

ARE SPANKING INJUNCTIONS SCIENTIFICALLY SUPPORTED?

ROBERT E. LARZELERE*

DIANA BAUMRIND**

I

INTRODUCTION

This special issue on corporal punishment addresses arguments for and against prohibitions of the historically widespread practice of disciplinary spanking by parents. A recent national survey estimated that ninety-four percent of American parents of four- and five-year-olds spanked their children at least occasionally.¹ Yet there is a growing trend for countries to ban corporal punishment by parents through family law or criminal law.² This article evaluates whether the current empirical evidence supports spanking prohibitions.

Does the scientific evidence show that spanking is invariably detrimental regardless of how it is used? Or can parents use spanking in nonharmful or beneficial ways, at least under some conditions? Should all corporal punishment be enjoined, or should a legal distinction be retained between spanking and physical abuse? These crucial questions compare the validity of two scientific perspectives, “anticorporal punishment” and “conditional corporal punishment,”³ both of which are represented in this issue.⁴ In this article, we will

Copyright © 2010 by Robert E. Larzelere and Diana Baumrind.

This article is also available at <http://www.law.duke.edu/journals/lcp>.

* Department of Human Development and Family Science, Oklahoma State University.

** Institute of Human Development, University of California, Berkeley.

1. Murray A. Straus & Julie H. Stewart, *Corporal Punishment by American Parents: National Data on Prevalence, Chronicity, Severity, and Duration, in Relation to Child and Family Characteristics*, 2 CLINICAL CHILD & FAM. PSYCHOL. REV. 55, 59–60 (1999).

2. See *Legal Reforms*, CENTER FOR EFFECTIVE DISCIPLINE, <http://www.stophitting.com/index.php?page=laws-main> (last visited January 6, 2010).

3. Corina Benjet & Alan E. Kazdin, *Spanking Children: The Controversies, Findings, and New Directions*, 23 CLINICAL PSYCHOL. REV. 197, 200–01 (2003). This article compared two recent literature reviews representing those two perspectives, one by Elizabeth Thompson Gershoff, *Corporal Punishment by Parents and Associated Child Behaviors and Experiences: A Meta-Analytic and Theoretical Review*, 128 PSYCHOL. BULL. 539 (2002), and the other by Robert E. Larzelere, *Child Outcomes of Nonabusive and Customary Physical Punishment by Parents: An Updated Literature Review*, 3 CLINICAL CHILD & FAM. PSYCHOL. REV. 199 (2000).

4. Elizabeth T. Gershoff, *More Harm Than Good: A Summary of Scientific Research on the Intended and Unintended Effects of Corporal Punishment on Children*, 73 LAW & CONTEMP. PROBS. 57 (Spring 2010).

use the terms *spanking prohibition* and *conditional spanking* to differentiate these two positions. The conditional-spanking viewpoint holds that spanking may be an appropriate disciplinary option under some conditions but not others. The conditions under which spanking may be a viable disciplinary method need to be investigated before applying a blanket prohibition. Because advocates of both positions are opposed to overly severe and abusive corporal punishment,⁵ evidence about the effects of excessively severe punishment does not differentiate the two positions and is not directly relevant to the desirability of spanking prohibitions. Evidence about using corporal punishment too severely would be indirectly relevant, however, if it could be shown that a spanking prohibition and conditional spanking differ in their abilities to prevent disciplinary actions from escalating to physical abuse, an issue addressed in section V.

Spanking-prohibition and conditional-spanking positions differ, too, on whether the use of disciplinary spanking is always or generally harmful in a cost-benefit analysis. The spanking-prohibition viewpoint necessarily implies that any nonharmful or beneficial subset of parental corporal punishment is so small a proportion or so minor in its benefits that it is outweighed by the detrimental effects of retaining any spanking option for parents.

To justify removing this option from parents, spanking prohibitionists first need to show causal evidence that spanking is detrimental in situations where it is considered most appropriate by parents, children, and psychologists.⁶ Second, prohibitionists need to compare the effects of spanking with the effects of alternative disciplinary tactics available to parents in the same disciplinary situations. Third, prohibitionists need evidence that parenting improves when parents are prevented from using disciplinary spanking. Fourth, prohibitionists need to show that adverse outcomes associated with spanking remain associated with spanking after eliminating the influences of several prevalent confounding variables, such as difficult child temperaments and socioeconomic disadvantages. If these confounding factors together account for the associations between spanking and adverse outcomes, those associations would be spurious and therefore misinterpreted as causal influences of spanking.

It is well known that children thrive under authoritative parenting,⁷ recently confirmed by ten-year outcomes from Baumrind's classic longitudinal data.⁸ *Authoritative* parenting combines nurturance, give-and-take communication,

5. Benjet & Kazdin, *supra* note 3, at 202.

6. Thomas F. Catron & John C. Masters, *Mothers' and Children's Conceptualizations of Corporal Punishment*, 64 CHILD DEV. 1815, 1819 (1993); Mark W. Roberts & Scott W. Powers, *Adjusting Chair Timeout Enforcement Procedures for Oppositional Children*, 21 BEHAV. THERAPY 257, 262 (1990).

7. Ross D. Parke & Raymond Buriel, *Socialization in the Family: Ethnic and Ecological Perspectives*, in 3 HANDBOOK OF CHILD PSYCHOLOGY: SOCIAL, EMOTIONAL, AND PERSONALITY DEVELOPMENT 429, 436-37 (W. Damon et al. eds., 2006); Laurence Steinberg, *We Know Some Things: Parent-Adolescent Relationships in Retrospect and Prospect*, 11 J. RES. ON ADOLESCENCE 1, 1 (2001).

8. Diana Baumrind et al., *Effects of Preschool Parents' Power Assertive Patterns and Practices on Adolescent Development*, 10 PARENTING: SCI. & PRAC. 167 (2010).

and support for age-appropriate independence with firm, confrontive discipline and maturity demands. It is critically distinguishable from *authoritarian* parenting, which is equally firm but shares none of the other aspects of authoritative parenting and is characterized instead by the use of hostile verbal discipline and severe corporal punishment. On average, authoritative parents spanked just as much as the average of all other parents.⁹ Undoubtedly, some parents can be authoritative without using spanking, but we have no evidence that all or even most parents can achieve authoritative parenting without an occasional spank. A crucial question is whether spanking prohibitions would undermine authoritative parenting for some parents. Would parents then use nonphysical disciplinary tactics more effectively than if they retained the spanking option? Or would parents enjoined from spanking become like *authoritarian* parents in using more verbal hostility, which is more detrimental than spanking,¹⁰ or, like *permissive* parents, become less able to enforce appropriate child cooperation? *Permissive* parents are nurturant and support age-appropriate independence but tend to avoid disciplinary confrontations and make few demands for mature responsibility and cooperation. Children of authoritarian parents were far less competent ten years later compared to children of authoritative parents, whereas children of permissive parents were the second least-competent group.¹¹

This article summarizes the scientific evidence on child outcomes of spanking, emphasizing causal evidence under conditions considered most appropriate for its use by parents and psychologists. Section II discusses the distinctions that must be made to answer crucial questions about the effects of spanking under those conditions. With those distinctions in mind, section III then contrasts two major literature reviews on corporal punishment. Section IV starts by summarizing the few definitive studies that have made all of the necessary distinctions outlined in section II and then summarizes and critiques the strongest causal evidence against ordinary spanking. Section V addresses other empirical issues, including the role of spanking in escalations of disciplinary actions toward abuse, the aversiveness of spanking compared to alternatives, and ethnic differences in the apparent outcomes of spanking. Section VI presents our conclusions.

II

NECESSARY DISTINCTIONS

Scientific support for injunctions against parents' use of disciplinary spanking must document that nonabusive spanking is harmful or ineffective when parents perceive the greatest need to use it—for example, when young children are persistently defiant even after parents try other disciplinary actions.

9. *Id.* at 179, 187.

10. *Id.* at 157, 178–83.

11. *Id.* at 157, 172–76, 84.

To be relevant for spanking prohibitions, empirical evidence must come from studies that discriminate three crucial issues correctly: (1) Corporal punishment must be implemented nonabusively (correct dosage), (2) it must be used in an appropriate disciplinary situation (appropriate presenting problems), and (3) the evidence must be causal, not correlational. Prohibitions of corrective medical actions would not be considered unless evidence came from studies making all three distinctions correctly. For example, a prohibition against radiation treatment would first need to show causal evidence of harm from appropriate dosages administered for appropriate presenting problems.

A. Appropriate Dosage

Any prohibition against spanking must likewise rely on evidence from nonabusive implementation, rather than evidence based on lumping spanking together with overly severe corporal punishment. The 1996 scientific-consensus conference on The Short- and Long-Term Consequences of Corporal Punishment defined corporal punishment as “bodily punishment of any kind as a form of discipline”¹² and spanking as a kind of corporal punishment that is “a. physically non-injurious; b. intended to modify behavior; and c. administered with an opened hand to the extremities or buttocks.”¹³ This article adopts this definition of spanking, which limits it to nonabusive usage, differentiated from severe corporal punishment.

To justify spanking prohibitions, research must first show the detrimental causal effects of spanking and then that less extreme injunctions cannot minimize those detrimental effects. Moreover, a cost-benefit analysis must weigh any unavoidable detrimental effects against any beneficial effects found for spanking. Radiation therapy has negative side effects, but it would not be prohibited unless studies showed that alternative treatments produced consistently better outcomes without increasing negative side effects, based on causal evidence of appropriate applications of radiation. In this article we will show that corporal punishment is associated with more-adverse outcomes than alternative disciplinary tactics only for severe and predominant use of corporal punishment.

B. Appropriate Presenting Problems

Just as radiation treatments are evaluated for specific types of cancer, spanking needs to be evaluated for its most appropriate presenting problems. Two kinds of evidence are relevant: situations in which parents are most likely to spank and situations in which psychologists have trained parents of young children when to spank appropriately. First, parents are most likely to spank

12. Stanford B. Friedman & S. Kenneth Schonberg, *Consensus Statements*, 98 PEDIATRICS 853, 853 (1996).

13. *Id.*

one- to nine-year-olds¹⁴ for defiance,¹⁵ especially when their misbehavior hurts someone else¹⁶ or puts the children themselves in danger.¹⁷ Second, from the late 1960s¹⁸ until the middle 1990s, clinical psychologists trained parents to use spanking¹⁹ to enforce time-out compliance in behavioral parent training programs for young children with disruptive-behavior diagnoses.²⁰ Two prominent practitioners in that period explained, “While we basically are opposed to physical punishment, we have found a mild spanking to be the most feasible backup for the child leaving the [time-out] chair.”²¹ These behavioral parent training programs—which feature time-out currently enforced with an alternative back-up—are currently recognized as some of the most effective treatments for young children with attention deficit–hyperactivity disorder (ADHD),²² oppositional defiant disorder, and conduct disorder.²³ We will show that the strongest causal evidence about spanking indicates that it is effective for reducing defiance in the most defiant two- to six-year-olds, which is crucial for their cooperation with time-out. The few other studies that focus on defiance or dangerous behaviors also document better outcomes for spanking than for most alternative tactics when dealing with defiance in young children.

C. Causal vs. Correlational Evidence

Corporal punishment is usually correlated with behavior problems such as antisocial behavior and aggression.²⁴ But correlation does not equal causation. What is open to dispute are the causal influences that explain those correlations. Making valid causal conclusions from correlations involving corrective actions is especially problematic, for correlations are biased against

14. Straus & Stewart, *supra* note 1, at 59.

15. Kathy L. Ritchie, *Maternal Behaviors and Cognitions During Discipline Episodes: A Comparison of Power Bouts and Single Acts of Noncompliance*, 35 DEVELOPMENTAL PSYCHOL. 580, 587 (1999).

16. George W. Holden et al., *Why 3-Year-Old Children Get Spanked—Parent and Child Determinants as Reported by College-Educated Mothers*, 41 MERRILL-PALMER Q.: J. DEVELOPMENTAL PSYCHOL. 431, 441 (1995).

17. Catron & Masters, *supra* note 6, at 1817, 1819–20; Ritchie, *supra* note 15, at 587.

18. TONI L. HEMBREE-KIGIN & CHERYL B. MCNEIL, PARENT–CHILD INTERACTION THERAPY 2 (1995).

19. *Id.* at 94–95; RUSSELL A. BARKLEY, DEFiant CHILDREN: A CLINICIAN’S MANUAL FOR PARENT TRAINING 117–18 (1987); EDWARD R. CHRISTOPHERSEN, LITTLE PEOPLE: GUIDELINES FOR COMMON SENSE CHILD REARING 151 (3d ed. 1988); REX L. FOREHAND & ROBERT J. MCMAHON, HELPING THE NONCOMPLIANT CHILD 79–80 (1981).

20. BARKLEY, *supra* note 19, at 2–3; CHRISTOPHERSEN, *supra* note 19, at 50, 152; FOREHAND & MCMAHON, *supra* note 19, at ix; HEMBREE-KIGIN & MCNEIL, *supra* note 18, at 7.

21. FOREHAND & MCMAHON, *supra* note 19, at 80.

22. William E. Pelham, Jr. & Gregory A. Fabiano, *Evidence-Based Psychosocial Treatments for Attention-Deficit/Hyperactivity Disorder*, 37 J. CLINICAL CHILD & ADOLESCENT PSYCHOL. 184, 187 (2008).

23. Sheila M. Eyberg et al., *Evidence-Based Psychosocial Treatments for Children and Adolescents with Disruptive Behavior*, 37 J. CLINICAL CHILD & ADOLESCENT PSYCHOL. 215, 226–29 (2008).

24. This is so both cross-sectionally and longitudinally. Gershoff, *supra* note 3, at 539.

corrective actions, a problem known as the intervention selection bias.²⁵ This selection bias occurs because of the poorer prognosis of those selected for the corrective action compared to the better prognosis of those not needing the corrective action.

To illustrate this bias, Table 1 summarizes data from the largest study of radiation treatment for Stage II endometrium cancer prior to 1992.²⁶ Women who received radiation treatment had a higher probability of dying in the next five years than did women of the same age in the general population. Either type of radiation treatment was therefore correlated with a higher rate of dying compared to women of the same age who did not receive radiation treatment. Using the program employed by Gershoff²⁷ to calculate effect-size statistics, this translates to large detrimental effect sizes of $d = .61$ and $d = 1.80$ for the two treatments.²⁸ By comparison, Gershoff's effect sizes for adverse outcomes of physical punishment ranged from $d = .09$ to $d = .69$,²⁹ which therefore appear less adverse than the radiation treatments in Table 1 when their effect size statistics are based inappropriately on unadjusted correlations. This shows that effect sizes based on longitudinal correlations are biased against all corrective actions because the comparison group includes many who did not need any corrective action. In the same way that having cancer causes women to be more likely to receive radiation treatment, children's oppositional behavior causes parents to be more likely to use all disciplinary tactics more frequently, not just spanking. Therefore the frequencies of all disciplinary tactics are correlated with more disruptive-behavior problems twenty months later, an association not distinctive of spanking.³⁰

25. Robert E. Larzelere et al., *The Intervention Selection Bias: An Underrecognized Confound in Intervention Research*, 130 PSYCHOL. BULL. 289, 289 (2004).

26. Perry W. Grigsby et al., *Stage II Carcinoma of the Endometrium: Results of Therapy and Prognostic Factors*, 11 INT'L J. RADIATION ONCOLOGY BIOLOGY PHYSICS 1915, 1918 (1985).

27. Gershoff, *supra* note 3, at 544.

28. Effect sizes in d estimate the difference a treatment is expected to make in terms of standard deviations of the outcome variable. In the social sciences, an effect size of $d = .20$ is considered small, $d = .50$ medium, and $d = .80$ large. JACOB COHEN, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES 25–26 (2d ed. 1988).

29. Gershoff, *supra* note 3, at 547.

30. Larzelere, *supra* note 25, at 290–91; Robert E. Larzelere et al., *Punishment Enhances Reasoning's Effectiveness as a Disciplinary Response to Toddlers*, 60 J. MARRIAGE & FAM. 388, 400 (1998).

Table 1. Five-Year Survival Rates for Stage II Endometrial Carcinoma^a

Treatment (or Comparison Condition)	5-year Survival Rate	5-year Death Rate	Equivalent Effect Sizes ^b	
			<i>r</i>	<i>d</i>
Radiation plus surgery	78%	22%	.29	.61
Radiation alone	48%	52%	.67	1.81
Actuarial survival, 65-year-old women ^c	93%	7%	--	--

^aBased on Grigsby et al.,³¹ the largest study of Stage II Endometrial Cancer in Glassburn et al.³²

^bCompared to actuarial survival in 65-year-old American women, using the statistical program³³ used in Gershoff's³⁴ meta-analysis to calculate effect sizes.

^cBased on 65-year-old American women, just above the median age of patients in Grigsby et al.³⁵

In summarizing the relevant empirical literature, then, it is important to distinguish correlational evidence from causally definitive and causally relevant results. *Causally definitive* results are those based on the kinds of randomized clinical trials that are widely recognized in science as providing more conclusive causal evidence than any other research strategy.³⁶ Because findings from randomized clinical trials are recognized as causally definitive, they are required by the Federal Drug Administration for new prescription drugs and by the Society of Clinical Child and Adolescent Psychology³⁷ to identify “evidence-based” psychosocial treatments for children and adolescents.³⁸ *Causally relevant* refers to studies that provide stronger causal evidence than unadjusted correlations yet do not use the randomization methods required for causally definitive conclusions. The most relevant example involves studies in which corporal punishment predicts a subsequent child outcome even after adjusting statistically for preexisting differences in that outcome, a research strategy

31. Grigsby et al., *supra* note 26, at 1918.

32. J. R. Glassburn et al., *Endometrium*, in PRINCIPLES AND PRACTICE OF RADIATION ONCOLOGY 1203, 1214-15 (C. A. Perez & L. W. Brady eds., 1992).

33. BLAIR T. JOHNSON, DSTAT: SOFTWARE FOR THE META-ANALYTIC REVIEW OF RESEARCH LITERATURE (1989).

34. Gershoff, *supra* note 3, at 544.

35. Grigsby et al., *supra* note 26, at 1916.

36. WILLIAM R. SHADISH, THOMAS D. COOK, & DONALD T. CAMPBELL, EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS FOR GENERALIZED CAUSAL INFERENCE 13 (2002).

37. The Society of Clinical Child and Adolescent Psychology is the fifty-third division of the American Psychological Association. See *Divisions*, AM. PSYCHOL. ASS'N, (Sept. 6, 2010), <http://www.apa.org/about/division/index.aspx>.

38. Wendy K. Silverman & Stephen P. Hinshaw, *The Second Special Issue on Evidence-Based Psychosocial Treatments for Children and Adolescents*, 37 J. CLINICAL CHILD & ADOLESCENT PSYCHOL. 1, 5 (2008).

known as *net-effects regression*.³⁹ Such statistical adjustments yield unbiased estimates of causal effects only when the process of selecting recipients for a corrective action is measured comprehensively⁴⁰ and without measurement error.⁴¹ Accordingly, epidemiologists recognize that residual confounding remains when confounds are only partially controlled for statistically.⁴² We will show that residual confounding can easily account for the strongest causal evidence against spanking.

D. Other Methodological Issues

Three other pervasive methodological issues warrant brief mention: (1) same-source bias, (2) confounding ineffectiveness with spanking frequency, and (3) the fact that very few studies have compared spanked children with never-spanked children. First, *same-source bias* occurs when the same person (for example, a parent) is the source of information for measures of spanking and of the child outcome. A mother who just told an interviewer that she spans her son frequently might try to justify it later in the interview by exaggerating her son's belligerence. Accordingly, same-source bias is known to artificially increase correlations of disciplinary tactics with adverse child outcomes, such as aggression.⁴³ Second, the more effectively any disciplinary tactic is used, the less need there will be to use it in the future. Therefore, frequency of a disciplinary tactic is partly due to *how ineffectively* a parent has used it previously, with more-effective implementations resulting in lower frequencies. This bias can make any disciplinary tactic appear to be more detrimental than it is, when based on measures of frequency of use.

Third, very few studies have actually compared a spanked group to a never-spanked one; most studies contrast frequent spanking with infrequent spanking. For example, the statistically controlled studies with the strongest causal evidence against customary spanking are all based on spanking frequency in the past week. Parents who spanked less than every other week would most likely be lumped together with never-spankers in the no-spanking group for that particular week.

39. A relevant example of *net effects regression* is the association between spanking at an initial time and aggression a year later, after removing what can be predicted about that aggression a year later from initial levels of aggression (that is, net of what can be predicted from initial aggression levels). Stephen Turner, "Net Effects": A Short History, in CAUSALITY IN CRISIS? STATISTICAL METHODS AND THE SEARCH FOR CAUSAL KNOWLEDGE IN THE SOCIAL SCIENCES 23–25 (Vaughn R. McKim & Stephen P. Turner eds., 1997).

40. James J. Heckman, *Sample Selection Bias as a Specification Error*, 47 ECONOMETRICA 153, 160 (1979).

41. DONALD T. CAMPBELL & DAVID A. KENNY, A PRIMER ON REGRESSION ARTIFACTS 155 (1999); David A. Freedman, *Statistical Models and Shoe Leather*, 21 SOC. METHODOLOGY 291, 302–04 (1991).

42. KENNETH J. ROTHMAN & SANDER GREENLAND, MODERN EPIDEMIOLOGY 62, 255–59 (2d ed. 1998).

43. MARIAN R. YARROW ET AL., CHILD REARING: AN INQUIRY INTO RESEARCH AND METHODS 80 (1968).

III

MAJOR LITERATURE REVIEWS

Of the six reviews of studies of corporal punishment published between 1996 and 2005,⁴⁴ only Gershoff⁴⁵ supports a spanking prohibition. Paolucci and Violato emphasized that the associations between corporal punishment and affective, cognitive, or behavioral child outcomes were very small,⁴⁶ concluding that the patterns of the causal evidence “seem to support Larzelere’s . . . contention that it is premature to impose guilt on the majority of parents who use ordinary spanking.”⁴⁷ Horn’s review of corporal punishment in African American families concluded, “[I]t is possible that there are benefits to nonabusive physical punishment for African-American children.”⁴⁸

The second Larzelere review and the Gershoff review were considered sufficiently important to be compared by Alan Kazdin, a recent president of the American Psychological Association, who called them both “exemplary in terms of scope, comprehensiveness, and scholarship.”⁴⁹ Kazdin and one of his colleagues said, “A top priority for research on spanking would seem to be a comparison of spanking with alternative procedures,”⁵⁰ a comparison made by Larzelere and Kuhn⁵¹ in their most recent literature review. Therefore, the discussion here will summarize this latest review and the contrasting one by Gershoff.⁵²

Gershoff’s review is thorough and has been cited more often than the other five reviews. But it fails to address the crucial question—should the use of disciplinary spanking be enjoined—because most studies upon which her review depended emphasized overly severe forms of corporal punishment and her effect sizes were based on unadjusted correlations with child outcomes. Although this problem was usually due to shortcomings of the original studies,

44. Gershoff, *supra* note 3, at 542; Ivor B. Horn et al., *Nonabusive Physical Punishment and Child Behavior Among African-American Children: A Systematic Review*, 96 JAMA 1162, 1163 (2004); Robert Larzelere, *Child Outcomes of Nonabusive and Customary Physical Punishment by Parents: An Updated Literature Review*, 3 CLINICAL CHILD & FAM. PSYCHOL. REV. 199, 200 (2000); Robert E. Larzelere, *A Review of the Outcomes of Parental Use of Nonabusive or Customary Physical Punishment*, 98 PEDIATRICS 824, 824 (1996); Robert E. Larzelere & Brett R. Kuhn, *Comparing Child Outcomes of Physical Punishment and Alternative Disciplinary Tactics: A Meta-analysis*, 8 CLINICAL CHILD & FAM. PSYCHOL. REV. 1, 4, 17 (2005); Elizabeth O. Paolucci & Claudio Violato, *A Meta-analysis of the Published Research on the Affective, Cognitive, and Behavioral Effects of Corporal Punishment*, 138 J. PSYCHOL. 194, 208–10 (2004).

45. ELIZABETH T. GERSHOFF, REPORT ON PHYSICAL PUNISHMENT IN THE UNITED STATES: WHAT RESEARCH TELLS US ABOUT ITS EFFECTS ON CHILDREN 26 (2008), available at <http://nospank.net/gershoff.htm>.

46. Paolucci & Violato, *supra* note 44, at 197.

47. *Id.* at 215.

48. Horn et al., *supra* note 44, at 1162.

49. Benjet & Kazdin, *supra* note 3, at 205.

50. *Id.* at 215.

51. Larzelere & Kuhn, *supra* note 44, at 17.

52. Gershoff, *supra* note 3.

Gershoff based her effect sizes on correlations even from studies with stronger causally relevant evidence.

To document one example, Gershoff based one effect size of $d = .30$ on Gunnoe and Mariner's⁵³ correlation of $r = .15$ ⁵⁴ between spanking and fighting. By doing so, Gershoff ignored that, after adjusting statistically for initial differences in bullying, spanking was associated with significantly *lower* aggression five years later for African Americans, for four- to seven-year-olds, and for girls, and higher aggression only for European Americans and for eight- to eleven-year-olds.⁵⁵ Gershoff's effect size thus indicated that spanking was associated with higher aggression even though the original authors' conclusion stated that "[f]or most children, claims that spanking teaches aggression seem unfounded."⁵⁶ Conclusions based on unadjusted longitudinal correlations rather than conflicting causal evidence would be like concluding that radiation treatment is ineffective for Stage II endometrium cancer,⁵⁷ regardless of stronger causal evidence to the contrary.

In her meta-analysis, Gershoff acknowledges the correlation problem, stating,

[b]ecause these meta-analyses are based primarily on correlational studies, parental corporal punishment cannot be identified definitively as the cause of these child behaviors. . . . [I]t is conceivable that the causal direction is reversed from what might be expected, such that children are driving the associations . . . [or] there might also be a third variable that predicts both parents' use of corporal punishment and child behaviors, such as parents' inconsistent style of discipline.⁵⁸

Notwithstanding her previous acknowledgement that "findings of correlation do not prove causation,"⁵⁹ in a recent advocacy publication, Gershoff treats the child correlates of corporal punishment as causal effects: "Taken together, the findings from these research studies support a causal link between parents' use of physical punishment and increases in children's future aggression, over and above the propensity for disobedient and aggressive child behavior to elicit parental physical punishment."⁶⁰

As illustrated in Table 1, the longitudinal correlations of radiation treatment appear detrimental because female recipients of radiation have a higher rate of dying than the general population of American women of that age. If reliance on correlations leads to the wrong causal conclusion about corrective medical actions such as radiation treatment, what is the basis for concluding that

53. Marjorie L. Gunnoe & Carrie L. Mariner, *Toward a Developmental-Contextual Model of the Effects of Parental Spanking on Children's Aggression*, 151 ARCHIVES PEDIATRICS & ADOLESCENT MED. 768, 772 (1997).

54. Gershoff, *supra* note 3, at 545.

55. Gunnoe & Mariner, *supra* note 53, at 772-73.

56. *Id.* at 768.

57. *See supra* table 1.

58. Gershoff, *supra* note 3, at 550.

59. *Id.*

60. GERSHOFF, *supra* note 45, at 14.

correlational evidence can lead to the correct causal conclusion about corrective disciplinary actions by parents? In both cases, the longitudinal correlations are biased against the corrective action because the group without the corrective action had a lower incidence of the problems that required some corrective action (cancer or oppositional behaviors). In assessing the child effects of corporal punishment, reliance on correlational evidence creates a selection bias due to child effects on parents. Nonetheless, Gershoff interprets the correlations as though they reflect only parent effects on children.

In her recent publication, Gershoff responded to the causal issue by citing three studies that show bidirectional longitudinal associations (child-to-parent and parent-to-child).⁶¹ But physical punishment was only one of five to nine items in the punishment measures used in those studies. Gershoff also cited an intervention study that she claimed showed that improvements in children's behavior were mediated by reductions in physical punishment. But physical punishment was only one of six items in the measure of "harsh parenting" on which she based her conclusions, and the composite harsh parenting was actually associated with subsequent behavioral improvements.⁶² Instead of showing that causal effects of spanking are uniquely detrimental, Gershoff's examples confirm that most statistical analyses are biased against all corrective disciplinary actions, even when controlling statistically for initial scores on the child outcomes. In addition, the study she cites for her strongest causal evidence actually found that the broad measure of harsh parenting was associated with greater *improvements* in behavior subsequently.

In addition, Gershoff's⁶³ meta-analysis consisted mostly of studies that not only failed to exclude overly severe corporal punishment, but emphasized severe usage in their measures. Out of fifty-two studies of broadly antisocial outcomes, sixty-five percent emphasized overly severe corporal punishment, ranging from spanking with an instrument to "beating with a stick,"⁶⁴ "hit[ting] . . . with a fist,"⁶⁵ or "slap[ping] in the face."⁶⁶ When Benjet and Kazdin compared Gershoff's review with the second Larzelere review,⁶⁷ they noted, "[b]ecause abusive behaviors are not excluded, the negative effects of severe

61. *Id.*

62. This is based on the *negative* predictive path coefficient (*b*) for DDI Harsh Parenting in Table 5 for path *b* in Figure 4 (Panel 3). Theodore P. Beauchaine et al., *Mediators, Moderators, and Predictors of 1-Year Outcomes Among Children Treated for Early-Onset Conduct Problems: A Latent Growth Curve Analysis*, 73 J. CONSULTING & CLINICAL PSYCHOL. 371, 383–84 (2005).

63. Gershoff, *supra* note 3; Diana Baumrind et al., *Ordinary Physical Punishment: Is it Harmful? Comment on Gershoff (2002)*, 128 PSYCHOL. BULL. 580, 581–82 (2002).

64. Anette Engfer & Klaus A. Schneewind, *Causes and Consequences of Harsh Parental Punishment: An Empirical Investigation in a Representative Sample of 570 German Families*, 6 CHILD ABUSE & NEGLECT 129, 133 (1982).

65. Robert T. Muller, *Family Aggressiveness Factors in the Prediction of Corporal Punishment: Reciprocal Effects and the Impact of Observer Perspective*, 10 J. FAM. PSYCHOL. 474, 477–78 (1996).

66. M. M. Lefkowitz et al., *Punishment, Identification and Aggression*, 9 MERRILL-PALMER Q. 159, 161 (1963); *see also* Baumrind et al., *supra* note 63, at 581.

67. Larzelere, *supra* note 3.

corporal punishment may cloud the effects of mild corporal punishment such as spanking.”⁶⁸

The latest review of the literature by Larzelere and Kuhn⁶⁹ improved on Gershoff’s⁷⁰ meta-analysis method by distinguishing among four types of corporal punishment and by obtaining causally relevant evidence rather than correlational evidence. First, they distinguished among conditional, customary, overly severe, and predominant use of corporal punishment as follows:⁷¹ “*Conditional* spanking was defined as physical punishment that was used primarily to back up milder disciplinary tactics (e.g., reasoning or time-out), used for defiance, or used in a controlled manner.”⁷² “*Customary* physical punishment was defined as typical parental usage (e.g., usage or frequency), without emphasizing its severity or predominance. . . . *Overly severe* physical punishment was based on measures that gave extra points for the severity of physical punishment . . . , [and] *predominant usage* of physical punishment included studies investigating predominant disciplinary tactics.”⁷³

Second, Larzelere and Kuhn dealt with the correlational–causal issue in two ways. First, they used the most causally relevant statistics from articles rather than relying primarily on correlations.⁷⁴ Second, they compared the effect size of each type of corporal punishment with effect sizes of alternative disciplinary tactics from the same studies.⁷⁵ If corporal punishment has *more* detrimental effects on children than alternative disciplinary tactics, this should be shown in differences between their effect sizes in predicting an outcome such as antisocial aggression. If, however, the apparently detrimental effects of corporal punishment are due to behaviorally difficult children causing parents to use all disciplinary punishments more frequently, then there should be no differences in how strongly varying disciplinary tactics are associated with aggression. Comparing the effect sizes of the two radiation treatments in Table 1 correctly identifies the more effective treatment (their *ds* differ by 1.20), whereas Gershoff’s method would conclude incorrectly that both radiation treatments are harmful because they cause a higher rate of deaths than in the no-treatment comparison group. Similar to the radiation-treatment example, comparing the effect sizes of corporal punishment with an alternative disciplinary tactic compares children who are similar in provoking corrective disciplinary actions. In contrast, the usual longitudinal correlation compares oppositional children who are spanked with a group of more-cooperative children who need few disciplinary tactics of any kind.

68. Benjet & Kazdin, *supra* note 3, at 204.

69. Larzelere & Kuhn, *supra* note 44.

70. Gershoff, *supra* note 3.

71. Larzelere & Kuhn, *supra* note 44, at 17.

72. *Id.*

73. *Id.*

74. *Id.* at 3, 17.

75. *Id.* at 3–4.

Comparing the effect sizes of corporal punishment with alternative disciplinary tactics within the same study is also fairer because it makes other factors equivalent. The comparisons of effect sizes are then based on the same statistical analyses on the same families. To adopt this strategy, Larzelere and Kuhn⁷⁶ included all twenty-six studies that appeared in either Larzelere's⁷⁷ previous review or Gershoff's⁷⁸ meta-analysis that investigated child outcomes of at least one alternative disciplinary tactic in addition to physical punishment.

The outcomes of physical punishment compared unfavorably with alternative, noncorporal disciplinary tactics only when it was the *primary* disciplinary method or was too severe (such as beating up a child or striking the face or head).⁷⁹ The outcomes of customary spanking were neither better nor worse than for any alternative tactic, except for one study in which spanking reduced drug abuse more than did nonphysical punishment.⁸⁰ Conditional spanking led to less noncompliance or antisocial behavior than ten of thirteen alternative disciplinary tactics and produced outcomes equivalent to those of the remaining three tactics.⁸¹ By definition, conditional spanking was used when children responded defiantly to other disciplinary tactics such as time-out (based on research on two- to six-year-olds).

IV

CAUSAL EVIDENCE RELEVANT TO SPANKING PROHIBITIONS

Section II established that the empirical evidence most relevant to spanking prohibitions is evidence about the causal effects of spanking in the most appropriate disciplinary situations, for example, when young children respond defiantly after parents have tried other disciplinary tactics. Subsection A below next summarizes evidence from the four studies with the most causally conclusive evidence of the use of spanking to enforce time-out in behavioral parent training. This is followed by evidence from five other studies that approximate the same type of spanking and provide outcome comparisons with alternative disciplinary tactics. Subsection B then summarizes and critiques the strongest causal evidence against spanking cited by spanking prohibitionists, consisting of seven studies that use net-effects regression⁸² to strengthen the causal evidence beyond unadjusted correlations. That subsection concludes with a second set of studies using net-effects regression to compare the outcomes of customary spanking with outcomes of alternative disciplinary tactics. Finally, subsection C briefly summarizes studies of overly severe or predominant use of

76. *Id.* at 4, 17.

77. Larzelere, *supra* note 3.

78. Gershoff, *supra* note 3.

79. Larzelere & Kuhn, *supra* note 44, at 1.

80. F. S. Tennant, Jr. et al., *Some Childhood Antecedents of Drug and Alcohol Abuse*, 102 AM. J. EPIDEMIOLOGY 377, 380, 382 (1975).

81. Larzelere & Kuhn, *supra* note 44, at 1.

82. Turner, *supra* note 39, at 23–25.

corporal punishment, which are the only types of physical punishment clearly associated with more detrimental effects than alternative disciplinary tactics.

A. Causal Effects of Conditional Spanking

Only four studies provide causally conclusive evidence about the effects of spanking in a disciplinary situation considered appropriate by clinical psychologists (that is, to back up time-out in a controlled manner with clinically defiant two- to six-year-olds).⁸³ Both major reviews of the literature recognized these studies as the only causally conclusive studies of corporal punishment.⁸⁴ These studies were part of Roberts' research program to identify the necessary components of behavioral parent training, which is recognized as one of the most effective treatments for disruptive-behavior diagnoses in young children, including oppositional defiant disorder, conduct disorder,⁸⁵ and ADHD.⁸⁶

1. Randomized Comparisons of Spank Back-Up for Time-Out Versus Alternatives

Roberts' four studies tested whether the traditional spank back-up for time-out was necessary in the Forehand-McMahon⁸⁷ version of behavioral parent training and what alternatives could be used instead of spanking to enforce cooperation with the time-out chair in the clinic. Alternatives included a child-determined release from time-out, a restraint procedure, and a brief, forced, room isolation. When using the spank back-up, children cooperated significantly more with time-out or parental commands than when using the child-release or the restraint back-up, and children's cooperation was the same as when using the room isolation. Overall, clinically defiant children required excessive repetitions of the enforcement procedure before cooperating with time-out in only 12% of cases with the spank back-up, 17% with the room isolation back-up, and 56% of cases with the restraint back-up.⁸⁸ Across all four studies, compliance rates to parental commands increased from 23% to 70% with the spank back-up, 21% to 72% with the room-isolation back-up, 18% to

83. Roberts & Powers, *supra* note 6; Mark W. Roberts, *Enforcing Chair Timeouts with Room Timeouts*, 12 BEHAV. MODIFICATION 353, 360 (1988); Arthur W. Bean & Mark W. Roberts, *The Effect of Time-Out Release Contingencies on Changes in Child Noncompliance*, 9 J. ABNORMAL CHILD PSYCHOL. 95, 101 (1981); D. E. Day & M. W. Roberts, *An Analysis of the Physical Punishment Component of a Parent Training Program*, 11 J. ABNORMAL CHILD PSYCHOL. 141, 147 (1983).

84. Gershoff, *supra* note 3, at 550 ("Because these meta-analyses are based primarily on correlational studies, parental corporal punishment cannot be identified definitively as the cause of these child behaviors and experiences, with the exception of immediate compliance." Three of Roberts' studies are included with two other studies of immediate compliance at 545 and constitute three of the four studies correctly identified in her Table 3 as using an experimental design (the other experimental study used a loud noise as the measure of "corporal punishment": J. C. LaVoie, *Punishment and Adolescent Self-Control*, 8 DEVELOPMENTAL PSYCHOL. 16, 18 (1973)); Larzelere & Kuhn, *supra* note 44, at 5 ("randomized clinical trials"), 17, 19.

85. Eyberg et al., *supra* note 23, at 226-29.

86. Pelham & Fabiano, *supra* note 22, at 187.

87. Forehand & McMahon, *supra* note 19.

88. Roberts & Powers, *supra* note 6, at 263.

52% with the restraint back-up, and 24% to 57% with the child-release procedure.⁸⁹ In contrast, children's compliance decreased from 27% to 13% in a control condition without any of the components of behavioral parent training.⁹⁰

Another benefit of effective enforcements for time-out is that the back-up procedure gets phased out quickly as the child learns to cooperate with time-out. The mean number of spanking or room-isolation back-ups decreased from 2.5 during the first implementation of time-out to 0.6 during its second implementation in the clinic,⁹¹ and from 0.7 during the first week at home to 0.15 in each of the third and fourth weeks after training in the clinic (the median decreased from 0.5 to 0.0).⁹² Skillful enforcement of time-out with spanking (or room isolation, the only equivalently effective alternative) results in parents' rapidly phasing out spanking, thereby moving into the low-spanking group.

In sum, training a child to cooperate with time-out is a crucial skill in behavioral parent training for young children with disruptive-behavior disorders. Roberts' series of randomized studies demonstrated causally conclusive evidence that this goal is achieved most effectively by the spank back-up and by the room-isolation back-up, both of which were significantly more effective than the restraint back-up or child-determined-release conditions.

2. The Importance of Roberts' Studies

Roberts' four studies are especially important for determining the advisability of spanking prohibitions for five reasons: they are the only studies of spanking that (1) are causally conclusive, (2) specify a nonabusive implementation of spanking in an appropriate disciplinary situation, (3) focus on the most behaviorally difficult children, (4) show how spanking can enforce preferred nonphysical tactics and then be phased out, and (5) show the benefit of multiple disciplinary options.

First, these experimental, randomized trials are the only studies of corporal punishment with causally conclusive results. In any other area, the results of randomized trials would trump conclusions from less causally conclusive studies.⁹³ For example, the Society of Clinical Child and Adolescent Psychology considers only "good group-design experiments"⁹⁴ to determine the effectiveness of psychosocial treatments. Studies with less-conclusive causal evidence are not considered at all. That means that if spanking were being

89. *Id.* at 264; Roberts, *supra* note 83, at 361; Bean & Roberts, *supra* note 83, at 101; Day & Roberts, *supra* note 83, at 148.

90. Bean & Roberts, *supra* note 83, at 101.

91. Roberts, *supra* note 83, at 363.

92. Roberts & Powers, *supra* note 6, at 266.

93. Samuel Shapiro, *Meta-analysis/Shmeta-analysis*, 140 AM. J. EPIDEMIOLOGY 771, 771-77 (1994); Samuel Shapiro, *Is Meta-analysis a Valid Approach to the Evaluation of Small Effects in Observational Studies?*, 50 J. CLINICAL EPIDEMIOLOGY 223, 223 (1997); James V. Lacey, Jr. et al., *Menopausal Hormone Replacement Therapy and Risk of Ovarian Cancer*, 288 JAMA 334, 337 (2002).

94. Silverman & Hinshaw, *supra* note 38, at 5.

evaluated for effectiveness by clinical child psychologist standards, only Mark Roberts' four studies would be considered at all; all of the studies cited for adverse effects by prohibitionists would be disregarded as inconclusive because of the inferiority of their methods for showing unbiased causal evidence.

Second, these are the only studies guaranteeing that the spanking is nonabusive (two open-handed swats to the buttocks under the supervision of a clinical psychologist) and that specify an appropriate disciplinary situation (to enforce compliance with time-out in children of about two to six years of age).

Third, these studies demonstrate effectiveness for the most behaviorally difficult young children. These children have the highest risk for delinquency and crime due to the stability of antisocial behavior after ages two to eight.⁹⁵ If spanking prohibitions were to undermine the effectiveness of parents' authority with behaviorally difficult children, a society would likely become more violent when those at-risk children grew up.

Fourth, Roberts' studies show that more-forceful tactics are often needed to make preferred disciplinary tactics such as reasoning and time-out more effective for behaviorally difficult young children. Few other studies show how skillful sequencing of disciplinary tactics is crucial for enhancing the effectiveness of preferred tactics. Larzelere and his colleagues showed that reasoning becomes more effective by itself with two- and three-year-olds to the extent that mothers back it up with nonphysical consequences at least ten percent of the time.⁹⁶ Roberts' series of studies extended that sequence by documenting that time-out must be consistently enforced by a forceful tactic such as a two-swat spanking or a room isolation to become effective in maintaining normal levels of cooperation. What is not known is how often a similar history of skillfully backing up nonphysical tactics with spanking is a factor underlying the normal levels of cooperation in nonclinical children who no longer need to be spanked.

Finally, Roberts and Powers showed the benefit for parents to have multiple options for enforcing time-out because two effective options provide an alternative if the first option is not sufficiently effective. First, the authors showed that each of the two most effective back-ups worked for children when the other back-up was slow in achieving compliance with time-out.⁹⁷ After seven escapes from time-out or seven time-outs without meeting compliance goals, the parent was switched to the other enforcement procedure. This adjustment occurred for fifteen (42%) of thirty-six children, with three switching from the

95. Daniel S. Shaw, *The Development of Aggression in Early Childhood*, in 1 THE CRISIS IN YOUTH MENTAL HEALTH: CRITICAL ISSUES AND EFFECTIVE PROGRAMS 183, 183–85 (Hiram E. Fitzgerald et al. eds., 2006); Sylvana M. Cote et al., *The Joint Development of Physical and Indirect Aggression: Predictors of Continuity and Change During Childhood*, 19 DEV. & PSYCHOPATHOLOGY 37, 44–49 (2007); Dan Olweus, *Stability of Aggressive Reaction Patterns in Males: A Review*, 86 PSYCHOL. BULL. 852, 863 (1979).

96. Larzelere et al., *supra* note 30, at 397–98.

97. Roberts & Powers, *supra* note 6, at 263–65.

spank back-up to the room isolation, and the other twelve switching from one of the other three back-up tactics to the spank back-up. The second back-up procedure always produced the desired compliance with time-out and with parental commands, which increased from 16% initially to 30% before switching to a second back-up to 79% after that switch.⁹⁸ The equivalence of these enforcement tactics also allowed parents to select the one they preferred to use at home. Most parents chose the spank back-up (64%), whereas 14% selected the room isolation, 9% selected the restraint procedure, and 14% selected an individualized combination.⁹⁹ Just as having multiple effective prescription drugs permits doctors and patients to work together to choose the most desirable drug for their situation and preferences, multiple effective back-up options for time-out likewise expand the effective options, including extra options to choose when the first option fails to produce cooperation with time-out.

3. Replies to Critiques of Roberts' Studies

The relevance of Roberts' studies has been criticized in several ways by spanking-prohibition advocates. This section replies to those criticisms.

Gershoff¹⁰⁰ acknowledges that Roberts' studies are the only causally definitive studies of corporal punishment and that they show that spanking reduces child noncompliance overall. Nonetheless, she has criticized the importance of Roberts' studies because their results (1) are limited to "immediate compliance," (2) show inconsistent effect sizes, and (3) Roberts himself concluded that physical punishment is not necessary.¹⁰¹

Gershoff's¹⁰² term *immediate compliance* suggests that spanking merely puts an immediate stop to noncompliance, thereby ending the discipline incident. When spanking had to be used in Roberts' studies, however, it was always due to noncompliance with both a parent command and time-out. Successful compliance was measured in Roberts' studies by compliance with parental commands or cooperation with time-out, situations in which the spank back-up never needed to be used.¹⁰³ Furthermore, compliance with time-out is essential for a parent to restore and maintain acceptable levels of cooperation with these behaviorally difficult children. Otherwise, continued defiance will keep the child at high risk for developmental pathways of authority conflicts, antisocial behavior, delinquency, and crime.¹⁰⁴

98. *Id.* at 265.

99. *Id.* at 266.

100. Gershoff, *supra* note 3, at 539, 550; Gershoff, *supra* note 45, at 13.

101. GERSHOFF, *supra* note 45, at 13.

102. Gershoff, *supra* note 3, at 539.

103. Roberts & Powers, *supra* note 6, at 261.

104. Rolf Loeber et al., *Developmental Pathways in Disruptive Child Behavior*, 5 DEV. & PSYCHOPATHOLOGY 103, 111 (1993); Jennifer L. White et al., *How Early Can We Tell? Predictors of Childhood Conduct Disorder and Adolescent Delinquency*, 28 CRIMINOLOGY 507, 513 (1990).

The inconsistency of Gershoff's effect sizes from Roberts' studies is due to differences in what the spank back-up was compared with. When compared with a no-treatment control group, behavioral parent training with the spank back-up had an unusually large effect size of $d = 3.39$.¹⁰⁵ When comparing two behavioral parent training protocols that differed only in whether they used the spank back-up or not in Roberts' original study, the effect size was also large ($d = 1.73$).¹⁰⁶ When comparing otherwise-identical behavioral parent-training protocols that differed only in using the room-isolation as opposed to the spank back-up, Roberts and Powers found no significant differences, concluding that the "Barrier and Spank procedures appeared equally effective, replicating prior studies."¹⁰⁷ This is a typical pattern for effective clinical treatments, showing more effectiveness than a no-treatment control group but equivalent effectiveness to another treatment. In fact, documenting outcomes equivalent to an effective treatment is one of the criteria for evidence-based treatments in clinical child psychology.¹⁰⁸

Finally, Gershoff quoted Day and Roberts as saying in 1983 that "there was no support for the necessity of the physical punishment,"¹⁰⁹ which actually meant that the traditional spank back-up was no longer the *only* maximally effective back-up for time-out. Roberts was still using the spank back-up in 1990 and switched parents to it when their originally assigned back-up procedure was not effective quickly enough (twelve (44%) of twenty-seven parents).¹¹⁰ Granted, Roberts now prefers the room-isolation back-up and so no longer uses the spank back-up. Given two equally effective enforcements for time-out, therapist preference is a valid reason for choosing one over the other. Another reason that most behavioral clinicians now use the room-isolation instead of the spank back-up is concern about their reputation, which influences clinical referrals. Neither Roberts' current practice nor reputation considerations negates the equivalent effectiveness of the spanking and room-isolation back-ups for time-out or the advantages of having both available as parental options, as shown by Roberts and Powers.¹¹¹ Clinics and families lacking an isolation room need to retain the spanking option, or they would have to use alternative back-ups for time-out that are either unproven or have been shown to be less effective. The effectiveness of the room-isolation has never been compared with the spank back-up for any setting other than the four- by-five-foot empty room with a four-foot-high plywood barrier used in Roberts' clinic.¹¹²

105. Gershoff, *supra* note 3, at 545.

106. Larzelere & Kuhn, *supra* note 44, at 5; Bean & Roberts, *supra* note 83, at 102.

107. Gershoff, *supra* note 3, at 545; Roberts & Powers, *supra* note 6, at 257.

108. Silverman & Hinshaw, *supra* note 38, at 5.

109. GERSHOFF, *supra* note 45, at 13; Day & Roberts, *supra* note 83, at 150.

110. Roberts & Powers, *supra* note 6, at 257, 269.

111. *Id.* at 257.

112. *Id.* at 260.

Roberts has also been criticized for using so many spanking repetitions, a mean of 8.6 spankings before complying with time-out in one study.¹¹³ This shows the difficulty of getting cooperation with time-out from clinically defiant young children who have learned to undermine all parental-control attempts, a key coercive process in the development of antisocial behavior.¹¹⁴ According to Gerald Patterson, the leading expert on coercive process, the key for treating this type of defiance is to teach parents “how to punish more effectively,”¹¹⁵ referring primarily to time-out. Effective time-out punishment requires the child to cooperate with it, which is accomplished by consistent use of the back-up enforcement for noncompliance with time-out. Roberts later improved the protocol by switching to either the spank back-up or the room-isolation back-up after seven escapes from time-out.¹¹⁶ So having multiple effective options enhances the ability of parents to discover what will work for their child in a particular situation, increasing the likelihood that the child will cooperate with time-out and thereby decreasing the probability that the parents will escalate the severity of their verbal or physical punishment.

4. Other Studies of Conditional Spanking

In their most recent meta-analysis on corporal punishment, Larzelere and Kuhn¹¹⁷ found five other studies that approximated conditional spanking on at least one of its following components: (1) clearly nonabusive spanking for (2) defiant refusal (3) to cooperate with other disciplinary tactics. Together with Roberts’ studies, the nine studies together showed that conditional spanking was associated with significantly less noncompliance or antisocial behavior than ten of thirteen alternative disciplinary tactics investigated in these studies. Table 2 summarizes one of those studies, which was unusual in highlighting defiant situations by asking mothers how they tried to get their child to cooperate when dealing with “extreme” or dangerous misbehavior.¹¹⁸ The extent to which they would use physical punishment in those situations was marginally associated with *less* aggression in preschool two months later, $r = -.19$ ($p < .10$ [which is generally considered marginally significant in science]), an indication of effectiveness matched only by privilege removal in that study. Both tactics showed significantly greater effectiveness at reducing preschool aggression than at least three of the other five disciplinary tactics ($p < .05$, the usual scientific standard for significant, reliable results that were unlikely to be chance occurrences). Other studies showed that conditional spanking was more effective than some alternative tactics at putting an immediate stop to

113. Day & Roberts, *supra* note 83, at 148.

114. JOHN REID ET AL., ANTISOCIAL BOYS 39–60 (1992).

115. GERALD R. PATTERSON, COERCIVE FAMILY PROCESS 111 (1982).

116. Roberts & Powers, *supra* note 6, at 259–60.

117. Larzelere & Kuhn, *supra* note 44, at 20–21.

118. YARROW ET AL., *supra* note 43, at 75 (“NIMH Study”), 154–57, 171, 179.

defiance¹¹⁹ or noncompliance,¹²⁰ and at reducing antisocial behavior and impulsivity.¹²¹

Table 2

Effect Sizes Predicting Aggression in Preschool from Disciplinary Tactics Used for “Extreme” or Dangerous Misbehaviors, from Yarrow et al. (1968)¹²²

Disciplinary Tactic	Effect Sizes		Difference from <i>d</i> for Physical Punishment
	<i>r</i>	<i>d</i>	
Isolation	.09	.18	.56*
Love withdrawal	.12	.24	.62**
Reasoning	.14 ^a	.28	.66**
Deprivation of privileges	-.19 ^b	-.38 ^a	.00
Physical punishment	-.19 ^b	-.38 ^a	n/a
Scolding	.12	.24	.62**
Diverting attention	.23 ^{a*}	.47*	.85***

^aThe sign was reversed from Yarrow et al., because their listed *r* is for the reverse-scored tactic.

^b*p* < .10.

p* < .05. *p* < .01. ****p* < .001.

In sum, the only four causally definitive studies of spanking for defiant refusal to comply with other tactics have shown that a two-swat spank to the buttocks is tied with one other enforcement tactic as the most effective tactic in that situation with two- to six-year-olds. Having two equivalently effective options is optimal for parents and therapists because each tactic works better for some children than the other tactic. Further, each option has proven to be effective when the other option was not effective in getting the child to

119. Ritchie, *supra* note 15, at 587; Robert E. Larzelere & Brett R. Kuhn, Immediate Effectiveness of Disciplinary Tactics by Type of Noncompliance (A Reanalysis of Ritchie, 1999), Presentation at the Annual Meeting of the American Psychological Association (2005).

120. Michael Chapman & Carolyn Zahn-Waxler, *Young Children's Compliance and Noncompliance to Parental Discipline in a Natural Setting*, 5 INT'L J. BEHAV. DEV. 81, 86 (1982).

121. M. A. Straus & V. E. Mouradian, *Impulsive Corporal Punishment by Mothers and Antisocial Behavior and Impulsiveness of Children*, 16 BEHAV. SCI. & L. 353, 362 (1998).

122. YARROW ET AL., *supra* note 43, at 75 (“NIMH study”), 154–57, 171, 179. Aggression was rated by preschool teachers two months after mothers described their disciplinary tactics with the following open-ended questions: “How do you go about getting [your child] to act as you want him to? . . . Think of some extreme things [your child] has done” and questions about “things that involve the total safety of a child, such as playing with matches, running in the street.” *Id.* Effect sizes (*ds*) were estimated from correlations with *Dstat*. JOHNSON, *supra* note 33. Since *d* estimates the association of each disciplinary tactic with aggression in standard deviations (*SDs*), the *ds* can be subtracted from each other to get differences between two tactics in their associations with preschool aggression, also in *SDs*. Positive *ds* indicate that the disciplinary tactic is associated with greater preschool aggression. Positive differences from the *d* for physical punishment indicates that the tactic is associated with more preschool aggression than is physical punishment.

cooperate with time-out. Either back-up enforces compliance with time-out, so that time-out can be relied upon more often, enabling parents to phase out the back-up tactic used to enforce time-out. We do not know how many infrequent spankers have gotten to that point by skillfully using spanking to enforce other disciplinary tactics, which they now rely on. Four of the five other studies that have at least one characteristic of this type of conditional spanking showed significantly greater effectiveness for spanking than for most alternative tactics in reducing defiance, noncompliance, or antisocial behavior.¹²³

B. Causally Relevant Evidence Against Spanking

The strongest causal evidence against spanking is from longitudinal studies showing that spanking frequency at a given initial time predicts more antisocial behavior problems one or two years later even after controlling statistically for initial differences on the behavior problems.¹²⁴ We are calling these *causally relevant* studies, because they have improved their causal evidence beyond that provided by unadjusted correlations but do not provide the causally definitive evidence required for evaluating clinical child treatments and new prescription drugs.

Seven longitudinal studies investigated whether customary spanking of children younger than thirteen predicted subsequent antisocial behavior or aggression after controlling statistically for initial levels of those outcomes. These studies showed nonsignificant,¹²⁵ small,¹²⁶ or mixed effects¹²⁷ of customary spanking on subsequent antisocial behavior or aggression. The small, significant effects were found mostly for non-Hispanic European Americans or in samples dominated by that group, with effect sizes of $\beta = .05$,¹²⁸ $.06$,¹²⁹ and $.07$,¹³⁰ equivalent to $d = .10$, $d = .12$, and $d = .12$, respectively. The study with mixed results was the only study that did not rely solely on parents' reporting all of the data.¹³¹ Using a distinct source of information for the child-outcome variable

123. Larzelere & Kuhn, *supra* note 44, at 20–21.

124. MURRAY A. STRAUS & VERA MOURADIAN, *BEATING THE DEVIL OUT OF THEM: CORPORAL PUNISHMENT IN AMERICAN FAMILIES AND ITS EFFECTS ON CHILDREN* 196–205 (2d ed. 2001).

125. Mary K. Eamon, *Poverty, Parenting, Peer, and Neighborhood Influences on Young Adolescent Antisocial Behavior*, 28 J. SOC. SERVICE RES. 1, 10 (2001); Larzelere et al., *supra* note 30, at 399–400; Mark F. Schmitz, *Influences of Race and Family Environment on Child Hyperactivity and Antisocial Behavior*, 65 J. MARRIAGE & FAM. 835, 841 (2003).

126. Lingxin Hao & Ross L. Matsueda, *Family Dynamics Through Childhood: A Sibling Model of Behavior Problems*, 35 SOC. SCI. RES. 500, 511 (2006); Matthew K. Mulvaney & Carolyn J. Mebert, *Parental Corporal Punishment Predicts Behavior Problems in Early Childhood*, 21 J. FAM. PSYCHOL. 389, 392 (2007); Murray A. Straus et al., *Spanking by Parents and Subsequent Antisocial Behavior of Children*, 151 ARCHIVES PEDIATRICS & ADOLESCENT MED. 761, 765 (1997).

127. Gunnoe & Mariner, *supra* note 53.

128. Hao & Matsueda, *supra* note 126, at 520 (lagged effects, effect size calculated from additional information from authors).

129. Mulvaney & Mebert, *supra* note 126, at 395.

130. Straus et al., *supra* note 126, at 765.

131. Gunnoe & Mariner, *supra* note 53, at 770, 772–73.

(that is, child reported aggression), it found that customary spanking significantly *reduced* aggression in the following subgroups: all four- to seven-year-olds, all African Americans aged four to eleven, and all girls aged four to eleven.¹³² That report also replicated the usual, small adverse effect of customary spanking on antisocial behavior when all the information was obtained from the parent.

None of these studies included a never-spanked contrast group, and their small, apparently detrimental effects can easily be explained by substantive or methodological shortcomings. All the studies with significant effects asked parents how often they had spanked their child in the past week.¹³³ Parents who spanked as often as twenty-five times annually were more likely than not to be included in the no-spanking group for that week. Therefore, all these small detrimental effects came from comparisons between less-frequent and more-frequent spanking, rather than between never-spanking and some spanking, which is necessary to support spanking prohibitions. The small detrimental effects could also be due to substantive factors. Overly severe and abusive parents were not excluded, so the small effects may be due entirely to them. These studies do nothing to distinguish appropriate from inappropriate situations in which to use spanking, such as limiting its use to dangerous behaviors and defiance and to enforce other disciplinary tactics. The small detrimental effects could also be due to parents who spank too frequently because they use it for an overly wide range of disciplinary situations.

In addition, the small effects could easily be due to methodological artifacts. Reliance solely on parental reports has been shown to inflate evidence against disciplinary tactics.¹³⁴ The only statistically controlled study that relied on a source of information in addition to parental reporting found that spanking predicted *reduced* aggression more often than it predicted increased aggression.¹³⁵

Also, these small effects could easily be caused by residual confounding.¹³⁶ Statistical controls eliminate all confounding and thus yield unbiased causal evidence only when the process of selecting recipients for a corrective action is measured comprehensively¹³⁷ and without measurement error.¹³⁸ For that reason, when epidemiologists make conclusions from similar data, they recognize that residual confounding remains after controlling statistically for fallible measures of confounding variables.¹³⁹ For example, residual confounding explained why

132. *Id.*

133. *E.g.*, Straus et al., *supra* note 126, at 762.

134. YARROW ET AL., *supra* note 43, at 80.

135. Gunnoe & Mariner, *supra* note 53, at 772–73.

136. Residual confounding is the part of the influence of the confounding variable that is left after being partially reduced. ROTHMAN & GREENLAND, *supra* note 42.

137. Heckman, *supra* note 40.

138. CAMPBELL & KENNY, *supra* note 41; Freedman, *supra* note 41.

139. ROTHMAN & GREENLAND, *supra* note 42.

the summer Head Start program appeared to be detrimental according to a major, early-evaluation study using similar statistical controls.¹⁴⁰

None of these studies investigated alternative disciplinary tactics with the same statistical analyses. If the association between the frequency of spanking and subsequent antisocial behavior is due to residual confounding with children's initial oppositional behavior, it follows that all disciplinary enforcements should show a similar association with antisocial behavior. This result would be consistent with Larzelere and Kuhn's meta-analysis, which found no differences in child outcomes of customary spanking compared with any alternative disciplinary tactic.¹⁴¹

Larzelere and his colleagues¹⁴² have recently implemented the first two studies known to compare the apparent effects of customary spanking with those of alternative tactics that parents could use instead. They replicated the small, apparently detrimental effects of spanking on subsequent antisocial behavior, controlling for preexisting differences. However, all kinds of nonphysical punishment also predicted higher antisocial behavior with the same controls. In their close replication of work by Straus et al.,¹⁴³ the first study found the following standardized regression coefficients predicting antisocial behavior from spanking and from alternative corrective actions when substituted for spanking in the same analyses with the same sample, controlling statistically for initial antisocial behavior: spanking: $\beta = .10, p < .05$; grounding: $\beta = .12, p < .01$; privilege removal: $\beta = .10, n.s.$; sending children to their room: $\beta = .09, p < .10$; psychotherapy during past year: $\beta = .24, p < .05$.¹⁴⁴ Using the Canadian National Longitudinal Survey of Children and Youth, the second replication found the following standardized regression coefficients predicting antisocial behavior two years later after controlling statistically for initial scores on antisocial behavior: physical punishment: $\beta = .07, p < .01$; nonphysical punishment: $\beta = .03, n.s.$; scolding or yelling: $\beta = .06, p < .05$; psychotherapy: $\beta = .07, p < .01$; and *Ritalin*: $\beta = .07, p < .01$.¹⁴⁵ In addition, *Ritalin*, nonphysical

140. WESTINGHOUSE LEARNING CORP. & OHIO UNIV., 1-2 THE IMPACT OF HEAD START: AN EVALUATION OF THE EFFECTS OF HEAD START ON CHILDREN'S COGNITIVE AND AFFECTIVE DEVELOPMENT 127 (1969) (report presented to the Office of Economic Opportunity Pursuant to Contract of B89-4536); Donald T. Campbell & Robert F. Boruch, *Making the Case for Randomized Assignment to Treatments by Considering the Alternatives: Six Ways in Which Quasi-Experimental Evaluations in Compensatory Education Tend to Underestimate Effects*, in EVALUATION AND EXPERIMENT: SOME CRITICAL ISSUES IN ASSESSING SOCIAL PROGRAMS 195, 209 (Carl A. Bennett & Arthur A. Lumsdaine eds., 1975).

141. Larzelere & Kuhn, *supra* note 44, at 1.

142. Robert E. Larzelere, Emilio Ferrer & et al., *Differences in Causal Estimates from Longitudinal Analyses of Residualized vs. Simple Gain Scores: Contrasting Controls for Selection and Regression Artifacts*, 34 INT'L J. BEHAV. DEV. 180 (2010); Robert E. Larzelere, Ronald B. Cox, Jr. & Gail L. Smith, *Do Nonphysical Punishments Reduce Antisocial Behavior More Than Spanking? A Comparison Using the Strongest Previous Causal Evidence Against Spanking*, 10:10 BMC PEDIATRICS 1 (2010).

143. Straus et al., *supra* note 126.

144. Larzelere, Cox & Smith, *supra* note 142, at 8.

145. Larzelere, Ferrer & et al., *supra* note 142, at 185.

punishment, and scolding or yelling all predicted significantly higher subsequent hyperactivity, whereas physical punishment did not, $\beta = .03$, *n.s.* Therefore, the strongest causal evidence against customary spanking is not unique to spanking, but applies as much to most corrective actions, including most corrective disciplinary actions by parents and corrective interventions by psychotherapists.

Other evidence from these two recent studies suggests that these small detrimental effects are biased due to residual confounding. First, all of these apparent effects became nonsignificant after improving the measure used to adjust statistically for preexisting differences in children.¹⁴⁶ Second, after reversing the bias by predicting simple changes in antisocial behavior during the following two years, all significant findings showed small beneficial effects, that is, in *reducing* antisocial behavior, albeit marginally for physical punishment and nonsignificantly for the professional interventions.¹⁴⁷ Third, this contradictory pattern of results was replicated in reversed time, that is, after reversing the temporal sequence of the data.¹⁴⁸ These last results would be expected from statistical artifacts,¹⁴⁹ not from actual causal effects, which can operate only forward in time.

Another strategy for obtaining causally relevant evidence from correlational studies is to compare the effect sizes of alternative disciplinary tactics with each other within the same studies. Using that strategy, Larzelere and Kuhn's meta-analysis found no differences in outcomes of customary spanking compared to any alternative disciplinary tactic studied except for one retrospective study favoring spanking over non-contact punishment for reducing substance abuse.¹⁵⁰

Grogan-Kaylor¹⁵¹ has claimed stronger causal evidence against customary spanking by using a statistical method called fixed-effects regression.¹⁵² However, his conclusions were based on associations between the frequency of spanking in the past week and the frequency of antisocial behavior during the past several months, and it is difficult to conclude that last week's spankings caused last month's antisocial behavior. His results cannot discriminate child effects on the parent from parent effects on the child. Accordingly, a replication

146. Larzelere, Cox & Smith, *supra* note 142, at 1, 8–16.

147. Larzelere, Ferrer & et al., *supra* note 142, at 180, 183–86.

148. *Id.* at 180, 185, 187.

149. CAMPBELL & KENNY, *supra* note 41, at 158–63.

150. Larzelere & Kuhn, *supra* note 44; Tennant, Jr. et al., *supra* note 80.

151. Andrew Grogan-Kaylor, *The Effect of Corporal Punishment on Antisocial Behavior in Children*, 28 SOC. WORK RES. 153, 160–61 (2004).

152. When applied to longitudinal data, fixed-effects regression basically subtracts the mean of each person's score from his scores at each occasion and then implements the regression analysis. The advantage of this method is that it eliminates the confounding influence of all between-subject differences on the analyses. However, it fails to control for the extent to which confounding variables vary across occasions for the same individual, unless explicitly in the regression analysis. Grogan-Kaylor's analyses reported associations between spanking and anti-social behavior during the same time period. *Id.*

of his analyses showed similar apparently detrimental “effects” for all types of nonphysical punishment and for psychotherapy.¹⁵³

In sum, the strongest causally relevant evidence against customary spanking yields small, apparently detrimental effects that can easily be due to a combination of several substantive and methodological factors that bias the results. Consistent with this, the first studies to use the same research methods for alternative disciplinary tactics showed similar results for corrective actions by both parents and professionals. In evaluating corrective actions by clinical psychologists, this type of evidence would not even be considered and would definitely not override the causally conclusive evidence that spanking can be effective for enforcing nonphysical disciplinary tactics, even in the most clinically defiant two- to six-year-olds.¹⁵⁴

C. Overly Severe and Predominant Use of Corporal Punishment

When comparing the outcomes of corporal punishment with outcomes of alternative disciplinary tactics, Larzelere and Kuhn’s meta-analysis found that the outcomes of physical discipline compared unfavorably with alternative disciplinary tactics only when it was the *primary* disciplinary method or was too severe (such as beating up a child or striking the face or head). Similarly, causally relevant studies of overly severe corporal punishment have generally found larger detrimental effects than have similar studies of customary spanking.¹⁵⁵ This supports what all professionals agree with—that overly severe corporal punishment is detrimental to children and should be avoided. Also, corporal punishment should not be the main disciplinary tactic used by parents.

153. Robert E. Larzelere & Ketevan Danelia, Similarity of Antisocial Outcomes for Nonphysical Punishment, Spanking, and Psychotherapy, Presentation at the Biennial Meeting of the Society for Research in Child Development (2009).

154. Silverman & Hinshaw, *supra* note 38, at 5.

155. Mara Brendgen et al., *Parent and Peer Effects on Delinquency-Related Violence and Dating Violence: A Test of Two Mediation Models*, 11 SOC. DEV. 225, 232 (2002); Sarah E. Fine et al., *Anger Perception, Caregivers’ Use of Physical Discipline, and Aggression in Children at Risk*, 13 SOC. DEV. 213, 220 (2004); Jennifer E. Lansford et al., *Friendship Quality, Peer Group Affiliation, and Peer Antisocial Behavior as Moderators of the Link Between Negative Parenting and Later Adolescent Externalizing Behavior*, 13 J. RES. ON ADOLESCENCE 161, 169 (2003); Jennifer E. Lansford et al., *Ethnic Differences in the Link Between Physical Discipline and Later Adolescent Externalizing Behaviors*, 45 J. CHILD PSYCHOL. & PSYCHIATRY 801, 804 (2004); Anna S. Lau et al., *Factors Affecting the Link Between Physical Discipline and Child Externalizing Problems in Black and White Families*, 34 J. COMMUNITY PSYCHOL. 89, 96 (2006); Liliana J. Lengua, *Anxiousness, Frustration, and Effortful Control as Moderators of the Relation Between Parenting and Adjustment in Middle-Childhood*, 17 SOC. DEV. 554, 562 (2008); Dustin A. Pardini et al., *The Development of Callous-Unemotional Traits and Antisocial Behavior in Children: Are There Shared and/or Unique Predictors?*, 36 J. CLINICAL CHILD & ADOLESCENT PSYCHOL. 319, 325 (2007).

V

OTHER EMPIRICAL ISSUES

A. Does Spanking Increase the Risk of Physical Child Abuse?

An important concern raised by spanking prohibitionists is that spanking might increase the risk of physical abuse. Most instances of physical abuse occur in situations that parents later describe as attempts to discipline their children.¹⁵⁶ Gershoff showed a strong association between corporal punishment and physical abuse,¹⁵⁷ but nine of the ten relevant studies in her meta-analysis were cross-sectional and the other study used retrospective data for the same child ages. Cross-sectional correlations do not prove causation.

Often, some version of the “stepping stone” argument is used to conclude that spanking must cause abuse because it typically precedes abuse. Such causal attribution from correlational evidence was shown to be specious when it was used to conclude that marijuana use is causally related to heroin use.¹⁵⁸ Since only a small proportion of the large majority of parents who spank their children ever abuse them and the most successful parents (authoritative parents) are average in their use of spanking, it is illogical to presume that abusive corporal punishment can be discouraged only by completely enjoining all spanking. To date, there has been no convincing evidence that spanking bans reduce physical child abuse.¹⁵⁹

To prevent escalations in frustration and risk of abuse, parents need to be able to get acceptable cooperation with nonabusive disciplinary tactics. Behavioral parent training accomplishes that by training parents to use time-out effectively, but it often requires an effective enforcement tactic. Four studies have documented a decrease in spanking following behavioral parent training. Two of the studies used the traditional spank back-up for time-out,¹⁶⁰ whereas two used alternative back-ups.¹⁶¹ Thus, the spank back-up and the room-isolation back-up are the two most effective enforcements to enhance compliance with time-out, which in turn prevents escalations toward abuse.

156. ALFRED KADUSHIN & JUDITH A. MARTIN, *CHILD ABUSE: AN INTERACTIONAL EVENT* 105 (1981).

157. GERSHOFF, *supra* note 45, at 17; Gershoff, *supra* note 3, at 546–47, 550.

158. Diana Baumrind, *Specious Causal Attributions in the Social Sciences: The Reformulated Stepping-Stone Theory of Heroin Use as Exemplar*, 45 J. PERSONALITY & SOC. PSYCHOL. 1289, 1296 (1983).

159. Robert E. Larzelere & Byron Johnson, *Evaluation of the Effects of Sweden's Spanking Ban on Physical Child Abuse Rates: A Literature Review*, 85 PSYCHOL. REP. 381, 386 (1999).

160. M. W. Roberts, *An Attempt to Reduce Timeout Resistance in Young Children*, 15 BEHAV. THERAPY 210, 212 (1984); S. Eyberg, *The Spank Back-Up in Time-Out with Preschool Children* (1993) (unpublished manuscript) (on file with the University of Florida).

161. Cheryl Bodiford McNeil et al., *Assessment of a New Procedure to Prevent Timeout Escape in Preschoolers*, 16 CHILD & FAM. BEHAV. THERAPY 27, 29–32 (1994); Carolyn Webster-Stratton, *Enhancing the Effectiveness of Self-Administered Videotape Parent Training for Families with Conduct-Problem Children*, 18 J. ABNORMAL CHILD PSYCHOL. 479, 481–85 (1990).

B. Is Spanking Necessarily More Aversive than Alternatives that Would Replace It?

The assumption underlying a spanking injunction is that spanking is necessarily more aversive than even the harshest alternative disciplinary tactic. There is little data on the aversiveness of typical spankings compared to alternatives that parents could use instead. The assumption that nonphysical punishment, such as time-out and privilege removal, is less aversive than customary spanking was not supported in a study of two- and three-year-olds.¹⁶² When used in combination with reasoning, nonphysical and physical punishments were tied as the most effective disciplinary responses at delaying recurrences of disobedience and fighting.¹⁶³ When used with reasoning, physical punishment resulted in a slightly higher level of child-distress intensity, but for a shorter period of time (an average of 4.2 minutes) than was the case following nonphysical punishment (an average of 5.2 minutes).¹⁶⁴

On average, preschoolers and fifth-graders considered a medium-to-hard spanking appropriate for dangerous misbehaviors, such as playing with matches, and moral transgressions, such as hitting or stealing.¹⁶⁵ On average, their mothers considered a medium-to-hard spanking appropriate only for dangerous behaviors by preschoolers and a light-to-medium spanking appropriate for moral transgressions and for dangerous behaviors by fifth graders.

If spanking is prohibited, do parents then use more-effective and less-aversive disciplinary tactics instead? Thirty years after Sweden passed the first spanking ban, there is still little supporting evidence. One comparison found that Swedish parents were less likely than American parents to use reasoning and behavior modification techniques and more likely to use physical restraint and “coercive verbal control.”¹⁶⁶ Support for disciplinary consequences has continued to erode in Sweden until only 31% of ten- to twelve-year-olds thought parents had the right to ground them and 53% thought parents had the right to remove their allowance.¹⁶⁷ Similarly, therapists reported that Norwegian parents showed “a surprisingly high prevalence of the permissive parenting form of child coercion. In these families, the parents often seem to be

162. Robert E. Larzelere et al., *Nonabusive Spanking: Parental Liberty or Child Abuse?*, 17 CHILDREN'S LEGAL RTS. J. 7, 12 (1997); Robert E. Larzelere & Jack A. Merenda, *The Effectiveness of Parental Discipline for Toddler Misbehavior at Different Levels of Child Distress*, 43 FAM. REL. 480, 483 (1994).

163. Robert E. Larzelere et al., *The Effects of Discipline Responses in Delaying Toddler Misbehavior Recurrences*, 18 CHILD & FAM. BEHAV. THERAPY 35, 43 (1996).

164. Larzelere et al., *supra* note 162, at 12.

165. Catron & Masters, *supra* note 6, at 1815, 1820.

166. K. Palmerus & S. Scarr, *How Parents Discipline Young Children: Cultural Comparisons and Individual Differences*, Presentation at the Biennial Meeting of the Society for Research in Child Development (1995), cited in Robert E. Larzelere, *Combining Love and Limits in Authoritative Parenting*, in PARENTHOOD IN AMERICA 85 (Jack C. Westman ed., 2001).

167. Staffan Janson, *Barn Och Misshandel: En Rapport Om Fropplig Bestraffning Och Annan Misshandel i Sverige Vid Slutet Av 1900-Talet* [*Children and Physical Abuse: A Report About Corporal Punishment and Other Physical Abuse in Sweden at the End of the 20th Century*], SOU 18, 58 (2001).

immobilized by unreasonable requests made by the child.”¹⁶⁸ A recent study in Quebec found that the annual percentage of mothers using minor physical discipline (usually spanking or slapping on the hand or arm) decreased from 48% to 43% from 1999 to 2004, but repeated psychological aggression increased from 48% to 52%.¹⁶⁹ Psychological aggression consisted mostly of yelling or screaming, but also included cursing at children or calling them “stupid.”

In sum, the available evidence suggests that spanking prohibitions may increase the use of verbal hostility, which has been shown to be one of the most detrimental forms of parental discipline,¹⁷⁰ with more detrimental effects than even physical child abuse in several studies.¹⁷¹ Spanking prohibitions may also increase the number of parents who cannot control their children’s coercive behavior, which puts those children at risk for delinquency and crime.¹⁷²

C. Ethnic Differences in Child Outcomes of Spanking

Another article in this special issue focuses on ethnic differences in the association of spanking with antisocial behavior. Many studies have found that spanking is more strongly associated with antisocial behavior and aggression in European Americans than in African Americans, with several studies indicating that spanking predicts significantly *lower* aggression in African Americans when reported by someone other than the parent.¹⁷³ This suggests that the typical way that African American parents use spanking produces better long-term outcomes than the typical way it is used in European American families. Ethnic differences in the apparent effectiveness of spanking may be due to its normative support within each ethnic group. More needs to be understood before imposing spanking prohibitions on ethnic groups that are under-represented among social scientists and policy makers.

168. G. R. Patterson & P. A. Fisher, *Recent Developments in Our Understanding of Parenting: Bidirectional Effects, Causal Models, and the Search for Parsimony*, in 5 HANDBOOK OF PARENTING: PRACTICAL ISSUES IN PARENTING 59, 74 (M. H. Bornstein ed., 2002).

169. Marie-Eve E. Clement & Claire Chamberland, *Physical Violence and Psychological Aggression Towards Children: Five-Year Trends in Practices and Attitudes from Two Population Surveys*, 31 CHILD ABUSE & NEGLECT 1001, 1006 (2007).

170. Baumrind et al., *supra* note 8, at 157, 178–83.

171. J. Douglas Bremner et al., *Development and Preliminary Psychometric Properties of an Instrument for the Measurement of Childhood Trauma: The Early Trauma Inventory*, 12 DEPRESSION & ANXIETY 1, 6 (2000); Martin H. Teicher et al., *Sticks, Stones, and Hurtful Words: Relative Effects of Various Forms of Childhood Maltreatment*, 163 AM. J. PSYCHIATRY 993, 994 (2006); Yvonne M. Vissing et al., *Verbal Aggression by Parents and Psychosocial Problems of Children*, 15 CHILD ABUSE & NEGLECT 223, 299 (1991).

172. REID ET AL., *supra* note 114, at 74.

173. Gunnoe & Mariner, *supra* note 53, at 768, 773; Lansford et al. (2004), *supra* note 155, at 801; Jodi Polaha et al., *Physical Discipline and Child Behavior Problems: A Study of Ethnic Group Differences*, 4 PARENTING: SCI. & PRACTICE 339, 348 (2004).

VI

CONCLUSIONS

Spanking must be understood in the broader context of the appropriate exercise of parental authority. Numerous studies have shown the superior effectiveness of authoritative parenting, especially compared with the extremes of authoritarian and permissive parenting.¹⁷⁴ We have recently extended that evidence by documenting that ten-year child outcomes vary greatly by these preschool parenting patterns.¹⁷⁵ Authoritative parents use their parental authority to empower their children's development. In contrast, authoritarian parents misuse their parental authority and permissive parents abdicate their parental authority. Authoritative parents combine nurturance, give-and-take communication, support for age-appropriate independence and autonomy, and firm confrontive discipline and maturity demands. Their use of confrontive discipline and maturity demands distinguish them from permissive parents, resulting in large differences in ten-year outcomes of those two parenting patterns.¹⁷⁶ Authoritarian parents, on the other hand, have low nurturance and use detrimental forms of power assertion, a combination associated with even worse ten-year outcomes than permissive parenting. The use of normative spanking did not distinguish authoritative parents from other parenting patterns, although it was used more by authoritarian parents than by permissive parents. We think that authoritative parenting can be implemented by some parents without the use of any spanking, but we have no evidence of that from our study, as all authoritative parents used spanking at least occasionally.

To support the firm control dimension of authoritative parenting, research must be capable of discriminating between effective and counterproductive corrective disciplinary actions. We suspect that *how* and *when* a disciplinary tactic is used will determine its effectiveness at least as much as whether the tactic is verbal, nonphysical, or physical. Multiple studies have shown that spanking is associated with adverse outcomes only when children perceive their parents as rejecting them.¹⁷⁷

In our study, the most detrimental forms of power assertion were verbal hostility and psychological control, which accounted for adverse outcomes in the full sample and were distinctive of authoritarian parents. Severe physical punishment and arbitrary discipline were also used more often by authoritarian parents and were associated with some long-term adverse outcomes, but did not

174. Steinberg, *supra* note 7.

175. Baumrind et al., *supra* note 8, at 157, 172–75.

176. *Id.* at 157, 172–76, 184.

177. Vonnie C. McLoyd & Julia Smith, *Physical Discipline and Behavior Problems in African American, European American, and Hispanic Children: Emotional Support as a Mediator*, 64 J. MARRIAGE & FAM. 40, 44 (2002); R. P. Rohner et al., *Effects of Corporal Punishment, Perceived Caretaker Warmth, and Cultural Beliefs on the Psychological Adjustment of Children in St. Kitts, West Indies*, 53 J. MARRIAGE & FAM. 681, 685 (1991).

predict those outcomes beyond what was accounted for by verbal hostility and psychological control.¹⁷⁸

The available evidence indicates that disciplinary reasoning is a crucial component of authoritative parenting and that children as young as two or three cooperate with reasoning more when it is backed up with time-out or privilege removal at least ten percent of the time.¹⁷⁹ Roberts' studies showed that even the most clinically defiant two- to six-year-olds will cooperate with time-out if enforced when necessary with an effective back-up tactic, such as a two-swat spank or room isolation. Skillful use of this sequence of increasingly forceful tactics can then lead to phasing out the back-up tactic as children learn to cooperate with time-out and pay more attention to their parents' verbal corrections. Therefore, some version of the sequencing used by Roberts' and other behavioral parent training programs could well be a process that produces well-behaved children whose parents rely primarily on reasoning and verbal correction.¹⁸⁰ To the extent this is the case, spanking prohibitions will inadvertently restrict the back-up options needed by some parents to enforce nonphysical tactics and reasoning. This may explain why some parents are at risk for extremely permissive parenting or for increased verbal hostility when they are prohibited from using spanking or equally effective back-up tactics.¹⁸¹

In this article we do not claim or imply that parents must use spanking to obtain compliance or that any kind of disciplinary punishment is necessary for all children. Parents should, however, retain the option to use spanking appropriately, unless they have abused that option. Current research indicates that customary spanking is not associated with child outcomes that are any more adverse than the outcomes of any other type of corrective discipline. The most empirically supported use for a two-swat spanking is when two- to six-year-olds respond defiantly to nonphysical disciplinary tactics, such as time-out, or when imposed to stop dangerous misbehavior. Spanking should never be used in an infant's first twelve months of life and rarely, if at all, before eighteen months of age. Parents should make sure their children know that any corrective discipline, including spanking, is motivated by love and concern for them. Parents must also be certain not to administer punishment too severely, whether physical or nonphysical. Finally, all punishment should be used in such a way that reduces the need to use it in the future. Every child is different, so not all disciplinary tactics will work as well with every child—or for every situation with the same child. Parents need to skillfully use a range of

178. Baumrind et al., *supra* note 8, at 157, 178–80.

179. Larzelere et al., *supra* note 30.

180. RICHARD Q. BELL & LAWRENCE V. HARPER, *CHILD EFFECTS ON ADULTS* (1977); Robert E. Larzelere, *Combining Love and Limits in Authoritative Parenting*, in *PARENTHOOD IN AMERICA* 85 (Jack C. Westman ed., 2001).

181. Clement & Chamberland, *supra* note 169, at 1001, 1006; Patterson & Fisher, *supra* note 168, at 74.

disciplinary options to help their children achieve their full potential, rather than to have effective options restricted unnecessarily.