

# Child Outcomes of Nonabusive and Customary Physical Punishment by Parents: An Updated Literature Review

Robert E. Larzelere<sup>1</sup>

---

This article updates the only previous systematic literature review of child outcomes of nonabusive and customary physical punishment by parents. The outcomes differ by methodologic, child, and subcultural factors as well as by how the physical punishment was used. All six studies that used clinical samples (including four randomized clinical studies) and all three sequential-analysis studies found beneficial outcomes, such as reduced noncompliance and fighting, primarily when nonabusive spanking was used to back up milder disciplinary tactics in 2- to 6-year olds. Five of eight longitudinal studies that controlled for initial child misbehavior found predominantly detrimental outcomes of spanking. However, those detrimental outcomes were primarily due to overly frequent use of physical punishment. Furthermore, apparently detrimental outcomes have been found for every alternative disciplinary tactic when investigated with similar analyses. Such detrimental associations of frequent use of any disciplinary tactic may be due to residual confounding from initial child misbehavior. Specific findings suggest discriminations between effective and counterproductive physical punishment with young children. More research is needed to clarify the role of spanking and alternative disciplinary tactics in control system aspects of parental discipline.

---

**KEY WORDS:** physical punishment; parental discipline.

Few social scientific topics involve such sharp contradictions as does parental spanking. Consider the following examples:

- About 94% of American 3- and 4-year olds have been spanked by a parent at least once during the past year (Straus & Stewart, 1999). Yet most clinical child psychologists would never suggest parental spanking, and one-third consider such a suggestion always unethical (Hyman, 1997; Schenck, Lyman, & Bodin, 2000).
- Several countries have banned all parental spanking, and others are considering such a ban. Yet if spanking were being evaluated as a psychological intervention, it would qualify as

“probably efficacious” (Lonigan, Elbert, & Johnson, 1998).

- Social scientists generally make sensitivity to cultural differences a top priority. Yet African-American families are often denigrated for spanking, even though most relevant research finds that moderate spanking has benign or beneficial outcomes in African Americans (Deater-Deckard & Dodge, 1997, and 11 commentaries; Gunnoe & Mariner, 1997; Whaley, 2000).
- A large number of social scientists consider the mildest spanking to be more harmful than alternative disciplinary tactics. Yet the only systematic review of nonabusive physical punishment found little evidence of differential harmfulness in direct comparisons with alternative disciplinary tactics (Larzelere, 1996).

Never before have social scientists advocated a total ban on a practice this widespread. Does the

---

<sup>1</sup>Psychology Department, Munroe–Meyer Institute, University of Nebraska Medical Center, and Girls and Boys Town. Correspondence concerning this article should be sent to Robert E. Larzelere, Psychology Department, Munroe–Meyer Institute, 685450 Nebraska Medical Center, Omaha, NE 68198-5450. Email: rlarzelere@unmc.edu.

available scientific evidence support a total ban on parental spanking, or does the evidence suggest an appropriate but limited role for nonabusive spanking in parental discipline? These are major questions underlying this review of the relevant scientific literature.

It is widely known that several methodologic problems are pervasive in much research on parental spanking. First, most research is cross-sectional and correlational, which cannot establish causation. Second, much research emphasizes overly severe forms of physical punishment. The controversy is about nonabusive physical punishment, not overly severe forms of physical punishment. Third, many studies incorporate spanking as one item in a broader measure of punitiveness. Such studies cannot determine the unique effects of spanking. Thus these three kinds of studies were excluded from this literature review.

It would have been ideal to focus exclusively on studies of explicitly nonabusive spanking. Because such studies are so few, this review also considers studies of customary spanking. These are studies that measure physical punishment without emphasizing the severity of its use.

This review updates the only previous systematic review of child outcomes of nonabusive or customary physical punishment by parents (Larzelere, 1996) and its unpublished predecessor (Lyons, Anderson, & Larson, 1993). It uses a qualitative "box-score" review for two reasons. First, a qualitative review clarifies whether outcomes vary by the causal conclusiveness of the designs, which is essential for conclusions about the *causal* effect of physical punishment. As noted by Miller and Pollock (1994), meta-analyses of causally inconclusive studies yield causally inconclusive summaries. Second, a qualitative review provides more details about the qualifying studies so that others can more easily critique the conclusions of this review. This is particularly important, given the controversial nature of the topic.

## METHOD

Several strategies were used to locate articles for this review. Articles from 1995 to February 2000 that addressed corporal punishment were identified using *PsychLit*, *Medline*, and references in those articles. In addition, all sources in two reviews of physical punishment's outcomes were considered as potential references (Gershoff, 1999; Larzelere, 1996). Finally, 21 leading research investigators were asked for re-

cent studies.<sup>2</sup> These procedures yielded 269 potential studies.

The selection criteria for relevant articles were as follows: First, a study had to be published in a peer-reviewed professional journal in the English language. Exceptions to this criterion were made for sources in the Gershoff (1999) review and for recent unpublished studies that otherwise qualified. Forty-one (41) of the 269 potential studies were ruled out because they included no original data on parental physical punishment. Second, studies had to have a child outcome variable for which beneficial versus detrimental outcomes were reasonably unambiguous. This eliminated 40 additional studies. This criterion excluded studies in which the only child outcome concerned the likelihood of using nonabusive physical punishment as a parent. Studies in which child outcomes concerned the child's subsequent use of overly severe or frequent physical punishment were included, however.

Third, a study had to include at least one measure of nonabusive or customary physical punishment by parents. This excluded measures dominated by severity or abusiveness (eliminating 37 more studies) and measures dominated by nonspanking tactics (e.g., washing a child's mouth out with soap, yelling, restraint, nonphysical punishment, which eliminated 39 additional studies). Fourth, the referent period for parental physical punishment had to precede the time period for the child outcome measure (eliminating 32 cross-sectional studies). Finally, the average age of the child when spanked had to be younger than 13 years (eliminating 42 otherwise eligible studies). This criterion also excluded retrospective studies unless their survey specified a referent age for spanking that averaged younger than 13 years.<sup>3</sup> Table I summarizes the 38 studies that met all five criteria.

Because two reviews of physical punishment came to rather different conclusions (Gershoff, 1999; Larzelere, 1996), it might be helpful to clarify which studies from Gershoff's (1999) review met the criteria

<sup>2</sup>Diana Baumrind, Kirby Deater-Deckard, Joan Durrant, Chris Ellison, Leonard Eron, James Garbarino, Elizabeth Thompson Gershoff, Anthony Graziano, George Holden, Irwin Hyman, John Knutson, Joan McCord, Vonnie McLoyd, Jerry Patterson, Kathy Ritchie, Mark Roberts, Rebecca Socolar, John Steley, Murray Straus, Claudio Violato, Gail Wallerstein.

<sup>3</sup>A prospective study that asked similar questions retrospectively at age 25 found that the retrospective data correlated most strongly with earlier maternal reports of physical punishment at ages 12 to 14 (Stattin, Janson, Klackenborg-Larsson, & Magnusson, 1995).

for this review and which criteria were not met by other studies in that review. Although the Gershoff (1999) review is still being revised for publication, it has been cited as "in press" for at least 2 years (e.g., Straus & Mouradian, 1998, citing an earlier version as Thompson, in press).

Only 16 of the 92 studies in Gershoff (1999) met the criteria for this review. Gershoff's (1999) other studies were excluded because of (i) the severity of the measure of physical punishment (17 studies), (ii) an overly broad measure of punishment (17 studies), (iii) a cross-sectional design (22 studies), (iv) a focus on physical punishment of teenagers (19 studies), or (v) no unambiguous child outcome variable (1 study). Thus, the present review focuses more specifically on causally relevant studies of nonabusive or customary physical punishment of young children than does the Gershoff (1999) review.

For the purposes of summarizing the data, a finding was counted as a beneficial outcome if physical punishment predicted a desirable outcome in the child (e.g., improved compliance) at the 0.05 significance level. A finding was counted as a detrimental outcome if physical punishment significantly predicted an undesirable outcome in the child (e.g., lower self-esteem, more delinquency). A study was summarized as finding predominantly beneficial or predominantly detrimental outcomes if (i) its only relevant significant outcome was in the specified direction, (ii) one of its two or three relevant analyses was significant and in the specified direction, or (iii) at least two of its relevant analyses were significant and in the specified direction. Otherwise, a study was summarized as showing a neutral outcome.

## RESULTS

Overall, the 38 qualifying studies (Table I) divided almost equally into predominantly beneficial child outcomes (32%), predominantly detrimental outcomes (34%), and neutral or mixed outcomes (34%). The following sections consider how various study characteristics discriminate beneficial outcomes from detrimental outcomes. Relevant study characteristics include methodologic characteristics, type of outcome, child characteristics, how the physical punishment was used, and the cultural context. Because different characteristics are often confounded with each other, a subsequent section highlights the 11 studies that are most relevant for untangling common confounds among these charac-

teristics. The final part of the Results section then compares the outcomes of physical punishment with those of alternative disciplinary techniques in these 38 studies.

### Research Design: Causal Conclusiveness

The child outcomes of physical punishment varied dramatically by the type of research design (see Table I). All 6 clinical studies (including 4 randomized clinical trials) found predominantly beneficial child outcomes, as did the 3 studies using sequential analyses. The beneficial outcomes included reduced noncompliance and fighting and, in one study, enhanced parental affection. However, 5 of the 8 controlled longitudinal studies found a broader range of predominantly detrimental child outcomes of physical punishment, and the remaining 3 studies found neutral outcomes or a mixture of both beneficial and detrimental outcomes. Almost half (47%) of the 15 uncontrolled longitudinal studies found predominantly detrimental child outcomes, and the rest (53%) found neutral outcomes. Two (2) of the 6 retrospective studies found beneficial outcomes, 1 found detrimental outcomes, and the remaining 3 found neutral outcomes.

Of the 12 retrospective studies in Larzelere's (1996) earlier review that did not qualify for this updated review, two-thirds found predominantly detrimental outcomes, and the other one-third found neutral outcomes. Thus detrimental outcomes have previously been most common in retrospective studies, especially when the referent child age is either unspecified or averaged 13 years old or more.

In general, the stronger the causal conclusiveness of the design, the more likely a study was to find beneficial child outcomes. There were, however, important exceptions to that overall tendency. Controlled longitudinal studies were the most causally conclusive studies that found mostly detrimental outcomes. Those studies differed from the randomized clinical trials on many other dimensions. The clinical studies tended to focus on using a 2-swat spank to enforce short-term compliance with time out in clinically referred 2- to 6-year olds. The controlled longitudinal studies were based on maternal reports of spanking frequency in older, nonclinical samples, using the following long-term outcomes: antisocial behavior, fighting, hostility, emotional problems, competence, and self-esteem. (A later section returns to these studies to begin untangling these confounds.)

**Table I. Studies of Child Outcomes of Nonabusive or Customary Physical Punishment by Parents**

Authors	Age <sup>a</sup>	N and gender <sup>b</sup>	Type of sample	Physical punishment (parent) <sup>c,d</sup>	Other discipline responses	Child outcome measures	Physical punishment association <sup>e</sup>	Summary of findings <sup>a,d</sup>
<b>Randomized clinical trials</b>								
<i>Preschoolers</i> Bean & Roberts (1981)	2-6	24 B,G	Clinic-referred, disruptive problems	2-swat spank back-up for time out (TO) (M)	Child release from TO, non-time-out control	Compliance with parent commands	+	Percentage compliance improved more with spank back-up than in either of other two conditions. Fewer TOs than Child Release group.
Day & Roberts 1983	2-5	16 B,G	Clinic-referred, disruptive	2-swat spank back-up for TO (M)	Room time out back-up for chair TO (barrier)	Compliance with parent commands	+	Spank back-up and room time out back-up equally effective in increasing compliance.
Roberts (1988)	2-6	18 B,G	Clinic-referred, disruptive	2-swat spank back-up for TO (M)	Room time out back-up for TO (barrier)	Compliance to parent commands, escapes from time out chair, verbalizations	+	The two back-ups equally effective in reducing escapes from time out and in increasing compliance. Room time-out back-up had fewer verbalizations in first chair time out.
Roberts & Powers (1990)	2-6	36 B,G	Clinic-referred, disruptive	2-swat spank back-up for TO (M)	Room TO, child release, hold as TO backups	Compliance, TO escapes, number of chair TOs needed	+	Spank and room TO yielded fewer escapes from TO and fewer TOs than hold or child release back-ups. Fifteen children needed alternative back-up (switch to spank in 12 cases), which increased compliance.
<b>Other clinical studies</b>								
<i>Preschoolers</i> Roberts (1982)	2-6	32 B,G	Clinic-referred, disruptive	2-swat spank back-up for TO (M)	None	Escapes from time out	+	Reduced mean number of escapes from time out from first to second time out attempt.
<i>School age</i> Bernal <i>et al.</i> (1968)	8	1 B	Clinic-referred, disruptive	Good, hard whack on rear (M)	Ignore	Compliance, aggression, liking of son	+	Increased compliance, decreased aggression, increased liking for son.
<b>Longitudinal studies controlling for initial misbehavior</b>								
<i>Preschoolers</i> Baumrind & Owens (2000)	3-5 (T1), 7-11 (T2), 12-16 (T3)	87-135 B,G	Children in pre-schools	Frequency of physical punishment from ratings of interviews and home observations (B)	Verbal punishment	General competence, cognitive competence, social agency, communal competence, total problems, externalizing problems, internalizing problems	-/0	After controlling for initial level of externalizing problems and cooperation, T1 spanking predicted lower general competence at T2 and T3, lower cognitive competence at T2, and lower communal competence at T2. Spanking frequency at T2 predicted lower general competence at T3 with the same control. Then, after dropping the 5% or 7% most frequent/severe cases of physical punishment, only one of those remained significant (T1 spanking and T2 general competence). This became nonsignificant after controlling also for either positive parenting or maladjusted parenting. Using equivalent analyses, T1 verbal punishment predicted detrimental outcomes in five T3 outcomes, four of which held up after dropping extreme verbal punishment cases.
Ellison <i>et al.</i> (1998)	2-4 (T1), 9-11 (T2)	467 B,G	Nationally representative sample	Frequency of spanking in previous week at T1 and T2 (mother report) (B)	None	Antisocial behavior, hostility, emotional problems	+/-	In conservative Protestant families, children who were spanked only at T1 had fewer T2 problems than children who were not spanked at either time. In other families, there were increased T2 problems for children spanked at both times, but no effects for those spanked at T1 but not at T2.

Gumroe & Mariner (1997)	4-11 (T1), 9-16 (T2)	1112 B,G	Nationally representative sample	Frequency of spanking in previous week (mother report) (B)	None	Antisocial behavior, youth-reported fighting at school	+/-	Spanking frequency predicted a decrease of subsequent fighting 5 years later for (1) African Americans, (2) 4-to 7-year olds, and (3) girls. Spanking frequency predicted an increased frequency of fighting 5 years later for (1) European Americans and (2) 8- to 11-year olds. It also predicted higher mother-reported antisocial behavior 5 years later, replicating Straus <i>et al.</i> (1997). All analyses controlled for T1 bullying.	
Larzelere <i>et al.</i> (1998b)	2-3 (T1), 4 (T2)	38 B,G	Volunteers in response to newspaper article	Slap hand or spank when used in combination with reasoning (mother report) (M)	Reasoning, non-physical punishment, combinations of the above, other	Delay until next recurrence of fighting or disobedience following subsequent reasoning by itself, disruptive behavior at T2	+/0	Backing up reasoning with either kind of punishment enhanced the subsequent effectiveness of reasoning by itself in delaying the next recurrence of either misbehavior. This effect occurred for a nonphysical back-up in 8 of 8 analyses and for physical punishment in 4 of 8 analyses. Children whose mothers used reasoning by itself frequently and also backed up reasoning with nonphysical punishment showed the largest decrease in T2 disruptive behavior. The same effect for physical punishment as a back-up was not significant. Reasoning without punishment predicted more T2 disruptive behavior.	
<i>School age</i> Adams (1995)	6-12 (T1), 10-16 (T2)	367 B,G	National sample with some oversampling/	Spanked twice or more in previous week or spanked once plus other endorsement of spanking (mother report) (M)	None	Academic achievement in reading and math, behavior problems, global self-esteem, scholastic self-esteem	-/0	Frequent spanking group had lower global self-esteem in boys and for all children. No significant differences on any of the other four outcome variables for boys. Nothing was significant for girls. Controlled for ethnicity; nurturance; cognitive enrichment; poverty; self-esteem and behavior problems; and other variables.	
Larzelere & Smith (2000)	6-9 (T1), 8-11 (T2)	981 B,G	National sample with some oversampling/	Spanking frequency in previous week (mother report) (M)	Grounding, sending to room, removing allowance, removing privileges in previous week	Antisocial behavior	-/0	The frequency of spanking or of any other the four alternatives each predicted a higher rate of antisocial behavior 2 years later, controlling for the original level of antisocial behavior. Predictions of T2 antisocial from each disciplinary tactic became nonsignificant when T1 externalizing problems were measured more thoroughly. Controlled for five other variables and interactions.	
McLeod <i>et al.</i> (1994)	4 or more (T1), 6 or more (T2)	1866 B,G	National sample with some oversampling/	Spanking frequency in previous week (mother report) (M)	None	Antisocial behavior	-/0	Spanking frequency at Time 2 predicted antisocial behavior at Time 2 in European Americans but not in African Americans, controlling for Time 1 antisocial behavior. T1 antisocial behavior predicted increased spanking in both races.	
Straus <i>et al.</i> (1997)	6-9 (T1), 8-11 (T1)	807 B,G	National sample with some oversampling/	Spanking frequency in previous week (mother report) (M)	None	Antisocial behavior	-	Spanking predicted more antisocial behavior 2 years later, controlling for initial level of antisocial behavior as well as five other demographic and parenting-quality measures and their interactions with spanking, due mostly to those spanked 3+ times weekly. Correlations similar for four other cohorts, including 3- to 5-year olds and children 10 or older.	
<b>Sequential analyses</b>									
<i>Preschoolers</i> Larzelere & Merenda (1994)	2-3	38 B,G	Volunteers in response to newspaper article	Slap hand or spank (mother report) (M)	Reasoning, non-physical punishment, combinations of above, and other	Delay until next recurrence of disobedience or fighting	+	Either type of punishment without reasoning delayed the next disobedience recurrence the most at high levels of child distress. Reasoning delayed disobedience recurrences the most at intermediate levels of child distress. Either type of punishment with reasoning at an intermediate child distress was the best at delaying disobedience recurrences overall.	
Larzelere <i>et al.</i> (1996)	2-3	38 B,G	Volunteers	Slap hand or spank (mother report) (M)	Reasoning, non-physical punishment, combinations of above, and other	Delay until next recurrence of disobedience or fighting	+	The combination of either physical or nonphysical punishment with reasoning was associated with a longer delay until the next misbehavior recurrence than following reasoning alone, punishment alone, or other discipline responses.	

**Table I. (Continued)**

Authors	Age <sup>a</sup>	N and gender <sup>b</sup>	Type of sample	Physical punishment (parent) <sup>c,d</sup>	Other discipline responses	Child outcome measures	Physical punishment association <sup>e</sup>	Summary of findings <sup>o,d</sup>
Ritchie (1999)	3	90 B,G	Recruited from birth records	Slap hand or spank (mother report) (M)	Ignore, power assert physically, power assert verbally, time out, threaten, remove privilege, reasoning, offer alternative	Ability to stop a non-compliant tactic suggested with (1) greater likelihood of using the tactic in response to a given child tactic (e.g., defiance) when (2) the disciplinary tactic was not followed by a significantly increased probability of the same child tactic; compliance	+	Time out and spank were the only tactics that mothers (1) were more likely to use for defiance and (2) that was not followed by an inflated probability of defiance. This pattern occurred for other disciplinary tactics only for ignore in response to whining and for privilege removal in response to simple refusals. Compliance occurred at higher rates following only ignore and power assert physically.
<b>Prospective longitudinal studies</b>								
<i>Preschoolers</i> Crowne <i>et al.</i> (1969)	5 (T1), 18 (T2)	83 B,G	Kindergarten	Frequency of spanking by each parent, extent of physical punishment (B)	Deprivation of privileges, withdrawal of love	Realistic aspirations (vs. overcautious or unrealistic), unusual shifts in aspirations	0	Mother's spanking frequency associated with realistic aspirations more than with unrealistic aspirations (females). Mother's spanking frequency +r with number of unusual shifts in aspirations. Love withdrawal +r with overly cautious aspirations (males).
Deater-Deckard & Dodge (1997)	1-5 (T1), 5.5-12 (T2 to T8)	585 B,G	Kindergarten students	No vs. mild to harsh physical discipline, distinguished from physical abuse (M)	Harsh discipline, ranging from nonrestrictive and positive to strict, severe, often physical	Externalizing problems (teacher rating)	-/0	Nonabusive mild to harsh physical punishment predicted lower externalizing problems in African Americans compared to children of nonspankers, whereas it predicted higher externalizing problems in European Americans (significance not given). The harsh discipline variable predicted subsequent externalizing problems in European Americans, but not in African Americans, unless African-American parents were nonnurturing.
Deater-Deckard & Pilon <i>et al.</i> (1996)	4 (T1), 4.5-7.5 (T2)	566 B,G	Kindergarten	Mother's spontaneous report of PP as responses to 5 vignettes (M)	None	Externalizing problems (teacher report), aggression and conflict with teacher (peer report)	-/0	Physical punishment +r with all 3 measures of acting out for European Americans, physical punishment -r with all three measures for African Americans, but nonsignificantly so.
Grinder (1962)	5 (T1), 11-12 (T2)	140 B,G	Kindergarten	Extent of physical punishment (M)	Isolation, love withdrawal, privilege deprivation	Resistance to temptation to cheat for prize in lab game	0	Physical punishment and other disciplinary responses were each unrelated to resistance to temptation.
Johannesson (1974)	9-24 mos. (T1), 10-13 (T2)	212 B,G	Birth cohort over 3 years	Extent of spanking from "never" to "constantly" (M)	None	Teacher's rating of tendency to start a quarrel and have conflict with peers	0	Spanking was not associated with subsequent school aggression.
McClelland & Pilon (1983)	5 (T1), 31-32 (T2)	78 B,G	Kindergarten	Extent of physical punishment (M)	Deprivation of privileges	Need for achievement, need for affiliation, need for power	0	Physical punishment -r with need for power for males only; deprivation of privileges unrelated to all adult needs.
McCord & Ensminger (1997)	6 (T1), 32 (T2)	953 B,G	Children in African-American inner city	Mother's report as to whether their child was spanked twice a week or more (M)	None reported	Criminal violence, depression, alcoholism	0/-	Frequent spanking predicted a greater probability of the three adulthood problems in only 1 of 6 analyses (only increased alcoholism for men; none of three outcomes were increased for women).
Michels <i>et al.</i> (1993)	M = 5.3 (T1), M = 5.9 (T2)	342 B,G	Kindergarten children	Parents' report that physical punishment worked best, Frequency of physical punishment (B)	Attention, time out, deprivation of privileges, reason, bribe/threaten, formal behavior management system	Disruptive classroom behavior (e.g., hitting, defiance, disruptiveness) as reported by the teacher	-	Children whose parents thought physical punishment worked best were more disruptive in the classroom than those who thought alternatives worked best or who considered discipline not a problem or who said nothing worked. Those physically punished weekly or more were more disruptive than those hit less often. Among those not physically punished, frequency of nonphysical punishment was not associated with disruptive behavior.

Sears (1961)	5 (T1), 12 (T2)	160 B.G	Kindergarten	Extent of physical punishment (M)	Deprivation of privileges, love withdrawal	Five types of aggression (e.g., antisocial, projected)	0	Of 10 correlations with later aggression, physical punishment +r only with prosocial aggression in girls. The other discipline responses also had one more +r than -r. Physical punishment not associated with subsequent self-esteem.	
Sears (1970)	5 (T1), 12 (T2)	160 B.G	Kindergarten	Extent of physical punishment (M)	None	Self-esteem	0	Girls in the high physical punishment group at both 12 and 36 months were lower on IQ than were those not in that group at both ages. The other 11 pairwise comparisons were not significant.	
Smith & Brooks-Gunn (1997)	12 & 36 mos. (T1), 36 mos. (T2)	715 B.G	Low-birth-weight children	Mother reported 2 or more physical punishments in past week or seen hitting or scolding (M)	None	IQ	0/-	Spanked children were higher than the 11 nonspanked/nonabused children on total aggression and reactive aggression, but not on bullying or instrumental aggression. Spanking frequency not related to any aggression measures.	
Strassberg et al. (1994)	4 (T1), 4.5 (T2)	273 B.G	Kindergarten	Spanking frequency during past year, including spanking with object (B)	Parental violence	Observed bullying and reactive, instrumental, and total aggression at school	-	Spanking frequency predicted poorer cognitive achievement at T3, controlling for cognitive measures at T1, age, ethnicity, gender, mother's education, father presence, cognitive stimulation, emotional support, etc. Reactively replicated without 1-year olds in the sample.	
Straus & Paschall (1998)	1-4 (T1), 3-6 (T2), 5-9 (T3)	960 B.G	National sample with some oversampling/	Frequency of spanking during previous week at T1 and T2 plus observed hitting (M)	None	Cognitive achievement in reading and math	-	Physical punishment was associated with more boys' dependency (desired attention and closeness) toward teacher, but not associated with aggression in either gender (except physical punishment for sibling aggression predicted lower aggression in total sample and tended toward significance for boys). Other disciplinary tactics also generally unrelated to outcomes.	
Yarrow et al. (1968)	M = 4.2 (T1), M = 4.4 (T2)	58 B.G	Nursery school children	9-point scale from little to very often used, rated from interview about six extreme examples of disobedience (M)	Isolation, reasoning, privilege removal, love withdrawal, isolation, scolding, diverting attention, modeling	Aggression and dependency as rated by two teachers in the nursery school	-/0	Spanking frequency predicted poorer cognitive achievement at T3, controlling for cognitive achievement at T1 plus controls for age, ethnicity, gender, mother's education, father presence, cognitive stimulation, emotional support, etc.	
<i>School age</i>									
Paschall & Straus (1998)	5-9 (T1), 7-11 (T2), 9-13 (T3)	711 B.G	National sample with some over sampling/	Frequency of spanking during previous week at T1 and T2 plus observed hitting (M)	None	Cognitive achievement in reading and math	-	Spanking frequency predicted more frequency and severity of physical punishment, were more likely to be hit when behaving, and were less likely to know why they were physically punished.	
<b>Retrospective studies</b>									
<i>Preschoolers</i>									
Rhue & Lynn (1987)	3-5 & 6-10 (T1), college age (T2)	53 B.G	Top and bottom 4% of college students on fantasy proneness + middle group	Severity, frequency of early physical punishment; hit when behaving; knew why they were physically punished (B)	None	Fantasy proneness	-	Percentage of punishment that was physical predicted lower rebelliousness, controlling for parental acceptance or parent or child orientation for physical punishment. Low rebelliousness could be reframed as extreme conformity. Physical punishment was unrelated to neuroticism and introversion.	
<i>School-age</i>									
Alibrando (1987)	7(T1), M = 22 (T2)	96 B.G	College students	Percentage of punishment that was physical, frequency of spanking or physical punishment (B)	None analyzed	Rebelliousness (vs. conformity), neuroticism, introversion (vs. extroversion)	+/0		

**Table 1. (Continued)**

Authors	Age <sup>a</sup>	N and gender <sup>b</sup>	Type of sample	Physical punishment (parent) <sup>c,d</sup>	Other discipline responses	Child outcome measures	Physical punishment association <sup>e</sup>	Summary of findings <sup>a,f</sup>
Holmes & Robins (1988)	6-13 (T1), 18-49 (T2)	200 B,G	Alcoholics, depressives, controls	Spanking and slapping used frequently or as usual mode of discipline (B)	Punched, hit with belt or stick, confinement, scolding, isolation, beatings	Differences in three groups (alcoholics or depressives vs. controls)	0	Groups similar in percentage that had experienced spanking frequently or as usual mode of punishment. Controls less likely to have been hit with object frequently or beaten, less likely to fear parental discipline, or to be punished in public. Their mothers were more likely to carry out their threats and share discipline responsibilities with their husbands.
Larzelere <i>et al.</i> (1989)	Before 13 (T1), <i>M</i> = 21 (T2)	157 B,G	Home economics college students	Frequency of spanking (before 13) (B)	Alternative punishment factor (time out, removing privileges, and restitution)	Self-esteem	0	Spanking unrelated to self-esteem, as was the alternative punishment factor. Parent-oriented motivation for spanking predicted low self-esteem.
MacIntyre & Cantrell (1995)	Before 13 (T1), <i>M</i> = 24 (T2)	240 B,G	College students	Major punishment style: physical punishment (No, with explanation, with verbal assault, or spontaneously) (B)	None	Approval of aggression toward objects or people	0	Major physical punishment styles prior to age 13 unrelated to approval of aggression toward either objects or people.
Tennant <i>et al.</i> (1975)	0-14 (T1), <i>M</i> = 23 (T2)	5044 B	U.S. Army soldiers in Germany	Punishment: spanking (Yes or no? or most common?) (B)	Noncontact punishment (e.g., stand in corner), punishment frequency	Usage of three drugs and alcohol (Never, occasional, or frequent)	+	Drug nonusers were more likely to report spanking than were frequent drug users, for three drugs. Amphetamine nonusers were less likely to report noncontact punishment than frequent users. No association with alcohol usage. Punishment frequency predicted more frequent use of two drugs.

<sup>a</sup>T1 = First data collection period (Time 1), T2 = second data collection period (Time 2), T3 = third data collection period (Time 3)

<sup>b</sup>B = boys, G = girls.

<sup>c</sup>M = mother, B = both mother and father.

<sup>d</sup>TO = time out, +r = positively associated (i.e., correlated), -r = negatively associated.

<sup>e</sup>+ = Predominantly beneficial child outcomes of physical punishment; - = predominantly detrimental child outcomes; 0 = generally neutral child outcomes; +/- = two or more beneficial child outcomes and two or more detrimental child outcomes; 0/- = generally neutral outcomes, but one outcome was detrimental out of four or more relevant analyses; -/0 = predominantly detrimental outcomes with important exceptions (e.g., disappears with controls, does not generalize across genders or ethnic groups); +/-0 = predominantly beneficial outcomes with important exceptions.

<sup>f</sup>Oversampling of young, single mothers; lower socioeconomic status; and ethnic minorities.



Uncontrolled longitudinal studies and retrospective studies found detrimental outcomes more often than beneficial outcomes. However, neither of those designs can rule out the possibility that higher levels of child misbehavior caused both the increased physical punishment and the subsequent detrimental child outcomes (Straus, Sugarman, & Giles-Sims, 1997). Because a primary purpose of this review is to clarify the causal effects of physical punishment, the remainder of the results section generally considers only the 17 studies that are most causally relevant. Thus the following sections exclude uncontrolled longitudinal and retrospective studies except for supplementary information.

The remaining sections sometimes include more than 17 findings. Studies contribute more than one finding when their results differ by the dimension summarized or when they have both beneficial and detrimental outcomes. For example, a study that investigated three types of child outcomes (i.e., behavior problems, mental health, and competence) would be counted three times when considering how the results vary by the type of outcome.

### Methodologic Issues

This section summarizes how the results of these 17 causally relevant studies varied by two methodologic variables: the type of data and whether the outcome was short or long term.

Beneficial child outcomes were most likely when the outcome measure was observational (6 of 6 studies) or a specific daily maternal report (3 of 3 studies). Detrimental child outcomes were most likely when all the measures in a study were based on the same person's global reports (5 of 7 results) or on different persons' global reports (3 of 4 results). Gunnoe and Mariner (1997) found mixed beneficial and detrimental outcomes when different sources were used for the spanking and outcome variable, but a uniformly detrimental outcome when basing all measures on maternal reports. Overall, beneficial outcomes were more likely for more objective, specific outcome measures.

On the second methodologic issue, beneficial child outcomes were more likely for outcomes measured less than 6 months after the use of physical punishment (9 of 10 findings). Detrimental outcomes were more likely for outcomes that were 6 months or more after the physical punishment (7 of 10 findings). The short- vs. long-term status of the outcomes was

confounded with several other dimensions, addressed in a later section.

### Substantive Issues

The findings in the studies in Table I varied by several substantive characteristics as well as by methodologic dimensions. These included the type of outcome, the age and clinical status of the child, how the physical punishment was used, and the larger cultural context.

#### *Type of Child Outcome*

The apparent effects of physical punishment were generally beneficial in reducing noncompliance and fighting and in enhancing parental warmth and milder disciplinary tactics. However, the apparent effects of physical punishment were generally detrimental in increasing externalizing behavior problems and mental health problems, and in reducing competencies.

All the studies that investigated noncompliance found that physical punishment reduced it. This included all 6 clinical studies, all 3 sequential studies, and the sequential analysis part of 1 controlled longitudinal study (Larzelere, Sather, Schneider, Larson, & Pike, 1998b).

Four of five findings on fighting indicated that it was reduced by previous physical punishment. This included the single-case clinical study (Bernal, Duryee, Pruett, & Burns, 1968), two sequential-analysis studies (Larzelere & Merenda, 1994; Larzelere, Schneider, Larson, & Pike, 1996), and four of seven significant findings in a controlled longitudinal study (Gunnoe & Mariner, 1997). Only Gunnoe and Mariner (1997) found that physical punishment increased subsequent fighting, and then only in three of seven significant findings (e.g., for 8- to 11-year olds and European-Americans as a subgroup).

A few findings suggest that physical punishment can enhance parental nurturance and the effectiveness of milder disciplinary tactics. The single-case study in Table I (Bernal *et al.*, 1968) found that spanking enhanced the mother's liking for her clinically disruptive 8-year-old son. All 6 clinical studies and Larzelere *et al.* (1998b) found that spanking was effective in enhancing the subsequent effectiveness of milder disciplinary tactics, such as time out or reasoning. This enhanced effectiveness of milder disciplin-

ary tactics should in turn prevent disciplinary problems from eroding parental nurturance toward the child. Further research is needed, however, to clarify that speculation.

In contrast to these generally beneficial outcomes, physical punishment tended to predict higher rates of externalizing problems, mental health problems, and lower competencies in the 17 causally relevant studies in Table I. Five (5) of 10 findings indicated that physical punishment significantly increased subsequent externalizing problems (Ellison, Musick, & Holden, 1998; Gunnoe & Mariner, 1997; Larzelere & Smith, 2000; McLeod, Kruttschnitt, & Dornfeld, 1994; Straus *et al.*, 1997). These detrimental outcomes were neutralized or reversed in African Americans (Gunnoe & Mariner, 1997; McLeod *et al.*, 1994; but see also Straus *et al.*, 1997) and conservative Protestants (Ellison *et al.*, 1998), two subcultures that view spanking more normatively.

The most consistent causally relevant evidence of detrimental outcomes of physical punishment has been a subsequent increase in a 6-item antisocial measure from controlled longitudinal studies of the National Longitudinal Study of Youth (NLSY; Larzelere & Smith, 2000; McLeod *et al.*, 1994; Straus *et al.*, 1997). It is therefore important to note that (i) similar increases in antisocial behavior were associated with each of four other disciplinary tactics included in the NLSY survey (Larzelere & Smith, 2000) and that (ii) all of these apparently detrimental outcomes became nonsignificant when the measure of initial child externalizing problems was expanded to a 16-item measure (Larzelere & Smith, 2000). The other four disciplinary tactics in the NLSY survey were removing privileges, removing allowance, sending to room, and grounding (Larzelere & Smith, 2000).

Two of four findings indicated that physical punishment predicted increased mental health problems. This included lower self-esteem (Adams, 1995) and increased hostility and emotional problems in part of the sample in Ellison *et al.* (1998). Both detrimental findings involved high spanking frequencies; twice a week for Adams' 6- to 12-year olds or weekly spanking at ages 9 to 11 in Ellison *et al.* (1998). For conservative Protestants, however, children who were spanked at ages 2 to 4 but not at ages 9 to 11 had lower hostility and emotional problems than children not spanked at either age (Ellison *et al.*, 1998). Baumrind and Owens (2000) found neutral outcomes on internalizing problems.

Finally, one of two studies found that physical punishment predicted lower subsequent compe-

tences. Baumrind and Owens (2000) found that spanking frequency at ages 3 to 5 predicted lower subsequent competencies in cognitive, communal, and general areas. All but one of these effects became nonsignificant after those who used spanking most frequently and severely were dropped from the analyses (5–7% of the sample).<sup>4</sup> Adams (1995) found no effect on academic achievement.

### *Child Characteristics*

The outcomes of physical punishment also varied by the child's age and whether the child was clinically disruptive. Studies of children averaging 6 or younger in Table I generally found beneficial outcomes, whereas studies of older children generally found detrimental outcomes. In children with mean ages under 6, 11 of 12 studies (92%) found predominantly beneficial outcomes of physical punishment, whereas the remaining study (8%) found predominantly detrimental outcomes. In children averaging from 7.5 to 10 years old, 6 studies (86%) found predominantly detrimental outcomes and only 1 study (14%) found beneficial outcomes. In the previous review, all 4 studies of the physical punishment of teenagers found detrimental outcomes (Larzelere, 1996). Since that review, the first controlled longitudinal study of spanking of teenagers found increased rates of dating-partner abuse associated with such spanking,  $p = 0.06$  with complete statistical control for initial delinquency (Simons, Lin, & Gordon, 1998; R. Simons, personal communication, 1999).

The results also varied by whether the child had clinical levels of disruptive behavior. All 6 clinical studies found beneficial outcomes of spanking, whereas the other 13 findings were evenly split between predominantly detrimental outcomes (54%) and predominantly beneficial outcomes (46%). The effect of spanking frequency in Straus *et al.* (1997) depended on initial level of antisocial behavior in 3 of the 5 cohorts (Straus, personal communication, 1996). In those 3 cohorts, spanking decreased subsequent antisocial behavior in the initially most antisocial group, but it increased antisocial behavior in the least antisocial group.

<sup>4</sup>In their publication, Baumrind and Owens (2000) are planning to control for some parenting characteristics before controlling for initial externalizing/cooperative behavior. The summary here applies to their data when controlling only for initial externalizing/cooperative behavior with and without the most frequent and severe users of physical punishment.

*How Physical Punishment Was Used*

The outcomes of physical punishment in the 17 strongest studies in Table I also varied by how physical punishment was used. Child outcomes tended to be beneficial when physical punishment was used nonabusively, not too frequently, primarily as a back-up to milder disciplinary tactics, and flexibly.

Not one of the 17 causally relevant studies found predominantly detrimental outcomes if they did anything to rule out parents who used physical punishment too severely. To be included in this review, the studies could not emphasize severity in their measure of physical punishment, but only 2 of the controlled longitudinal studies made any attempt to rule out abusive physical punishment. The 6 clinical studies and 2 of the 3 sequential analysis studies made some attempt to exclude overly severe physical punishment. Nine (9) of the 10 findings that ruled out abuse were predominantly beneficial, and the remaining study had generally neutral findings after dropping the 5% to 7% of their sample that used physical punishment most frequently and severely (Baumrind & Owens, 2000). Of the studies that did not rule out abuse, most findings (7 of 11) indicated detrimental outcomes, 2 indicated beneficial outcomes, and 2 showed neutral outcomes.

Studies that emphasized the severity of physical punishment found detrimental outcomes, but they have been excluded from this review (e.g., Weiss, Dodge, & Bates, 1992). Straus and Mouradian (1998) addressed this issue in an important cross-sectional study. Mothers were asked, "When you had to spank or hit [your child], how often did you spank because you were so angry that you 'lost it'?" (p. 357). Mothers who reported "losing it" half of the time or more showed a much stronger association between frequency of spanking and the child's antisocial and impulsive behavior than those never losing it. For mothers who never lost control due to anger, their frequency of spanking was correlated about zero with the child's antisocial and impulsive behavior.

Other studies have investigated the effects of mild or moderate physical punishment and the effects of severe physical punishment in the same study. The types of severe physical punishment with detrimental outcomes included (i) whipping, punching, kicking, or beating up (Bryan & Freed, 1982); (ii) beatings (Holmes & Robins, 1988); (iii) yelling, throwing things, or attempting to injure someone when frustrated or annoyed (McCord, 1988); and (iv) hitting or beating up a child (Strassberg, Dodge, Pettit, &

Bates, 1994). None of these studies found detrimental outcomes for nonabusive or customary physical punishment, except for Strassberg *et al.* (1994). In that study, the 11 children (4%) who were not spanked during the year preceding kindergarten were less aggressive during kindergarten than the 96% who had been spanked. All other studies in Table I that excluded abusive parents found beneficial or neutral outcomes of nonabusive physical punishment.

Detrimental outcomes were associated with overly frequent physical punishment as well as overly severe physical punishment. The 8 controlled longitudinal studies all investigated spanking frequency or an approximation of it. Linear associations between spanking frequency and subsequent outcomes may be significant due solely to the most frequent spankers. Most (5 of 8) of the controlled longitudinal studies found detrimental child outcomes, but they all tested linear associations of spanking frequency. Three (3) of the controlled longitudinal studies provided information on various spanking frequencies. They indicated that the detrimental outcomes of physical punishment did not become significantly different from those not spanked until the spanking frequency reached one to three times a week (6- to 12-year olds: Adams, 1995; 3- to 5-year olds: Baumrind & Owens, 2000; 6- to 9-year olds: Straus *et al.*, 1997).

In contrast, spanking had predominantly beneficial outcomes when it was used conditionally, primarily to back up milder disciplinary tactics. All 9 studies that emphasized spanking as a back-up for a milder tactic found beneficial child outcomes. First, the series of clinical studies by Roberts and colleagues specified only one use for spanking: a back-up for noncompliance with the time-out procedure. The beneficial outcomes from the Larzelere series of studies occurred primarily when spanking was used in combination with reasoning, usually with the spanking coming after the reasoning. Both sets of studies found that the milder disciplinary tactic became more effective by itself after being backed up by spanking. A conditional spanking back-up was also used by Bernal *et al.* (1968) that led to improved compliance and parental affection and reduced fighting.

The conditional use of spanking after other tactics have failed is also consistent with how mothers change their tactics during extended disciplinary incidents with 3-year olds (Ritchie, 1999). Mothers were more likely to use verbal commands and reasoning and offer alternatives early in an extended disciplin-

ary incident. The longer an extended disciplinary incident lasted, however, the more likely mothers were to impose negative consequences such as physical power assertion, time out, a spank, or privilege removal. Other tactics, such as threatening, consenting, and ignoring, became more frequent during the middle of an extended incident, but then decreased subsequently. Putting those patterns together, mothers tend to use mild verbal tactics at first. Then they decide whether to consent, ignore, or continue to insist on child cooperation. Finally, they implement negative consequences or other power assertion as a last resort in an extended incident. Resorting to spanking too quickly might lead to its overly frequent use and detrimental outcomes. In contrast, skilled use of spanking as an occasional back-up for milder disciplinary tactics with 2- to 6-year olds is more effective.

Finally, flexible use of nonabusive spanking and alternatives is associated with better child outcomes than primary or exclusive use of physical punishment. Flexible use is best illustrated in the study of clinically noncompliant 2- to 6-year olds by Roberts and Powers (1990). They investigated four alternative approaches to enforcing compliance with time-out: a 2-swat spank, a 1-minute room isolation, a physical hold, and a child-determined release from time out. The first two back-ups for time out were equally effective overall, and more effective than the last two procedures. Some clinically defiant preschoolers would persist in their noncompliance to time-out even after repeated uses of the assigned back-up procedure. Roberts and Powers (1990) dealt with this by switching to an alternative back-up procedure if the initial back-up was repeated 6 times for the same time-out. The spank back-up was changed to a brief room isolation, whereas the brief room isolation and the physical hold were changed to the spank back-up. Switching to an alternative back-up tactic was sufficient to gain compliance with time-out in all cases. This suggests that parents should switch disciplinary tactics when the first one is not working, rather than increasing the intensity of the first tactic. It also implies that parents need more disciplinary options, not fewer ones, to maximize their flexible use of nonabusive alternatives.

### *Cultural Context*

Finally, the child outcomes of physical punishment differed by the cultural context. Three of the

controlled longitudinal studies investigated the effects of physical punishment with ethnic minorities, usually African Americans. All three studies found significantly differential effects of spanking by ethnicity. Physical punishment never predicted predominantly detrimental outcomes in ethnic minorities, except for ethnically diverse 6- to 9-year olds spanked three times a week or more (Straus *et al.*, 1997). In contrast, lower spanking frequency predicted significantly lower rates of fighting 5 years later in African-American children in one study (Gunnore & Mariner, 1997). The other study found neutral outcomes of spanking frequency in African-American children but detrimental outcomes in European-American children (McLeod *et al.*, 1994).

Two uncontrolled longitudinal studies also found significantly distinct outcomes of spanking for European Americans and for African Americans (Deater-Deckard & Dodge, 1997; Deater-Deckard, Dodge, Bates, & Pettit, 1996). The outcomes were predominantly detrimental for European Americans but neutral for African Americans. The strongest evidence of ethnic differences seems to occur in predicting school aggression (see also Polaha, 1998).

One study found differences in outcomes of spanking by religious groups (Ellison *et al.*, 1998). The results varied by whether spanking persisted from about age 3 to about age 10 and by whether the family was conservative Protestant. Beneficial outcomes occurred for early, discontinued spanking relative to those not spanked at either age, but only for conservative Protestants. Detrimental outcomes occurred for persistent spankers, but only for the part of the sample that was not conservative Protestants.

These ethnic and religious subcultural differences in the outcomes of spanking probably depend on how spanking is used and its normative acceptance in those subcultures. For example, spanking is more likely to be perceived as evidence of parental concern in African-American families, whereas it is more likely to be seen as an indication of parental rejection in European-American families (Deater-Deckard *et al.*, 1996). The tradition within the African-American subculture views spanking as a means of establishing appropriate disciplinary control to prevent parents from the need to yell at their children (Mosby, Rawls, Meehan, Mays, & Pettinari, 1999). Understanding these ethnic differences in the outcomes of spanking has been the subject of numerous recent articles (Deater-Deckard & Dodge, 1997, and 11 commentaries; Baumrind, 1996; Mosby *et al.*, 1999; Whaley, 2000).

*Confounding Factors*

To this point, this review examines the results of the 17 causally relevant studies of physical punishment one dimension at a time. However, those dimensions were often confounded with each other. In the previous review (Larzelere, 1996), beneficial child outcomes were found consistently in the causally relevant studies (i.e., four randomized clinical studies), whereas detrimental child outcomes were more likely in causally ambiguous studies, such as uncontrolled longitudinal and retrospective studies. However, several dimensions were confounded with causal conclusiveness as potential explanations for discriminating beneficial from detrimental child outcomes of physical punishment. Studies showing beneficial child outcomes of physical punishment were also more likely to (i) focus on short-term outcomes (e.g., during the next discipline incident or the next week), (ii) emphasize compliance as the child outcome, (iii) study children from the ages of 2 to 6, (iv) investigate the use of spanking to back up milder disciplinary tactics, and (v) make some attempt to rule out abusive use of physical punishment. In contrast, detrimental child outcomes were most likely in studies that were opposite on each of those dimensions as well as being causally ambiguous. The controlled longitudinal studies are directly relevant to untangling some of these confounds. Table II lists them and three other causally relevant studies that varied the characteristics that have been confounded with the causal conclusiveness of the studies.

Predominantly detrimental long-term outcomes never occurred in the causally relevant studies in Table II under the following conditions: if overly severe users of physical punishment were removed from the analyses, if spanking was measured as a conditional back up rather than as a frequency measure, or if the sample of children was initially high on externalizing problems. Only one study found detrimental outcomes for children under the mean age of 6 (Baumrind & Owens, 2000), and it no longer had predominantly detrimental outcomes after those parents who overused physical punishment were removed from the analyses. In contrast, 6 of 9 findings for children over the mean age of 6 indicated predominantly detrimental outcomes, and only the single-case study found beneficial outcomes for children in this age range (Bernal *et al.*, 1968).

Almost all the outcomes in Table II were neutral or beneficial for subcultural groups that are more likely to endorse the use of spanking (3 beneficial

findings, 1 detrimental finding in Straus *et al.*, 1997). The outcomes in this review did not appear to be affected by positive parenting characteristics, although this warrants further investigation.

Detrimental outcomes were more likely when outcomes occurred 6 months or more after the physical punishment (8 detrimental vs. 3 beneficial findings). Beneficial outcomes occurred more often when outcomes occurred less than 6 months later (e.g., during the next discipline incident or the next day; all three findings were beneficial). Thus detrimental outcomes were more likely for long-term outcomes than for short-term outcomes, but this time dimension was often confounded with the other dimensions in Table II.

The outcomes of physical punishment varied little by type of outcome in Table II. Beneficial and detrimental outcomes were equally likely for the category of subsequent behavior problems. There was only one more detrimental outcome than beneficial outcomes for mental health or competence outcomes.

**Comparisons with Alternative Disciplinary Tactics**

A comparison of the outcomes of physical punishment with outcomes of alternative disciplinary tactics is important for sorting out the causality issues. If the apparently detrimental outcomes of nonsevere physical punishment are artifacts actually caused by the initial levels of child misbehavior, then the apparent outcomes of alternative disciplinary tactics should be similar to those of spanking when analyzed in the same way. However, if the detrimental outcomes are unique to physical punishment, this would strengthen the case for uniquely detrimental effects of physical punishment. Twenty (20) studies in Table I investigated alternative disciplinary responses as well as physical punishment. They included all 6 clinical studies, all 3 sequential-analysis studies, 3 of the 8 controlled longitudinal studies, 6 of the 15 uncontrolled longitudinal studies, and 2 of the 6 retrospective studies.

Only three alternative disciplinary tactics ever had more beneficial outcomes than did physical punishment, and each of those tactics compared unfavorably with physical punishment elsewhere. When used as a back-up with 2- and 3-year olds, nonphysical punishment enhanced the subsequent effectiveness of disciplinary reasoning by itself in 9 of 10 analyses, whereas physical punishment enhanced reasoning's effectiveness in only 4 of 10 analyses (Larzelere *et*

**Table II.** Causally Relevant Studies of Physical Punishment Pertinent to Major Confounding Variables

Authors	Association by outcome type		Competence <sup>e</sup>	Long-term?	Severe use of punishment removed	Back-up use vs. frequency of spanking	Positive parenting	Clinically disruptive sample	Child's age (years)
	Behavior problems <sup>a</sup>	Mental health <sup>a</sup>							
<b>Controlled longitudinal studies</b>									
Adams (1995)	0	—	0	Yes	No	Frequency	Controlled	No	6-12
Baumrind & Owens (2000)	0	0	—	Yes	Then 0 effects <sup>f</sup>	Frequency	0 effect when controlled	No	4-5
Ellison <i>et al.</i> (1998)	+ (Cons Pros <sup>b</sup> ) - (Others <sup>b</sup> )	+ (Cons Pros <sup>b</sup> ) - (Others <sup>b</sup> )		Yes	No	Frequency	Not controlled	No	2-4 and 9-11
Gunnoe & Mariner (1997)	+ (Younger, Blacks, girls) - (Older, Whites)			Yes	No	Frequency	Controlled for praise	No	4-7 and 8-11
Larzelere <i>et al.</i> (1998)	0			Yes	Yes	Back-up	Controlled	No	2-3
Larzelere & Smith (2000)	- <sup>d</sup>			Yes	No	Frequency	Controlled	No	6-9
McLeod <i>et al.</i> (1994)	- (Whites) 0 (Blacks)			Yes	No	Frequency	Not controlled	No	6 or older
Straus <i>et al.</i> (1997)	-			Yes	No	Frequency	Controlled	No <sup>e</sup>	6-9
<b>Other causally relevant studies</b>									
Bernal <i>et al.</i> (1968)	+			No	Yes	Back-up	Enhanced by spank	Yes	8
Larzelere & Merenda (1994)	+			No	Yes	Back-up	Partially controlled <sup>e</sup>	No	2-3
Larzelere <i>et al.</i> (1996)	+			No	Yes	Back-up	Partially controlled <sup>e</sup>	No	2-3

*Note:* The most causally conclusive studies (Roberts' four randomized clinical trials) are not included because the purpose of this table is to evaluate confounds that differentiate those studies (with uniformly beneficial outcomes) from uncontrolled longitudinal and retrospective studies (with predominantly detrimental outcomes).

<sup>a</sup>+ = predominantly beneficial effect of physical punishment; - = predominantly detrimental effect of physical punishment; 0 = neutral effect of physical punishment.

<sup>b</sup>Cons Pros = Conservative Protestants. Beneficial effects only for Conservative Protestants spanked only at ages 2-4. Detrimental effects only for others spanked both at ages 2-4 and at 9-11.

<sup>c</sup>Initial antisocial behavior interacted with spanking in three of five cohorts such that spanking decreased subsequent antisocial in the most antisocial children, but increased subsequent antisocial in the least antisocial children (Straus, personal communication, 1996).

<sup>d</sup>The apparent effect of spanking disappeared with an improved statistical control for initial externalizing problem behavior.

<sup>e</sup>Partly within-subject analyses.

<sup>f</sup>Only 1 of 21 associations were significantly detrimental.

*al.*, 1998b). However, a retrospective study found that spanking predicted a lower likelihood of using three illegal drugs, whereas noncontact punishment predicted a greater likelihood of using one illegal drug (Tennant, Detels, & Clark, 1975). Ignoring and physical power assertion both elicited a significantly increased probability of immediate compliance in Ritchie (1999) that was not achieved by spanking or any other disciplinary tactic in that study. However, ignoring was ineffective before spanking was tried with an 8-year old's clinically defiant referral (Bernal *et al.*, 1968), and physical restraint was a significantly less effect back-up for time-out than was a spank back-up (Roberts & Powers, 1990). In addition to these three tactics, grounding had more beneficial outcomes with teenagers than did physical punishment (Caesar, 1988; Joubert, 1992; Larzelere, 1996).

Six additional disciplinary responses were found to have less beneficial outcomes than did physical punishment in at least one study: a child-determined release from time-out with 2- to 6-year olds (Bean & Roberts, 1981; Roberts & Powers, 1990); reasoning without punishment, punishment without reasoning, and disciplinary responses other than punishment or reasoning with 2- and 3-year olds (Larzelere *et al.*, 1996, 1998b); love withdrawal with 5-year olds (Crowne, Conn, Marlowe, & Edwards, 1969); and verbal punishment with 3- to 5-year olds (Baumrind & Owens, 2000).

Focusing only on the 8 controlled longitudinal studies, detrimental outcomes of spanking were found as strongly for both undesirable alternatives (e.g., verbal punishment: Baumrind & Owens, 2000) and recommended alternatives (e.g., privilege removal, grounding, allowance removal, sending to room: Larzelere & Smith, 2000; reasoning: Larzelere *et al.*, 1998b).

In sum, the relative child outcomes in direct comparisons of spanking and alternatives varied by age. Nonabusive spanking compared favorably with six alternatives in 2- to 6-year olds. Four recommended alternatives show outcomes equivalent to spanking during the ages 6 to 9. Grounding has been replicated as a more effective disciplinary alternative than spanking with teenagers (Larzelere, 1996).

## DISCUSSION

There are two major current perspectives on parental use of nonabusive physical punishment: an unconditional antispanking perspective (Straus *et al.*,

1997) and an evidence-based perspective that attempts to differentiate between effective vs. counterproductive spanking (Larzelere, Baumrind, & Polite, 1998a). The unconditional antispanking viewpoint holds that spanking invariably has detrimental child effects regardless of how it is used, the age of the child, the disciplinary situation, the parent-child context, or the cultural context. The evidence-based perspective holds that there may be some parental spanking that would enhance child outcomes, or at least not detract from them. The evidence-based perspective also questions whether the current scientific evidence is adequate for imposing an antispanking value on all parents. What are the implications of this review for these contrasting perspectives? The next section first considers evidence for the unconditional antispanking perspective.

### Is Spanking Invariably Detrimental?

The view that detrimental child outcomes invariably follow from nonabusive spanking is contradicted by several patterns in these results. First, the stronger the causal conclusiveness of the studies, the more likely they are to detect beneficial outcomes of spanking, not detrimental outcomes. Second, causally relevant studies (e.g., controlled longitudinal studies or better) have never found detrimental child outcomes under any one of the following conditions: if abusive parents were removed from the spanking group, if spanking was measured as a back-up for milder disciplinary tactics, or if the sample were clinically defiant children. Most causally relevant studies of children averaging less than 6 years old have found beneficial child outcomes of spanking. The only exception changed to neutral outcomes after removing the 5% to 7% most frequent/severe spankers and controlling for positive parenting (Baumrind & Owens, 2000).

The strongest evidence of invariably detrimental child outcomes came from controlled longitudinal studies of the effect of spanking 6- to 9-year olds on their antisocial behavior 2 to 5 years later (Gunnoe & Mariner, 1997; Larzelere & Smith, 2000; McLeod *et al.*, 1994; Straus *et al.*, 1997). Together, these studies found that the detrimental effect did not generalize consistently to African-American families or to child-reported fighting at school. Only Straus *et al.* (1997) found consistently detrimental outcomes of spanking frequency at ages 6 to 9, predicting increased subsequent antisocial behavior according to maternal reports. The significance of their effect seems to be

solely due to those spanked at the rate of three times or more a week. Overall, the unique effect of spanking accounted for only 1.3% of the variance in subsequent antisocial behavior in their study (Larzelere *et al.*, 1998a).

Larzelere and Smith (2000) replicated and extended the Straus *et al.* (1997) study, using the same dataset. They showed that four recommended alternative disciplinary tactics predicted the same small increase in antisocial behavior 2 years later, when analyzed in the same way as spanking. Such effects for spanking and all four alternatives became nonsignificant when initial child misbehavior was controlled for more adequately, with a 16-item measure of externalizing problems rather than a trichotomous measure of antisocial behavior. Thus, the most consistent evidence of detrimental outcomes of spanking on subsequent antisocial behavior (Straus *et al.*, 1997) seems to be due to residual confounding, which occurs when an inadequate measure of the confounding variable is used as a covariate (Rothman, 1986).

The equivalence of child outcomes for spanking and alternatives is sometimes used as an argument that spanking can be banned because equally effective alternatives are available (Graziano, Hamblen, & Plante, 1996). By itself, this type of argument would be considered insufficient for banning nonparental interventions. For example, the equivalent effectiveness of a new drug (e.g., Tylenol) would not be sufficient grounds for banning a traditional drug (aspirin). Rather, two drugs whose effectiveness is equivalent on the average would each be preferable in particular cases. Second, each would provide an alternative when the other proved ineffective in a particular case. Third, some applications (e.g., treating high fevers) would involve the combined use of both drugs. The application of all three principles to spanking and alternatives are illustrated in Roberts and Powers (1990). As a backup for time-out, spanking and a brief room isolation proved equally effective on the average. Each one worked with some children better than the alternative. Each alternative (spanking or a brief room isolation) was effective when the other alternative was slow in accomplishing the goal of compliance with the time-out procedure. This suggests that parents can be more effective with multiple options in disciplinary tactics, just as they can be more effective with multiple drugs for treating fevers.

Another concern is that negative side effects are more likely for spanking than for alternative disciplinary tactics. Unintended side effects are not a special focus of this review, but conclusive evidence of

unavoidable negative side effects is difficult to find. Reviews of side effects of punishment in general have found such effects to be limited and readily avoided (Newsom, Favell, & Rincover, 1983; Walters & Grusec, 1977). For example, Newsom *et al.* (1983) concluded

Punishment procedures are avoided and underutilized more often from uninformed fears of hypothetical, all-powerful negative side effects than from knowledgeable appraisals of their generally limited and manageable negative side effects. The result is often the continuation of serious behavior problems for months and years when they might be eliminated, to the client's immense long-term benefit, in a matter of days or weeks. (pp. 285–286)

Walters and Grusec (1977) came to a similar overall conclusion, but noted that increased aggression was a likely side effect of physical punishment. This review found mixed evidence on that point. Three causally relevant studies (Bernal *et al.*, 1968; Gunnoe & Mariner, 1997; Larzelere *et al.*, 1996) found that physical punishment reduced subsequent fighting in some circumstances, whereas one of those studies (Gunnoe & Mariner, 1997) found that it also increased subsequent fighting in other circumstances. Walters and Grusec (1977) did not have access to causally relevant studies on spanking and aggression.

The strongest evidence for a negative side effect of physical punishment occurred for the broader outcome of antisocial behavior, which was considered previously (Ellison *et al.*, 1998; Gunnoe & Mariner, 1997; McLeod *et al.*, 1994; Straus *et al.*, 1997). The detrimental effect on antisocial behavior tended to be small, was contradicted in two subcultural groups, and was replicated by every alternative disciplinary tactic investigated to date (Larzelere & Smith, 2000). This particular side effect (increased antisocial behavior) seems to be contradicted by the replicated effectiveness of behavioral parent training for treating conduct disorder (Brestan & Eyberg, 1998; Kazdin, 1995). Its effectiveness has been documented with either the spank back-up or an alternative back-up for time-out. So the empirical evidence for negative side effects of nonabusive spanking is sparse and inconsistent.

Of course, the unconditional antispanking perspective sometimes considers empirical evidence irrelevant due to the ethical value of protecting children from all hitting. This is certainly commendable, but it is not applied absolutely in other areas. There is no widespread advocacy against piercing a child's skin for medical shots or surgery. So the value of



minimizing bodily pain to children is a relative value that gets balanced with the presumed value of such practices. Thus evidence for the effectiveness of spanking is relevant to an appropriate balance of antipain values with competing values.

Moreover, some ethnic and religious groups have explicit values supporting spanking. There has been little careful scholarship on the justification for imposing one set of values on groups with differing values, given various levels of empirical scientific justification. Instead, the discussions about spanking are often based on a superficial understanding of the empirical data or on a simplistic absolute stance against any pain, which is applied inconsistently to other issues.

When 94% of parents use physical punishment at least occasionally with 3- and 4-year olds (Straus & Stewart, 1999), social scientists must clarify these kinds of issues carefully. The inadequacy of the social scientific evidence suggests the possibility that unconditional antispanking advocates are inadvertently imposing one set of values on a very complex issue. Reasoning and nonphysical disciplinary tactics may work better for highly verbal parents in the wealthy suburbs than for many less-advantaged parents. We must be sensitive to cultural, religious, and socioeconomic distinctions before imposing our values on other parents on this important issue.

In sum, Diana Baumrind's (1996) assessment still seems applicable: "A blanket injunction against disciplinary spanking by parents is not scientifically supportable" (p. 828). That is not equivalent, of course, to endorsing or even tolerating all forms of customary physical punishment (Bauman & Friedman, 1998).

### Effective vs. Counterproductive Physical Punishment

This review clarifies one form of effective spanking and one indication of generally counterproductive use of physical punishment. Additional research is needed to clarify the appropriate boundaries between these two extremes of customary physical punishment. A few empirically based hints about this middle gray area are summarized from this review and related studies.

First, spanking has consistently beneficial outcomes when it is nonabusive (e.g., two swats to the buttocks with an open hand) and used primarily to back up milder disciplinary tactics with 2- to 6-year

olds by loving parents. This is consistent with the conditions under which mothers are most likely to use spanking; that is, after milder disciplinary tactics have failed (Goodenough, 1931; Mosby *et al.*, 1999; Ritchie, 1999). The series of studies by Roberts and colleagues show that the spank back-up is one of the two most effective tactics for backing up time-out in clinically defiant children in this age range. The series of studies by Larzelere and colleagues (1998b) showed a similar effect in enforcing disciplinary reasoning among nonclinical 2- and 3-year olds. A benefit of using spanking as a back-up is that the milder disciplinary tactic becomes more effective by itself, thus rendering spanking less necessary subsequently.

At the other extreme, most detrimental outcomes in causally relevant studies are due to overly frequent use of physical punishment. This suggests that overly frequent spanking or its correlates are indicative of counterproductive ways of using spanking. Because the same apparently detrimental effects occur for overly frequent use of four recommended alternative discipline tactics, the dysfunctionality associated with overly frequent punishments generalizes across disciplinary responses rather than being unique to spanking. Further research is needed to determine the nature of this dysfunctionality. An initial study suggested that these apparently detrimental effects are methodologic artifacts (e.g., residual confounding when the initial level of child misbehavior is imperfectly controlled for: Larzelere & Smith, 2000).

Between these two extremes, the research evidence leaves a gray area with a few empirical clues. The clues suggest that the following guidelines are more characteristic of effective spanking than of counterproductive physical punishment:

1. Not overly severe.
2. Under control, not in danger of "losing it" from anger (Straus & Mouradian, 1998).
3. During ages 2 to 6, not during the teenage years (Gunnoe & Mariner, 1997; Simons *et al.*, 1998). Although conclusive evidence is scarce, spanking should be phased out as soon as possible between the ages of 7 and 12.
4. Used with reasoning (Larzelere *et al.*, 1996), preferably eliciting an intermediate rather than a high level of child distress (Larzelere & Merenda, 1994).
5. Used privately (Holmes & Robins, 1988).
6. Motivated by concern for the child, not by parent-oriented concerns (e.g., from frustration, to show who is boss: Larzelere, Klein, Schumm, & Alibrando, 1989).

7. Used after a single warning, generalizing from Roberts (1982). Roberts showed that a single warning before time-out reduced the necessary time outs by 74% without sacrificing any effectiveness of the behavioral parent training.
8. Used flexibly. If spanking does not work, parents should try other approaches and other tactics rather than increasing the intensity of the spanking (Roberts & Powers, 1990).

The first four points are based on multiple studies, whereas the last four points are based primarily on one study each. Thus more and better research is needed, which is the topic of the next section.

### Research on Physical Punishment

A major implication of this review is that better research is needed on physical punishment and on parental discipline in general. Others have noted major gaps in the empirical evidence for currently accepted views of optimal parental discipline (e.g., Grusec & Goodnow, 1994; Yarrow, Campbell, & Burton, 1968).

The following problems seem pervasive in research on physical punishment: (i) failure to take a development perspective; (ii) failure to discriminate between nonabusive vs. overly severe physical punishment; (iii) failure to control adequately for the initial level of child misbehavior; (iv) failure to compare spanking directly with the alternatives that parents could use for similar incidents; (v) failure to distinguish between disciplinary tactics at the initial, middle, and later stages of extended disciplinary incidents; (vi) failure to consider whether the effects of spanking depend on a positive parenting context, such as reasoning or nurturance; and (vii) failure to appreciate subcultural differences in the outcomes and risks of physical punishment.

Along with correcting those problems, two other research issues seem particularly pressing. The first is to understand escalation processes within disciplinary incidents and the role of physical punishment in them, when used in different ways. Most cases of physical abuse occur during discipline incidents (Kadushin & Martin, 1981), but there is little solid evidence about what increases or decreases the risk of escalation toward abuse.

Banning all physical punishment seems to be a plausible means to reduce child abuse, just as the

Prohibition Amendment seemed a plausible way to reduce alcohol abuse (see also Baumrind, 1983). However, the Roberts series of studies and the Larzelere series of studies have shown that firmer tactics are sometimes necessary to enhance the subsequent effectiveness of milder disciplinary tactics. The increased effectiveness of milder disciplinary tactics should then reduce the risk of escalation during subsequent discipline incidents. This suggests the possibility that *how* spanking is used is the critical factor in whether it increases the risk of escalation toward abuse. Of course, how other tactics are used within disciplinary incidents would also affect changes in the risk of escalation (see Snyder, Edwards, McGraw, Kilgore, & Holton, 1994). Baumrind (1973) noted that it was permissive parents who admitted more often to “explosive attacks of rage in which they inflicted more pain or injury upon the child than they had intended. . . . Permissive parents apparently became violent because they felt that they could neither control the child’s behavior nor tolerate its effect upon themselves” (p. 35).

The second pressing issue is the need for a rigorous evaluation of legislation banning parental spanking. To my knowledge, there has been very little empirical evaluation of the first spanking ban in Sweden (Larzelere & Johnson, 1999) and no evaluation of such bans in other countries. One of the stated purposes of a major survey on corporal punishment in Sweden was to see whether the spanking ban had even decreased the prevalence of corporal punishment 15 years after the ban (Statistics Sweden, 1996). That survey found that more than 30% of children raised after the spanking ban had experienced corporal punishment from their parents. Furthermore, this percentage had dropped very little compared to the generation before the spanking ban (e.g., 32% by fathers, compared to 34% in the next oldest generation). Most of the decrease in physical punishment in Sweden occurred before the spanking ban, not after it. The biggest changes after the spanking ban was a small decrease in support for mild spanking, whereas the most problematic types of spanking (e.g., spanking of teenagers) had not decreased at all.

It is possible that the prohibition of all spanking eliminates a type of mild spanking that prevents further escalation within discipline incidents. Milder disciplinary tactics do not get backed up as effectively and thus are less likely to be used. Palmerus and Scarr (1995) compared American and Swedish parents on the disciplinary tactics they used. As expected, Swedish parents used less corporal punishment, but they

also used less reasoning and fewer behavior modification tactics, more yelling, and more restraining.<sup>5</sup> These kinds of changes in the entire disciplinary system might partially account for the 489% increase from 1981 to 1994 in Swedish criminal rates of physical child abuse of children under 7 (Wittrock, 1992, 1995). Many other factors could also have accounted for those increases. The main point is that we need rigorous, unbiased evaluations of such spanking bans, especially given the lack of scientific support for the unconditional antispanking perspective.

### Research on Parental Discipline in General

In addition to better research on spanking, we need better research on parental discipline in general. To understand the role of physical punishment or any disciplinary tactic, we must understand it in the context of parental discipline as a whole.

Much research on physical punishment as well as other disciplinary tactics views a given tactic as invariably effective or invariably counterproductive. There may be more potential in identifying more vs. less effective ways to use each disciplinary tactic. The effect of any tactic may depend on the overall disciplinary style and the parent-child, family, and cultural contexts.

A major article by Grusec and Goodnow (1994) had implications that overlap with this review, even though it emphasized the opposite end of the continuum of traditional disciplinary tactics. Grusec and Goodnow concluded that the greater effectiveness of disciplinary reasoning than various forms of power assertion was supported inconsistently in the empirical literature. Empirical support for the relative effectiveness of reasoning came primarily from samples with low rates of disruptive behavior problems (e.g., girls in middle-class families). In contrast, reasoning often did not look relatively more effective than power-assertive methods in samples with high rates of disruptive behavior problems (e.g., families with lower socioeconomic status; children with difficult temperaments).

This review focuses on disciplinary spanking, at the opposite end of the continuum of traditional disciplinary tactics. Similar to Grusec and Goodnow's

(1994) conclusions about reasoning, it finds the empirical support for commonly held conclusions about nonabusive spanking to be surprisingly weak. An implication shared with Grusec and Goodnow is that we must move beyond inconsistently supported over-generalizations to conceptualize and investigate disciplinary responses more innovatively.

One possibility is a conditional sequence model of optimal disciplinary responses (Larzelere, *in press*). This model says that parents should begin with the mildest disciplinary tactic that they think will be effective in producing appropriate cooperation from a child (e.g., reasoning). If the disciplinary issue is important and neither an appropriate cooperation or negotiation occurs, then the parent should back up the gentle disciplinary tactic (e.g., reasoning) with a nonphysical punishment (e.g., time-out). Only if the child fails to comply with the nonphysical punishment should the parent resort to nonabusive spanking (e.g., two open-handed swats to the buttocks). Each back-up step should be preceded by a single warning (Roberts, 1982). If something like this conditional sequence is used at ages 2 to 6, then the parent should gradually phase out the ultimate back-up tactic (spanking) and then the intermediate back-up tactic (nonphysical punishment). In this way, gentle disciplinary tactics such as reasoning will be used for most disciplinary incidents and they will be effective in keeping the child within acceptable behavioral limits.

This conditional sequence model is consistent with several developmental perspectives whose implications for integrating gentle and firmer disciplinary tactics have not been fully exploited (Baumrind, 1973; Bell & Harper, 1977; Hoffman, 1977; Patterson, 1982; Valsiner, 1987). In general, these developmental mini-theories recognize that disciplinary responses act as part of a control system process, somewhat similar to a thermostat-controlled heating system. Traditional methods of statistical analyses are ill suited for investigating such control system processes, whether cross-sectionally or longitudinally. To illustrate, the average morning temperature in my home during the past year was 61°F when the furnace was running and 67°F when it was not running. This cross-sectional analysis suggests superficially that running the furnace causes my house to be colder. I could design a longitudinal version of this study to prove that running the furnace is either positively or negatively correlated with subsequent house temperatures, depending on the time between data collection points (6 or 12 months). This illus-

<sup>5</sup>Palmerus (personal communication, April 2000) thinks that the yelling and restraining done by Swedish parents was less aversive than that done by American parents. However, she has not divided her variables into milder and more aversive categories to test this impression.

trates that our traditional analytic methods can easily come to incorrect conclusions about control system processes, depending on how we incorporate the initial presenting problem into the analyses and whether our analyses fit the actual time lag of the causal mechanisms.

Parents use disciplinary tactics as part of a control system process. Their disciplinary responses vary depending on their assessment of such things as the importance of the disciplinary issue, whether the child has been cooperating recently, the appropriateness of the child's negotiation attempts, and the child's defiance. Accordingly, Ritchie (1999) showed that maternal tactics differ in the beginning, middle, and end of an extended disciplinary incident.

The control system perspective has many research implications. More information is needed about optimal ways for parents to maintain children's behavior in the appropriate range, ways to prevent children from testing the limits, ways to make gentler disciplinary tactics effective when the limits are tested, and so on. Especially for children with more difficult temperaments or disruptive behavior problems, parents must back up those gentle disciplinary tactics with power-assertive tactics occasionally, especially during the ages from 2 to 6.

The most consistent finding of beneficial child outcomes of nonabusive spanking is consistent with such a control system process. Spanking was effective in reducing subsequent noncompliance and fighting in 2- to 6-year olds when it was used primarily to back up milder disciplinary tactics such as reasoning or time-out. Roberts and Powers (1990) showed that, on the average, a brief room isolation was equally effective as spanking as a back-up of last resort.

The control system perspective also has important implications for last-resort disciplinary tactics. At some point, eliminating all forceful backup tactics renders the entire disciplinary control system ineffective. This seems to be illustrated by mothers who used reasoning frequently with 2- and 3-year olds without ever backing the reasoning up with negative consequences (Larzelere *et al.*, 1998b). Their children's disruptive behavior increased the most during the next 20 months. Such a failure of last-resort backups could easily lead to the development of "nattering," which Patterson (1982) identified as a typical parenting pattern in families of antisocial boys. *Nattering* is defined as nagging or scolding irritably with no intention of following through (Patterson, 1982, p. 112; Patterson, Reid, & Dishion, 1992, p. 66).

A control system perspective also implies that

more forceful tactics should not be overused as initial or primary disciplinary responses. This is consistent with the fact that overly frequent physical punishment has the most consistent evidence of detrimental causal effects on children's behavior.

An optimal disciplinary control system fits an authoritative parenting style, which combines parental nurturance and give-and-take communication with firm control when necessary. There are a range of appropriate ways to combine parental nurturance, communication, and control with approximately similar outcomes for the child (Baumrind, 1991). It is the extreme parenting styles that yield detrimental child outcomes, whether the extreme is permissive or authoritarian. Similarly, there are many ways for parents to implement a reasonably effective disciplinary control system within a loving parent-child relationship. The extremes are most likely to be counterproductive. One extreme would be never or rarely using firm tactics to set effective limits. The other extreme would be rarely using milder disciplinary tactics as a preferred disciplinary response before resorting to harsher disciplinary tactics.

A control system process illustrates only one possible innovative way to investigate optimal disciplinary responses. It seems to be counterproductive and simplistic to continue viewing some disciplinary tactics as invariably good and others as invariably bad, as long as those tactics are nonabusive. Another innovative approach would consider how disciplinary response tactics influence the parent-child relationship quality and how that quality in turn influences the effects of disciplinary tactics (Kuczynski & Lollis, *in press*). The main point is that parental discipline research needs to break out of some boxes, both methodologically and conceptually.

In conclusion, this review indicates that we have a lot to learn about the complexities of appropriate parental discipline. Research on spanking does not support an unconditional antispanking position at this time. It is more difficult to establish conclusive boundaries between effective and counterproductive physical punishment. Spanking as a back-up for milder disciplinary tactics in 2- to 6-year olds seems to produce beneficial child outcomes, at least in reducing noncompliance and fighting. Overly frequent spanking predicts a wider range of detrimental outcomes, but to a degree matched by overly frequent use of every alternative disciplinary tactic investigated to date. Future research must distinguish between effective vs. counterproductive physical punishment. It must also distinguish between effective

vs. counterproductive alternatives and how those alternatives compare to nonabusive spanking in similar disciplinary situations. Research must also explore how proactive discipline, disciplinary reasoning, firmer disciplinary tactics, and parental nurturance mutually influence each other. Several theoretic perspectives in developmental psychology suggest ways to conceptualize such an integration, but those leads must be more fully exploited.

## ACKNOWLEDGMENTS

I thank Sheila Eyberg and two anonymous reviewers for very helpful critiques of a previous draft of this article.

## REFERENCES

- Adams, M. J. (1995). *Youth in crisis: An examination of adverse risk factors affecting children's cognitive and behavioral/emotional development, children ages 10–16*. Unpublished doctoral dissertation, University of Texas at Dallas, Dallas.
- Alibrando, S. A., Jr. (1987). *The effects of corporal punishment and contextual parental characteristics on rebelliousness, neuroticism and introversion*. Unpublished doctoral dissertation, Biola University, LaMirada, CA.
- Bauman, L. J., & Friedman, S. B. (1998). Corporal punishment. *Pediatric Clinics of North America*, 45, 403–414.
- Baumrind, D. (1973). The development of instrumental competence through socialization. In A. D. Pick (Ed.), *Minnesota Symposia on Child Psychology* (Vol. 7, pp. 3–46). Minneapolis: University of Minnesota Press.
- Baumrind, D. (1983). Specious causal attributions in the social sciences: The reformulated stepping-stone theory of heroin use as exemplar. *Journal of Personality and Social Psychology*, 45, 1289–1298.
- Baumrind, D. (1991). The influence of parenting style on adolescent competence and substance use. *Journal of Early Adolescence*, 11, 56–95.
- Baumrind, D. (1996). A blanket injunction against disciplinary use of spanking is not warranted by the data. *Pediatrics*, 98, 828–831.
- Baumrind, D., & Owens, E. B. (2000). *Does normative physical punishment by parents cause detrimental child outcomes: A prospective longitudinal study*. Manuscript in preparation.
- Bean, A. W., & Roberts, M. W. (1981). The effect of time-out release contingencies on changes in child noncompliance. *Journal of Abnormal Child Psychology*, 9, 95–105.
- Bell, R. Q., & Harper, L. V. (1977). *Child effects on adults*. Hillsdale, NJ: Erlbaum.
- Bernal, M. E., Duryee, J. S., Pruett, H. L., & Burns, B. J. (1968). Behavior modification and the brat syndrome. *Journal of Consulting and Clinical Psychology*, 32, 447–455.
- Brestan, E. V., & Eyberg, S. M. (1998). Effective psychosocial treatments of conduct-disordered children and adolescents: 29 years, 82 studies, and 5,272 kids. *Journal of Clinical Child Psychology*, 27(2), 180–189.
- Bryan, J. W., & Freed, F. W. (1982). Corporal punishment: Normative data and sociological and psychological correlates in a community college population. *Journal of Youth & Adolescence*, 11, 77–87.
- Caesar, P. (1988). Exposure to violence in the families-of-origin among wife-abusers and maritally nonviolent men. *Violence and Victims*, 3, 49–63.
- Crowne, D. P., Conn, L. K., Marlowe, D., & Edwards, C. N. (1969). Some developmental antecedents of level of aspiration. *Journal of Personality*, 37, 73–92.
- Day, D. E., & Roberts, M. W. (1983). An analysis of the physical punishment component of a parent training program. *Journal of Abnormal Child Psychology*, 11, 141–152.
- Deater-Deckard, K., & Dodge, K. A. (1997). Externalizing behavior problems and discipline revisited: Nonlinear effects and variation by culture, context, and gender. *Psychological Inquiry*, 8, 161–175.
- Deater-Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1996). Physical discipline among African American and European American mothers: Links to children's externalizing behaviors. *Developmental Psychology*, 32, 1065–1072.
- Ellison, C., Musick, M., & Holden, G. (1998). *A longitudinal study of the effects of corporal punishment: The moderating effect of conservative Protestantism*. Unpublished manuscript.
- Gershoff, E. T. (1999). *The effects of parental corporal punishment on children: A process model and meta-analytic review*. Unpublished manuscript.
- Goodenough, F. L. (1931). *Anger in young children*. Minneapolis: University of Minnesota Press.
- Graziano, A. M., Hamblen, L., & Plante, W. A. (1996). Subabusive violence in child rearing in middle-class American families. *Pediatrics*, 98(4), 845–848.
- Grinder, R. E. (1962). Parental childrearing practices, conscience, and resistance to temptation of sixth grade children. *Child Development*, 33, 803–820.
- Grusec, J. E., & Goodnow, J. J. (1994). Impact of parental discipline methods on the child's internalization of values: A reconceptualization of current points of view. *Developmental Psychology*, 30, 4–19.
- Gunnoe, M. L., & Mariner, C. L. (1997). Toward a developmental-contextual model of the effects of parental spanking on children's aggression. *Archives of Pediatrics and Adolescent Medicine*, 151, 768–775.
- Hoffman, M. L. (1977). Moral internalization: Current theory and research. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 85–133). New York: Academic Press.
- Holmes, S. J., & Robins, L. N. (1988). The role of parental disciplinary practices in the development of depression and alcoholism. *Psychiatry*, 51, 24–36.
- Hyman, I. A. (1997). *The case against spanking*. San Francisco: Jossey-Bass.
- Johannesson, I. (1974). Aggressive behavior among school children related to maternal practices in early childhood. In J. deWit & W. W. Hartup (Eds.), *Determinants and origins of aggressive behavior* (pp. 413–426). The Hague: Mouton.
- Joubert, C. E. (1992). Antecedents of narcissism and psychological reactance as indicated by college students' retrospective reports of their parents' behaviors. *Psychological Reports*, 70, 1111–1115.
- Kadushin, A., & Martin, J. A. (1981). *Child abuse: An interactional event*. New York: Columbia University Press.
- Kazdin, A. E. (1995). *Conduct disorders in childhood and adolescence* (2nd ed.). Thousand Oaks, CA: Sage.
- Kuczynski, L., & Lollis, S. (in press). Four foundations for a dynamic model of parenting. In J. R. M. Gerris (Ed.), *Dynamics of parenting*. Hillsdale, NJ: Erlbaum.
- Larzelere, R. E. (1996). A review of the outcomes of parental use of nonabusive or customary physical punishment. *Pediatrics*, 98, 824–828.
- Larzelere, R. E. (in press). Combining love and limits in authoritative parenting: A conditional sequence model of disciplinary

- responses. In J. C. Westman (Ed.), *Parenthood in America*. Madison, WI: University of Wisconsin Press.
- Larzelere, R. E., Baumrind, D., & Polite, K. (1998a). Two emerging perspectives of parental spanking from two 1996 conferences. *Archives of Pediatrics and Adolescent Medicine*, *152*, 303–305.
- Larzelere, R. E., & Johnson, B. (1999). Evaluation of the effects of Sweden's spanking ban on physical child abuse rates: A literature review. *Psychological Reports*, *85*, 381–392.
- Larzelere, R. E., Klein, M., Schumm, W. R., & Alibrando, S. A., Jr. (1989). Relations of spanking and other parenting characteristics to self-esteem and perceived fairness of parental discipline. *Psychological Reports*, *64*, 1140–1142.
- Larzelere, R. E., & Merenda, J. A. (1994). The effectiveness of parental discipline for toddler misbehavior at different levels of child distress. *Family Relations*, *43*, 480–488.
- Larzelere, R. E., Sather, P. R., Schneider, W. N., Larson, D. B., & Pike, P. L. (1998b). Punishment enhances reasoning's effectiveness as a disciplinary response to toddlers. *Journal of Marriage and the Family*, *60*, 388–403.
- Larzelere, R. E., Schneider, W. N., Larson, D. B., & Pike, P. L. (1996). The effects of discipline responses in delaying toddler misbehavior recurrences. *Child & Family Behavior Therapy*, *18*, 35–57.
- Larzelere, R. E., & Smith, G. L. (2000, August). *Controlled longitudinal effects of five disciplinary tactics on antisocial behavior*. Paper presented at the annual convention of the American Psychological Association, Washington, DC.
- Lonigan, C. J., Elbert, J. C., & Johnson, S. B. (1998). Empirically supported psychosocial interventions for children: An overview. *Journal of Clinical Child Psychology*, *27*, 138–145.
- Lyons, J. S., Anderson, R. L., & Larson, D. B. (1993). *The use and effects of physical punishment in the home: A systematic review*. Paper presented at the meeting of the American Academy of Pediatrics, Washington, DC.
- MacIntyre, D. I., & Cantrell, P. J. (1995). Punishment history and adult attitudes towards violence and aggression in men and women. *Social Behaviour and Personality*, *23*, 23–28.
- McClelland, D. C., & Pilon, D. A. (1983). Sources of adult motives in patterns of parent behavior in early childhood. *Journal of Personality and Social Psychology*, *44*, 564–574.
- McCord, J. (1988). Parental aggressiveness and physical punishment in long term perspective. In G. Hotaling, D. Finkelhor, J. Kilpatrick, & M. Straus (Eds.), *Family abuse and its consequences*. Newbury Park, CA: Sage.
- McCord, J., & Ensminger, M. E. (1997). Multiple risks and comorbidity in an African-American population. *Criminal Behaviour and Mental Health*, *7*, 339–352.
- McLeod, J. D., Kruttschnitt, C., & Dornfeld, M. (1994). Does parenting explain the effects of structural conditions on children's antisocial behavior? A comparison of blacks and whites. *Social Forces*, *73*, 575–604.
- Michels, S., Pianta, R., & Reeve, R. (1993). Parent self-reports of discipline practices and child acting-out behaviors in kindergarten. *Early Education and Development*, *4*, 139–144.
- Miller, N., & Pollock, V. E. (1994). Meta-analytic synthesis for theory development. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 457–484). New York: Russell Sage Foundation.
- Mosby, L., Rawls, A. W., Meehan, A. J., Mays, E., & Pettinari, C. J. (1999). Troubles in interracial talk about discipline: An examination of African American child rearing narratives. *Journal of Comparative Family Studies*, *30*, 489–521.
- Newsom, C., Favell, J. E., & Rincover, A. (1983). Side effects of punishment. In S. Axelrod & J. Apsche (Eds.), *The effects of punishment on human behavior* (pp. 285–316). New York: Academic Press.
- Palmerus, K., & Scarr, S. (1995). *How parents discipline young children. Cultural comparisons and individual differences*. Paper presented at the conference of the Society for Research on Child Development, Indianapolis, IN.
- Paschall, M. J., & Straus, M. A. (1998). *Corporal punishment by mothers and child's cognitive development: A second longitudinal study*. Unpublished abstract.
- Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia Press.
- Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys*. Eugene, OR: Castalia.
- Polaha, J. A. (1998). *The relationship between physical discipline and child behavior problems: A study of group differences*. Unpublished doctoral dissertation, Auburn University, Auburn, AL.
- Rhue, J. W., & Lynn, S. J. (1987). Fantasy proneness: Developmental antecedents. *Journal of Personality*, *55*, 121–137.
- Ritchie, K. L. (1999). Maternal behaviors and cognitions during discipline episodes: A comparison of power bouts and single acts of noncompliance. *Developmental Psychology*, *35*, 580–589.
- Roberts, M. W. (1982). The effects of warned versus unwarned time-out procedures on child noncompliance. *Child & Family Behavior Therapy*, *4*, 37–53.
- Roberts, M. W. (1988). Enforcing chair timeouts with room timeouts. *Behavioral Modification*, *12*, 353–370.
- Roberts, M. W., & Powers, S. W. (1990). Adjusting chair timeout enforcement procedures for oppositional children. *Behavior Therapy*, *21*, 257–271.
- Rothman, K. S. (1986). *Modern epidemiology*. Boston, MA: Little, Brown and Co.
- Schenck, E. R., Lyman, R. D., & Bodin, S. D. (2000). Ethical beliefs, attitudes, and professional practices of psychologists regarding parental use of corporal punishment: A survey. *Children's Services: Social Policy, Research, and Practice*, *3*(1), 23–38.
- Sears, R. R. (1961). Relation of early socialization experiences to aggression in middle childhood. *Journal of Abnormal and Social Psychology*, *63*, 466–492.
- Sears, R. R. (1970). Relation of early socialization experiences to self-concepts and gender role in middle childhood. *Child Development*, *41*, 267–289.
- Simons, R. L., Lin, K.-H., & Gordon, L. C. (1998). Socialization in the family of origin and male dating violence: A prospective study. *Journal of Marriage and the Family*, *60*, 467–478.
- Smith, J. R., & Brooks-Gunn, J. (1997). Correlates and consequences of harsh discipline for young children. *Archives of Pediatrics and Adolescent Medicine*, *151*, 777–786.
- Snyder, J., Edwards, P., McGraw, K., Kilgore, K., & Holton, A. (1994). Escalation and reinforcement in mother-child conflict: Social processes associated with the development of physical aggression. *Development and Psychopathology*, *6*, 305–321.
- Statistics Sweden. (1996). *Spanking and other forms of physical punishment* (Demography, the Family, and Children 1996:1.2). Stockholm: Statistics Sweden.
- Stattin, H., Janson, H., Klackenber-Larsson, I., & Magnusson, D. (1995). Corporal punishment in everyday life: An intergenerational perspective. In J. McCord (Ed.), *Coercion and punishment in long-term perspectives* (pp. 315–347). Cambridge, England: Cambridge University Press.
- Strassberg, Z., Dodge, K. A., Pettit, G. W., & Bates, J. E. (1994). Spanking in the home and children's subsequent aggression toward kindergarten peers. *Development and Psychopathology*, *6*, 445–461.
- Straus, M. A., & Mouradian, V. E. (1998). Impulsive corporal punishment by mothers and antisocial behavior and impulsiveness of children. *Behavioral Sciences and the Law*, *16*, 353–374.
- Straus, M. A., & Paschall, M. J. (1998, August 1). *Corporal punishment by mothers and child's cognitive development: A longitu-*

- dinal study*. Paper presented at the 14th World Congress of Sociology, Montreal, Canada.
- Straus, M. A., & Stewart, J. H. (1999). Corporal punishment by American parents: National data on prevalence, chronicity, severity, and duration, in relation to child and family characteristics. *Clinical Child and Family Psychology Review*, 2(2), 55–70.
- Straus, M. A., Sugarman, D. B., & Giles-Sims, J. (1997). Spanking by parents and subsequent antisocial behavior of children. *Archives of Pediatrics and Adolescent Medicine*, 151, 761–767.
- Tennant, F. S., Jr., Detels, R., & Clark, V. (1975). Some childhood antecedents of drug and alcohol abuse. *American Journal of Epidemiology*, 102, 377–385.
- Valsiner, J. (1987). *Culture and the development of children's actions: A cultural–historical theory of developmental psychology*. New York: Wiley.
- Walters, G. C., & Grusec, J. E. (1977). *Punishment*. San Francisco: W. H. Freeman.
- Weiss, B., Dodge, K. A., & Bates, J. E. (1992). Some consequences of early harsh discipline: Child aggression and a maladaptive social information processing style. *Child Development*, 63, 1321–1335.
- Whaley, A. L. (2000). Sociocultural differences in the developmental consequences of the use of physical discipline during childhood for African Americans. *Cultural Diversity and Ethnic Minority Psychology*, 6(1), 5–12.
- Wittrock, U. (1992). Barnmisshandel i kriminalstatistiken, 1981–1991 [Violent crimes against children in criminal statistics, 1981–1991]. *KR Info*, 1992: 7.
- Wittrock, U. (1995). Barnmisshandel, 1984–1994 [Violent crimes against children, 1984–1994]. *KR Info*, 1995: 5.
- Yarrow, M. R., Campbell, J. D., & Burton, R. V. (1968). *Child rearing: An inquiry into research and methods*. San Francisco: Jossey-Bass.

Copyright of Clinical Child & Family Psychology Review is the property of Kluwer Academic Publishing and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.