longitudinal Study, Kindergarten Cohort 1998–1999 (ECLS-K), a nationally representative study of children entering kindergarten that year. Spanking was measured in the spring by mothers' responses to the question "About how many times, if any, have you spanked [your child] in the past week?" The outcome was children's externalizing behavior as rated by teachers in

the spring of third grade using an adapted version of the Social Skills Rating Scale (Gresham & Elliott, 1990); teachers reported how often the child did such things as argue, fight, or get angry, with response options ranging from *never* (1) to *most of the time* (4). The effect size was therefore longitudinal and used separate raters (mother and teacher) for spanking and the child outcome.

The final study included in the meta-analyses was by McLeod et al. (1994). This study also used the NLSY-C but reported data from a different wave (1988) and for a different child outcome (antisocial behavior); 1,330 White families and 536 Black families were included. Mothers reported how often in the previous week they had spanked their child. They also reported on their child's antisocial behavior using six items from the Behavior Problems Index (Peterson & Zill, 1986). Both measures came from the same wave, and thus the bivariate association between spanking and child antisocial behavior was cross-sectional and the raters for each variable were the same (mothers).

Coding of Effect Sizes

Study-level effect sizes were calculated independently by each of the authors; for all effect sizes, agreement was achieved to at least the third decimal place. When discrepancies occurred in effect-size calculations, the discrepancy was discussed and then each author independently recalculated the effect size. This process was repeated, if necessary, until consensus was achieved. Study-level effect sizes were transformed into standardized mean difference effect sizes to allow combination across effect sizes using Cohen's formula for d (Cohen, 1988; Sterne, 2009):

$$\text{Cohen's } d = \frac{\text{mean}_{\text{treatment}} - \text{mean}_{\text{comparison}}}{sd_{\text{pooled}}}$$

where sd_{pooled} was calculated as

$$sd_{\text{pooled}} = \sqrt{\frac{((n_1 - 1) * sd_1^2) + ((n_2 - 1) * sd_2^2)}{n_1 + n_2 - 2}}$$

Calculation of Cohen's d was uncomplicated when an article reported the sample size, mean, and standard deviation of a group exposed to spanking as well as a group that had never been

spanked. However, some articles did not report effects as comparisons between two groups. Therefore, we made use of formulas found in Borenstein et al. (2009) and Johnson (1993) to convert quantitative measures of association such as correlations and differences of proportions to Cohen's d effect sizes. For each study, we also calculated the standard error of the estimate of Cohen's d utilizing formulas given in Sterne (2009).

Meta-Analytic Procedure

Once all study effect sizes had been converted to the metric of Cohen's d , effect sizes were combined in a meta-analysis. Each study was entered into the model, weighted by its precision ($1/se_d$), and combined into a weighted average of effect. The meta-analysis reported in this article utilized two different meta-analytic strategies. First, a random effects model (Borenstein et al., 2009; DerSimonian & Laird, 1986) using the Stata command *metan* (Bradburn, Deeks, & Altman, 2009) was employed to calculate an overall effect size for the study as well as a separate effect size for each subgroup. The random effects model for meta-analysis does not assume that there is a single underlying effect size of the studies being analyzed; rather, it allows effect sizes to differ across studies to account for the fact that study samples differ by characteristics such as age, gender, race, ethnicity, and nationality. The random effects meta-analytic model calculates the mean of the effect sizes, an estimate of statistical significance of these results, and a measure of the heterogeneity of effect sizes in terms of their variation around the estimated mean effect size. Subsequently, we employed metaregression (Borenstein et al., 2009; Harbord & Higgins, 2009) to examine the degree to which effect sizes for Black families differed from effect sizes for White families.

RESULTS

Given the small number of studies, the meta-analyses necessarily combined effect sizes across child outcomes. One of the studies (Foshee et al., 2005) contributed two effect sizes, and two other studies (Christie-Mizell et al., 2008; McLeod et al., 1994) used the same data set (NLSY-C). Four of the six effect sizes were based on longitudinal data,

namely spanking measured at one time point and the child outcome measured at a later time point.

The results of the meta-analyses by race are presented in Table 2; the effect sizes for each study are presented along with the weighted mean effect sizes across studies overall and the mean effect sizes for the White and Black subsamples. At the level of individual study, five of the six effect sizes for the White subsamples were statistically different from zero, all in the direction of detrimental outcomes. Among the six subsamples of Black families, four of the six effect sizes were statistically different from zero, and all four indicated associations with detrimental outcomes. The weighted mean effect sizes were $d = .274$ for the White subsamples, $d = .300$ for the Black subsamples, and $d = .279$ overall; all three mean effect sizes were statistically different from zero, as indicated by confidence intervals that did not include zero. The weighted mean effect size for the Black subsamples was not statistically different from that for the White subsamples, $t = 0.18, p = .861$.

Because the Deater-Deckard et al. (1996) study included the use of objects in its measure of physical punishment and this has been raised as a concern with previous meta-analyses (Baumrind, Larzelere, & Cowan, 2002), we ran the meta-analyses again with this study excluded to examine differences between the race groups when studies only examined parents' use of spanking (see Table 2). The deleted effect size from Deater-Deckard et al. (1996) for the White subsample was positive and statistically different from zero ($d = .399$), and the deleted effect size for the Black subsample was negative but not statistically different from zero ($d = -.063$). When the Deater-Deckard et al. study was deleted from the analyses, the weighted mean effect sizes overall and for the White subsamples were slightly reduced to $d = .278$ and $d = .270$, respectively, and the weighted mean effect size for the Black subsamples rose slightly to $d = .311$. All three mean effect sizes remained statistically different from zero. The weighted mean effect size for the Black subsamples without Deater-Deckard et al. remained not statistically different from that for the White subsamples, $t = 0.66, p = .526$, which indicates that spanking was similarly associated with detrimental outcomes for both White and Black children.

Table 2. Study-level Effect Sizes and Mean Effect Sizes by White and Black Subsamples

| Study | Child outcome | Race group | Subsample <i>n</i> | <i>d</i> | 95% CI | | Beneficial outcomes | | Detrimental outcomes | |
|---|---------------------------------|------------|--------------------|----------|--------|-------|---------------------|------|----------------------|------|
| | | | | | | | -0.50 | 0.00 | 0.50 | 1.00 |
| Christie-Mizell et al. (2008) | Depressive symptoms | White | 1,139 | 0.181 | 0.064 | 0.297 | | | | |
| | | Black | 713 | 0.242 | 0.094 | 0.390 | | | | |
| Deater-Deckard et al. (1996) | Externalizing behavior problems | White | 372 | 0.399 | 0.192 | 0.607 | | | | |
| | | Black | 88 | -0.063 | -0.483 | 0.358 | | | | |
| Foshee et al. (2005) | Dating violence | White | 958 | -0.060 | -0.187 | 0.067 | | | | |
| | | Black | 188 | 0.242 | -0.047 | 0.531 | | | | |
| Foshee et al. (2005) | Victim of physical abuse | White | 958 | 0.494 | 0.364 | 0.625 | | | | |
| | | Black | 188 | 0.494 | 0.199 | 0.790 | | | | |
| Gershoff et al. (2012) | Externalizing behavior problems | White | 7,057 | 0.283 | 0.236 | 0.330 | | | | |
| | | Black | 1,352 | 0.303 | 0.195 | 0.411 | | | | |
| McLeod et al. (1994) | Antisocial behavior | White | 1,330 | 0.366 | 0.257 | 0.475 | | | | |
| | | Black | 536 | 0.387 | 0.214 | 0.560 | | | | |
| Weighted mean effect sizes, all | | Overall | 14,879 | 0.279 | 0.247 | 0.312 | | | | |
| | | White | 11,814 | 0.274 | 0.237 | 0.310 | | | | |
| | | Black | 3,065 | 0.300 | 0.228 | 0.372 | | | | |
| Weighted mean effect sizes, with Deater-Deckard et al. (1996) removed | | Overall | 14,419 | 0.278 | 0.245 | 0.311 | | | | |
| | | White | 11,442 | 0.270 | 0.233 | 0.307 | | | | |
| | | Black | 2,977 | 0.311 | 0.238 | 0.384 | | | | |

DISCUSSION

The question of whether spanking is linked with fewer negative child outcomes across cultural groups in which the practice is common remains one of the most controversial issues surrounding physical forms of discipline (Durrant, 2008). Following the cultural normativeness perspective, physical punishment should be linked with fewer negative outcomes for children who are from cultural groups in which the use of physical punishment is normative, such as Black families in the United States, than for children from groups in which it is not, such as White families in the United States (Deater-Deckard & Dodge, 1997). The meta-analyses presented in this article tested this argument directly by examining whether spanking is differentially associated with child outcomes in White versus Black families.

Four of the five analyzed studies reported race differences in the frequency with which parents spanked their children (see Table 1); in each case, Black parents spanked their children more often than did White parents. This finding supports the argument that spanking tends to

be more normative among Black families than among White families. Yet, contrary to what was predicted by cultural normativeness theory, spanking was associated with adverse outcomes at similar magnitudes for both Black and White samples of children. Two of the study-level effect sizes for the Black subsamples were not statistically different from zero but four were, and all four of these reported an association between spanking and a detrimental child outcome. Importantly, none of the study-level effect sizes indicated that spanking was associated with beneficial outcomes for the Black subsample.

We found that the strength, direction, and statistical significance of the associations between spanking and child outcomes did not differ for White and Black families. These results suggest that, although there are differences between racial groups in how often parents spank, there are no differences between Black and White families in the way in which spanking is related to child behavioral and mental health problems. In other words, frequency in usage of spanking is a culture-specific behavior, but

there is equifinality in the linkages of spanking with children's outcomes, which indicates culture-common processes (Bornstein, 2012). One of the studies included in the meta-analyses demonstrated this idea directly: In models in which differences in spanking frequency were taken into account, there were no differences across four racial or ethnic groups (White, Black, Hispanic, and Asian) within the United States in the extent to which spanking predicted increases in children's behavior problems over and above initial levels (Gershoff et al., 2012). This finding is consistent with a growing body of research that has found similarities, rather than differences, across racial and ethnic groups in the extent to which spanking is associated with adverse child outcomes (Berlin et al., 2009; Grogan-Kaylor, 2004, 2005; Lau et al., 2006; McLoyd & Smith, 2002; Pardini et al., 2008). While the analyses were limited by the low number of studies reporting relevant data, our findings challenge the argument that the child outcomes associated with spanking depend on the cultural group in which the spanking takes place.

The Deater-Deckard et al. (1996) study was the first to suggest that spanking may be beneficial for Black children and not for White children, and it was the basis for the Deater-Deckard and Dodge article making the case for cultural normativeness as a moderator of parenting impacts on children. Placed alongside several other studies, however, the Deater-Deckard et al. study appears to be an outlier; when it was removed from the analyses, the mean effect size for spanking associated with detrimental outcomes was actually slightly larger for Black children than it was for White children. This appears to be mainly a result of the very large standard error for the effect size for the Black sample (see Table 2), which in turn is likely attributable to the small size of the Black subsample ($n = 88$), which was only a quarter the size of the White subsample ($n = 372$). Of the studies to date that report bivariate effect sizes separately by racial group, there are no other studies that empirically replicate the findings from Deater-Deckard et al.

The notion that families' cultural beliefs or racial group might modify the effectiveness of disciplinary practices has been widely accepted by family researchers and by the public at large. The original studies are highly cited, with Deater-Deckard and Dodge (1997) cited 683 times and Deater-Deckard et al. (1996) cited

778 times per Google Scholar as of mid-2016. It is beyond the scope of this article to speculate why this idea took such firm root in our understanding of family processes, but the results of our analyses suggest that what has become a truism about race and discipline appears not, in fact, to be true.

The analyses presented here have some limitations, the first of which is the small number of applicable studies. Only five studies were found that reported bivariate associations between spanking and child outcomes. Several studies examined moderation by race but did not report the bivariate statistics needed for the meta-analyses (e.g., Berlin et al., 2009). A second limitation is that none of the studies included children younger than age 5; spanking peaks at age 3 (Straus & Stewart, 1999), and so it will be important in future work to examine these association with younger children. A third limitation is the fact that effect sizes for a range of child outcomes were combined together; this was necessary given the low number of studies. In the future, more studies of spanking and child outcomes are needed that report bivariate associations separately by race; both of these limitations could be redressed in a future replication of the analyses presented here.

A fourth limitation is a concern about the research on spanking generally, namely that of uncertain causality; it may be that children with more problematic behaviors elicit more spanking from their parents (Larzelere, Kuhn, & Johnson, 2004). This concern is lessened somewhat in the present analyses by the fact that four of the six effect sizes were from longitudinal studies that measured spanking at one point and child outcome at a later point, thus removing the possibility that the child outcomes predicted the frequency of spanking in these studies. However, there is still the possibility that children's behavior at the first wave may have elicited more spanking at that wave and that this earlier behavior predicts the child's future behavior as well. This possibility of such a third variable effect was tested directly in one of the five studies included in the meta-analyses; using cross-lagged models that estimated the links from early spanking to later externalizing behavior and from early externalizing behavior to later spanking simultaneously, Gershoff et al. (2012) found both paths to be statistically significant and of similar magnitude (age 5 spanking to age 8 externalizing behavior, $\beta = .05$, $p < .001$;

age 5 externalizing behavior to age 8 spanking, $\beta = .06, p < .001$). Thus, although there is evidence that children with more problem behavior elicit more spanking, there is also evidence that spanking is associated with subsequent behavior problems.

A final limitation is one we noted at the outset; namely, these studies used race as a marker of cultural normativeness—they did not incorporate actual racial differences in norms about spanking but rather made the assumption that rates and norms varied by race. There have been few attempts to directly examine cultural norms about spanking as a potential moderator of the association between spanking and child outcomes. Using a sample of mothers and children from six countries, one study did directly assess mothers' beliefs about the normativeness of physical punishment in their communities and found that, although high norms did slightly decrease the association between physical punishment (spanking, slapping, or hitting) and child aggression, more physical punishment was associated with more child aggression for both the high and low normativeness groups (Gershoff et al., 2010). There is likely to be as much variation in spanking attitudes and behaviors within racial groups as there is between them (Kelley et al., 1992), and assumptions should therefore not be made about the normativeness of any parenting practice without measuring it directly.

Implications for Practice and Policy

The finding that spanking is not associated with positive outcomes for Black children or White children, and instead is associated with negative outcomes, indicates that parents of both groups should be advised not to spank their children. Several organizations of professionals who work with children and families have recommended that all parents, regardless of cultural group, reduce or end their use of spanking and that professionals should advise parents on alternatives to spanking (American Academy of Child and Adolescent Psychiatry, 2012; American Academy of Pediatrics, 1998; National Association of Pediatric Nurse Practitioners, 2011). As with all practice with families, these efforts should be culturally sensitive but at the same time empirically informed (Fontes, 2005). Evidence that parent education interventions can successfully reduce physical punishment

and increase sensitive parenting across diverse cultures in the United States (Gross, Garvey, Julion, Fogg, Tucker, & Mokros, 2009) and in countries around the world (Kner, Gardner, & Cluver, 2013; Skar, Sherr, Clucas, & von Tetzchner, 2014) makes clear that such efforts are both feasible and effective. Multitiered interventions to reduce the use of spanking and physical punishment that include public education campaigns, such as the Triple P parenting intervention (see Prinz, Sanders, Shapiro, Whitaker, & Lutzker, 2009), are promising strategies for reducing spanking and increasing positive and effective parenting in communities and cultural groups across the United States.

Spanking and physical punishment by parents remain legal throughout the United States, but they are not legal in all countries. A total of 49 countries have banned all physical punishment of children (Global Initiative to End Corporal Punishment of Children, 2016) as a reaction to mounting evidence that spanking is harmful to children and to declarations by the United Nations that spanking violates children's human rights (UN Committee on the Rights of the Child, 2007). Given the continued strong support for spanking among the American public (Child Trends Databank, 2015) and the United States' status as the only country not to ratify the Convention on the Rights of the Child (United Nations, 2015), a legal ban is not a likely means of reducing physical punishment in the United States.

CONCLUSION

These meta-analyses provide evidence that, contrary to the cultural normativeness perspective, spanking is not associated with beneficial outcomes for Black children and rather is associated with detrimental outcomes for Black children to the same extent as it is for White children. This finding is consistent with a growing body of research that has found racial and ethnic similarities, rather than differences, in the extent to which spanking is associated with adverse child outcomes (Berlin et al., 2009; Grogan-Kaylor 2004, 2005; Lau et al., 2006; McLoyd & Smith, 2002; Pardini et al., 2008). Although the analyses were limited by the low number of studies reporting relevant data, our findings challenge the argument that the child outcomes associated with spanking depend on the cultural group in which the spanking takes place.

REFERENCES

Studies used in the meta-analyses are preceded by an asterisk.

- Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL 4–18, YSR, and TRF profiles*. Burlington, VT: Department of Psychiatry, University of Vermont.
- American Academy of Child and Adolescent Psychiatry. (2012, July 30). *Policy statement on corporal punishment*. Retrieved from https://www.aacap.org/aacap/Policy_Statements/2012/Policy_Statement_on_Corporal_Punishment.aspx
- American Academy of Pediatrics, Committee on Psychosocial Aspects of Child and Family Health. (1998). Guidance for effective discipline. *Pediatrics*, *101*, 723–728.
- Baumrind, D., Larzelere, R. E., & Cowan, P. A. (2002). Ordinary physical punishment: Is it harmful? Comment on Gershoff (2002). *Psychological Bulletin*, *128*, 580–589. doi: 10.1037/0033-2909.128.4.580
- Berlin, L. J., Ispa, J. M., Fine, M. A., Malone, P. S., Brooks-Gunn, J., Brady-Smith, C., ... Bai, Y. (2009). Correlates and consequences of spanking and verbal punishment for low-income White, African American, and Mexican American toddlers. *Child Development*, *80*, 1403–1420. doi: 10.1111/j.1467-8624.2009.01341.x
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to meta-analysis*. West Sussex, UK: Wiley.
- Bornstein, M. H. (2012). Cultural approaches to parenting. *Parenting: Science and Practice*, *12*, 212–221. doi: 10.1080/15295192.2012.683359
- Bornstein, M. H., Putnick, D. L., Suwalsky, J. T. D., Venuti, P., de Falco, S., de Galperín, C. Z., ... Tichovolsky, M. H. (2012). Emotional relationships in mothers and infants: Culture-common and community-specific characteristics of dyads from rural and metropolitan settings in Argentina, Italy, and the United States. *Journal of Cross Cultural Psychology*, *43*, 171–197. doi: 10.1177/00220221110388563
- Bradburn, M. J., Deeks, J. J., & Altman, D. G. (2009). Metan: A command for meta-analysis in Stata. In J. A. C. Sterne (Ed.), *Meta-analysis in Stata: An updated collection from the Stata Journal* (pp. 3–28). College Station, TX: Stata Press.
- Bronfenbrenner, U., & Morris, P. (1998). The ecology of developmental processes. In R. M. Lerner (Ed.), *Handbook of child psychology. Vol. 1: Theoretical models of human development* (5th ed., pp. 993–1028). New York, NY: Wiley.
- Child Trends Databank. (2015). *Attitudes toward spanking*. Retrieved from <http://www.childtrends.org/?indicators=attitudes-toward-spanking>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- *Christie-Mizell, C. A., Pryor, E. M., & Grossman, E. R. B. (2008). Child depressive symptoms, spanking and emotional support: Differences between African American and European American youth. *Family Relations*, *57*, 335–350. doi: 10.1111/j.1741-3729.2008.00504.x
- Dasen, P. R., & Mishra, R. C. (2000). Cross-cultural views on human development in the third millennium. *International Journal of Behavioral Development*, *24*, 428–434. doi: 10.1080/016502500750037973
- Deater-Deckard, K., & Dodge, K. A. (1997). Externalizing behavior problems and discipline revisited: Nonlinear effects and variation by culture, context, and gender. *Psychological Inquiry*, *8*, 161–175. doi: 10.1207/s15327965pli0803_1
- *Deater-Deckard, D., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1996). Physical discipline among African American and European American mothers: Links to children's externalizing behaviors. *Developmental Psychology*, *32*, 1065–1072. doi: 10.1037/0012-1649.32.6.1065
- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*, 177–188. doi: 10.1016/0197-2456(86)90046-2
- Durrant, J. E. (2008). Physical punishment, culture and rights: Current issues for professionals. *Journal of Developmental and Behavioral Pediatrics*, *29*, 55–66. doi: 10.1097/dbp.0b013e318135448a
- Flynn, C. P. (1998). To spank or not to spank: The effect of situation and age of child on support for corporal punishment. *Journal of Family Violence*, *13*, 21–37. doi: 10.1023/A:1022808716048
- Fontes, L. A. (2005). *Child abuse and culture: Working with diverse families*. New York, NY: Guilford.
- *Foshee, V. A., Ennett, S. T., Bauman, K. E., Benefield, T., & Suchindran, C. (2005). The association between family violence and adolescent dating violence onset: Does it vary by race, socioeconomic status, and family structure? *Journal of Early Adolescence*, *25*, 317–344. doi: 10.1177/0272431605277307
- Gershoff, E. T., Grogan-Kaylor, A., Lansford, J. E., Chang, L., Zelli, A., Deater-Deckard, K., & Dodge, K. A. (2010). Parent discipline practices in an international sample: Associations with child behaviors and moderation by perceived normativeness. *Child Development*, *81*, 487–502. doi: 10.1111/j.1467-8624.2009.01409.x
- *Gershoff, E. T., Lansford, J. E., Sexton, H. R., Davis-Kean, P. E., & Sameroff, A. J. (2012). Longitudinal links between spanking and children's externalizing behaviors in a national sample of White, Black, Hispanic, and Asian American Families. *Child Development*, *83*, 838–843. doi: 10.1111/j.1467-8624.2011.01732.x
- Global Initiative to End All Corporal Punishment of Children. (2016). *States which have prohibited all corporal punishment*. Retrieved from

- <http://www.endcorporalpunishment.org/progress/prohibiting-states>
- Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system manual*. Circle Pines, MN: American Guidance Service.
- Grogan-Kaylor, A. (2004). The effect of corporal punishment on antisocial behavior in children. *Social Work Research, 28*, 153–162. doi: 10.1093/swr/28.3.153
- Grogan-Kaylor, A. (2005). Corporal punishment and the growth trajectory of children's antisocial behavior. *Child Maltreatment, 10*, 283–292. doi: 10.1177/1077559505277803
- Grogan-Kaylor, A., & Otis, M. D. (2007). The predictors of parental use of corporal punishment. *Family Relations, 56*, 80–91. doi: 10.1111/j.1741-3729.2007.00441.x
- Gross, D., Garvey, C., Julion, W., Fogg, L., Tucker, S., & Mokros, H. (2009). Efficacy of the Chicago Parent Program with low-income African American and Latino parents of young children. *Prevention Science, 10*, 54–65.
- Gunnoe, M. L., & Mariner, C. L. (1997). Toward a developmental-contextual model of the effects of parental spanking on children's aggression. *Archives of Pediatric and Adolescent Medicine, 151*, 768–775. doi: 10.1001/archpedi.1997.02170450018003
- Hall, G. C., Yip, T., & Zárate, M. A. (2016). On becoming multicultural in a monocultural research world: A conceptual approach to studying ethnocultural diversity. *American Psychologist, 71*, 40–51. doi: 10.1037/a0039734
- Harbord, R. M., & Higgins, J. P. T. (2009). Meta-regression in Stata. In J. A. C. Sterne (Ed.), *Meta-analysis in Stata: An updated collection from the Stata Journal* (pp. 70–96). College Station, TX: Stata Press.
- Horn, I. B., Joseph, J. G., & Cheng, T. L. (2004). Nonabusive physical punishment and child behavior among African-American children: A systematic review. *Journal of the National Medical Association, 96*, 1162–1168.
- Johnson, B. T. (1993). *DSTAT 1.10: Software for the meta-analytic review of research literatures*. Hillsdale, NJ: Erlbaum.
- Kelley, M. L., Power, T. G., & Wimbush, D. D. (1992). Determinants of disciplinary practices in low-income Black mothers. *Child Development, 63*, 573–582. doi: 10.2307/1131347
- Knerr, W., Gardner, F., & Cluver, L. (2013). Improving positive parenting skills and reducing harsh and abusive parenting in low- and middle-income countries: A systematic review. *Prevention Science, 14*, 352–363. doi: 10.1007/s11121-012-0314-1
- Lansford, J. E., Chang, L., Dodge, K. A., Malone, P. S., Oburu, P., Palmérus, K., ... Quinn, N. (2005). Cultural normativeness as a moderator of the link between physical discipline and children's adjustment: A comparison of China, India, Italy, Kenya, Philippines, and Thailand. *Child Development, 76*, 1234–1246. doi: 10.1111/j.1467-8624.2005.00847.x
- Lansford, J. E., Deater-Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (2004). Ethnic differences in the link between physical discipline and later adolescent externalizing behaviors. *Journal of Child Psychology and Psychiatry, 45*, 801–812. doi: 10.1111/j.1469-7610.2004.00273.x
- Lansford, J. E., Sharma, C., Malone, P. S., Woodlief, D., Dodge, K. A., Oburu, P., ... Di Giunta, L. (2014). Corporal punishment, maternal warmth, and child adjustment: A longitudinal study in eight countries. *Journal of Clinical Child and Adolescent Psychology, 43*, 670–685. doi: 10.1080/15374416.2014.893518
- Lansford, J. E., Wager, L. B., Bates, J. E., Pettit, G. S., & Dodge, K. A. (2012). Forms of spanking and children's externalizing behaviors. *Family Relations, 61*, 224–236. doi: 10.1111/j.1741-3729.2011.00700.x
- Larzelere, R. E., Kuhn, B. R., & Johnson, B. (2004). The intervention selection bias: An underrecognized confound in intervention research. *Psychological Bulletin, 130*, 289–303. doi: 10.1037/0033-2909.130.2.289
- Lau, A. S., Litrownik, A. J., Newton, R. R., Black, M. M., & Everson, M. D. (2006). Factors affecting the link between physical discipline and child externalizing problems in Black and White families. *Journal of Community Psychology, 34*, 89–103. doi: 10.1002/jcop.20085
- *McLeod, J. D., Kruttschnitt, C., & Dornfeld, M. (1994). Does parenting explain the effects of structural conditions on children's antisocial behavior? A comparison of Blacks and Whites. *Social Forces, 73*, 575–604. doi: 10.2307/2579822
- McLoyd, V. C., & Smith, J. (2002). Physical discipline and behavior problems in African American, European American, and Hispanic children: Emotional support as a moderator. *Journal of Marriage and the Family, 64*, 40–53. doi: 10.1111/j.1741-3737.2002.00040.x
- Mistry, J., Chaudhuri, J. H., & Diez, V. (2003). Ethnotheories of parenting: At the interface between culture and child development. In R. M. Lerner, D. Wertlieb, & F. Jacobs (Eds.), *Handbook of applied developmental science. Vol. 1: Applying developmental science for youth and families: Historical and theoretical foundations* (pp. 233–257). Thousand Oaks, CA: Sage.
- Mosby, L., Rawls, A. W., Meehan, A. J., Mays, E., & Pettinari, C. J. (1999). Troubles in interracial talk about discipline: An examination of African American child rearing narratives. *Journal of Comparative Family Studies, 30*, 489–521.

- National Association of Pediatric Nurse Practitioners. (2011). NAPNAP position statement on corporal punishment. *Journal of Pediatric Health Care*, 25, e31–e32. Retrieved from <http://download.journals.elsevierhealth.com/pdfs/journals/0891-5245/PIIS0891524511002288.pdf>
- Pardini, D. A., Fite, P. J., & Burke, J. D. (2008). Bidirectional associations between parenting practices and conduct problems in boys from childhood to adolescence: The moderating effect of age and African-American ethnicity. *Journal of Abnormal Child Psychology*, 36, 647–662. doi: 10.1007/s10802-007-9162-z
- Peterson, J. L., & Zill, N. (1986). Marital disruption, parent–child relationships, and behavior problems in children. *Journal of Marriage and Family*, 48, 295–307. doi: 10.2307/352397
- Prinz, R. J., Sanders, M. R., Shapiro, C. J., Whitaker, D. J., & Lutzker, J. R. (2009). Population-based prevention of child maltreatment: The US Triple P System population trial. *Prevention Science*, 10, 1–12. doi: 10.1007/s11121-009-0123-3
- Rogoff, B. (2003). *The cultural nature of human development*. New York, NY: Oxford University Press.
- Runyan, D. K., Shankar, V., Hassan, F., Hunter, W. M., Jain, D., Paula, C. S., ... Bordin, I. A. (2010). International variations in harsh child discipline. *Pediatrics*, 126, e701–e711. doi: 10.1542/peds.2008-2374
- Skar, A. M. S., Sherr, L., Clucas, C., & von Tetzchner, S. (2014). Evaluation of follow-up effects of the International Child Development Programme on caregivers in Mozambique. *Infants & Young Children*, 27, 120–135. doi: 0.1097/IYC.0000000000000006
- Slade, E. P., & Wissow, L. S. (2004). Spanking in early childhood and later behavior problems: A prospective study of infants and young toddlers. *Pediatrics*, 113, 1321–1330. doi: 10.1542/peds.113.5.1321
- Smith, T. W., Marsden, P., Hout, M., & Kim, J. (2016). *General social surveys, 1972–2014* [Machine-readable data file]. Retrieved from <http://gss.norc.org>
- Sterne, J. A. C. (Ed.). (2009). *Meta-analysis in Stata: An updated collection from the Stata Journal*. College Station, TX: Stata Press.
- Straus, M. A. (1979). Measuring intrafamily conflict and violence: The Conflict Tactics Scale. *Journal of Marriage and Family*, 41, 75–85. doi: 10.2307/351733
- Straus, M. A., & Stewart, J. H. (1999). Corporal punishment by American parents: National data on prevalence, chronicity, severity, and duration, in relation to child and family characteristics. *Clinical Child and Family Psychology Review*, 2, 55–70. doi: 10.1023/A:1021891529770
- Super, C. M., & Harkness, S. (1986). The developmental niche: A conceptualization at the interface of child and culture. *International Journal of Behavioural Development*, 9, 545–569. doi: 10.1177/016502548600900409
- Umaña-Taylor, A. J., Quintana, S. M., Lee, R. M., Cross, W. E., Rivas-Drake, D., Schwartz, S. J., ... Ethnic and Racial Identity in the 21st Century Study Group. (2014). Ethnic and racial identity during adolescence and into young adulthood: An integrated conceptualization. *Child Development*, 85, 21–39. doi: 10.1111/cdev.12196
- UN Committee on the Rights of the Child. (2007, March 2). *CRC general comment no. 8 (2006): The right of the child to protection from corporal punishment and other cruel or degrading forms of punishment* (U.N. Doc. No. CRC/C/GC/8). Retrieved from <http://www.refworld.org/docid/460bc7772.html>
- United Nations. (2015, January 20). *UN lauds Somalia as country ratifies landmark children's rights treaty*. Retrieved from <http://www.un.org/apps/news/story.asp?NewsID=49845#.VjNnrerQsP>