

**DOES CAUSALLY RELEVANT RESEARCH SUPPORT A BLANKET INJUNCTION
AGAINST DISCIPLINARY SPANKING BY PARENTS?**

Diana Baumrind, Ph.D.
Institute of Human Development
University of California,
Berkeley, California

Invited Address at the 109th Annual Convention of the American Psychological
Association
August 24, 2001

* I wish to acknowledge the invaluable help of the students who worked on this project – most notably Nadia Sorkhabi who trained the raters and together with Ellen Middaugh developed the Manual for the Parent Disciplinary Rating Scale, and Sally Lim for her important contribution to data coding and management. Dr. Elizabeth Owens conducted the data analyses and co-authors the report of this study, which is in the process of peer review. Dr. Owens and myself are grateful to the William T. Grant Foundation who funded our efforts, and to Dr. Robert Larzelere with whom we consulted on substantive and methodological matters.

** The data included in this address have not completed peer review. Therefore the readers and media should reserve judgment.

Does causally relevant research support a blanket injunction against disciplinary spanking by parents?

The title of this talk -- “Does causally relevant research support a blanket injunction against disciplinary spanking by parents?” -- encompasses the question to be addressed today. The longitudinal case records from the Family Socialization and Developmental Competence Project, which I will hereafter refer to by its initials – FSP – were recoded for the purpose of measuring the effects of normative physical punishment on the participating children and adolescents.

It is generally acknowledged that the methodology used to address the effects of spanking on children’s adjustment is too problematic to support a causal argument. Nevertheless findings from these problematic studies have been used, as though they were causally relevant, to support an unconditional anti-spanking position by organizations such as EPOCH-Worldwide and the Canadian Foundation for Children, Youth and the Law. For that reason it is timely to remind ourselves of the elementary methodological criteria that correlational data must meet to support causal conclusions, especially conclusions intended to affect social policy – in this case to criminalize a normative disciplinary practice, namely spanking. The study I will discuss today was designed expressly to meet these elementary methodological criteria by mining the unusually comprehensive FSP archival case records and data base, to measure and then control third variables that could threaten the validity of causal conclusions concerning spanking effects on child outcomes.

I will use the term “spanking” (See Friedman and Schonberg [1996, p.853], and Strassberg and colleagues do [1994, p.446].) to refer to striking the child on the buttocks or extremities with an open hand without inflicting physical injury with the intention to modify behavior. Spanking is intended to be aversive, but not necessarily by inflicting physical pain.

There is a significant and convincing body of evidence that physical and sexual abuse of children is harmful and traumatic (e.g., Baumrind, 1994, 1995). Unconditional anti-spanking advocates such as Murray Straus argue that any level of physical punishment is harmful so that in effect the consequences of normative and abusive physical punishment differ in degree, not in kind. However, Larzelere (2000) concluded from his review of child outcomes associated with nonabusive physical punishment that “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” (p.209).

What then, according to such experts as Cook and Campbell, Abelson, and Hill, are the elementary methodological criteria that studies must meet in order to support a causal argument that corrective, non-abusive physical punishment by parents is harmful, (See Abelson, 1995; Cook and Campbell, 1979; Rothman and Greenland, 1998; and Hill, 1965.); and how was our study designed to meet those elementary criteria?

First, parents who use physical punishment abusively must be distinguished from those whose use is normative in frequency and intensity. Therefore, when examining the effects of “spanking”, the FSP sample was limited to families in which the severity and frequency of physical punishment was normative for that population.

Second, in order to address the critical issue of causal direction, a control for baseline child misbehavior should be included in the analyses. Without that control we have no way of knowing if the initial level of child misbehavior caused an increase in both spanking and subsequent child problems. The FSP provided highly reliable measures of baseline child misbehavior which were then used as covariates in analyses intended to be causally relevant¹.

Third, in order to identify spanking as the culprit in the event that spanking does correlate significantly with detrimental child outcomes, the influence of artifactual third variables such as parental rejection must be accounted for. Therefore highly reliable measures of positive and negative parenting practices were entered as plausible third variables in the analyses.

Fourth, there must be no shared source variance between measures of parent and child behavior that would inflate or distort the magnitude of their relationship. The information on both parent and child are obtained from the same source in most of the social survey data analyzed by Straus and others, whereas the data collected in this study on parents and children are completely independent within and across time periods, thus preventing the bias introduced by shared source variance.

Fifth, the joint impact of fathers’ as well as mothers’ use of physical punishment should be assessed. In this study, unlike in most, both parents’ use of physical and verbal punishment are assessed.

Sixth, in order to justify an injunction against the use of spanking based on the presumption that it is uniquely harmful, the impact of physical punishment should be contrasted with the impact of other disciplinary interventions. Therefore we assessed the impact of verbal punishment in addition to that of physical punishment.

Seventh, to prevent residual confounding which gives the appearance of controlling artifacts without actually doing so, the measures of possible artifactual or third variables must be of high quality. The time-intensive assessments of each family at each time period in the FSP data base provide high-quality measures that meet the common threats to construct validity and internal validity that beset self-report population-based survey studies of physical punishment. Information about parents and children in the FSP came from direct observation in naturalistic settings (such as the home, classroom, and school playground); from intensive semi-structured interviews with parents, children, and teachers; and from standardized and project-designed psychological tests.

Observations and interviews were conducted by highly trained professionals who were selected for their heterodox views and life experiences. Using transcripts of the entire

battery of interview and observational procedures at a particular assessment point, the primary observer/interviewer and a second rater completed project-designed Q-sorts for children and Likert-type rating scales for parents². Although theoretical constructs were initially used to develop the child and parent ratings items, the final measure scores were determined by empirical reduction of the individual items using cluster analytic techniques. Measures of the full range of possible child and adolescent outcomes including measures of children's baseline misbehavior were available in the FSP archives which also provided highly reliable measures of relevant parenting behavior. These serve as controls of third variables that could account for the correlations between spanking and child outcomes.

THE STUDY

For this particular study concerning effects of parents' disciplinary practices, a new instrument, the Parent Disciplinary Rating Scale (PDRS), was created. 71 rating items reflecting how characteristic it is for each parent to use various discipline strategies and behaviors comprised the scale. At each time point, after reading the transcribed mother and father interviews and the descriptions of home observations of the entire family, two trained coders rated each mother and father.

Sample

The FSP sample is ideally suited to evaluate the correlates of normative physical punishment in a socially-advantaged population. We reasoned that if physical punishment is harmful per se, rather than a proxy for economic stress and inept parenting, its detrimental effects would be discernable in this socially advantaged population. Our findings do not address the effects on child outcomes of legally abusive physical punishment, but rather of frequency and intensity of mild to moderately severe, but not legally abusive, physical punishment administered during each of the three time periods – preschool (T1), early primary school (T2) and early adolescence (T3), in a low-risk social context. The FSP sample is of moderate size ranging from an n of 79 for some longitudinal analyses to an n of 164 for some cross-sectional analyses.

Recourse to some physical punishment was normative in the FSP sample, despite the liberal politics of the Berkeley community, and the high educational level and social status of the parents. Although by Time 3 when the children were 14 and 15, 62% of parents used no physical punishment, only 4% of the parents had never used physical punishment at Time 1, when the children were preschoolers, and only 16% had never used physical punishment between Time 1 and Time 2 when the children were ages 8 and 9. There was a considerable range of frequency and severity of use of physical punishment by the FSP parents, with a small minority, between 4% and 7%, at each time period resorting to non-normative, although not legally abusive, physical punishment. The analyses that were intended to refer to child outcomes associated with normative physical punishment excluded those parents.

Measures

Frequency of physical punishment at each time period was measured using two items from the Parent Disciplinary Rating Scale or PDRS which together enabled parent dyads to be placed in ordinally-arranged groups reflecting the frequency with which they used physical punishment. At T1 and T2 seven such groups were created. At T3, only four groups were needed because most couples used no physical punishment. The groups are operationally defined in Table 1 of your handout.

For theoretical reasons pertaining to Straus' claim that any physical punishment was harmful we chose to differentiate between parents who never spanked during the time period and those who very seldom did, that is who had spanked the child one to three times, or three to five times. Straus did not use planned contrasts to test his claim that any spanking was harmful.

In order to detect overly severe use of physical punishment by either parent a separate score for intensity of physical punishment at each age was created from two items: a) Parent uses paddle or other instrument to strike the child, or strikes the child on the face or torso, and b) Parent lifts the child and throws or shakes the child. Average rates of parental physical punishment intensity in this sample were very low (under 1.10 on a 1- to 3-point scale), with the large majority of families displaying no intensity of physical punishment at any age, even at the highest level of frequency. However, a small number of parent pairs did report using intense physical punishment sometimes, and a few pairs received intensity scores that were notably higher than the rest.

A second measure of physical punishment was created reflecting both the frequency and intensity of physical punishment, in order to categorize parents into theoretically contrasting parent categories. For these category assignments we used the metaphor of Green, Yellow and Red to signify, respectively, "OK to go", "caution suggested", and "stop". We further divided "caution suggested" into Yellow and Orange. Although the operational definitions of these categories vary slightly from T1 through T3, general descriptions are as follows: **Green** = little or no physical punishment and no intensity; **Yellow** = occasional physical punishment with no or little intensity; **Orange** = some to often use of physical punishment with little or no intensity, or frequent physical punishment with no intensity; **Red** = frequent physical punishment with at least some intensity.

The operational definitions of these four parent categories appear as Table 2 of your handout.

A measure of frequency of an alternative disciplinary tactic, verbal punishment, was created using three items: a) Parent yells or shouts, b) Parent belittles the child by the use of sarcasm, and c) Parent engages in nattering (pointless and disapproving chatter).

Additionally, two control variables to be used as covariates were created. A Positive Parenting score for each family at each age was derived from 23 items of the PDRS that reflected planful and supportive discipline, use of reason to persuade, flexible tactics, coherent and consistent discipline policies, adjustment of discipline to developmental needs, together with archival scores on clusters measuring parents' responsiveness and psychological differentiation. A Maladjusted Parenting score was also created at each age from 3 archival scores on the Internalizing/Maladjusted, Externalizing/Exploitative, and Intrusive clusters.

At each time period six primary parenting types had been derived from the archival data based on patterns of scores on the parent clusters that loaded most highly on the two primary factors of Demandingness and Responsiveness³. The six primary parenting types are labelled: Authoritative, Democratic, Directive, Good Enough, Permissive, and Rejecting/Neglecting. The patterns of scores for Demandingness and Responsiveness that define these types are depicted schematically in Table 3 of your handout.

In order to statistically control for earlier child misbehavior when assessing the effects of physical punishment on children, two variables from the archival data base were used – Externalizing Problem Behavior and Uncooperative. Externalizing Problem Behavior includes as high loading items Gets into fights, Bullies and Socially Disruptive. The Uncooperative cluster includes as high loading items Disobedient, Untrustworthy and Defiant.

Table 4 of your handout provides a summary of the data reduction procedures for all the parent and child measures.

Data Analyses

Our general approach to data analysis was first to identify the unconditional relations between parents' physical punishment and child outcomes within and across time periods and then by controlling for potentially confounding variables to successively probe these relations to see if they "held" once alternative explanations were considered. In order to evaluate the net effects of normative physical punishment on child outcomes, after removing the Red zone families, regression-based adjustments for baseline child misbehavior and causally relevant parenting variables were subsequently included.

In view of the controversial nature of our subject, our approach to data analyses favored transparency over elegance, as well as efforts to maximize the probability of disconfirming the null hypothesis by using one-tailed t – tests where specific directional hypotheses concerning the detrimental outcomes of normative physical punishment were probed.

We began our analyses with computation of correlations between the dimensional (frequency) measure of physical punishment at each time point and the set of measures of child and adolescent outcomes. Determining whether such relations are spurious or real was then addressed by examining the relations between physical punishment and child outcomes while controlling for potentially relevant confounding variables. First, in order to examine the effects of normative socially accepted spanking, the few Red zone parents whose use of physical punishment was unusually severe for this population were removed and the correlations of what we then called "normative physical punishment", or NPP, with child outcomes were re-run. Then, in order to control for the effect of relevant third variables that could plausibly account for the apparent effects of NPP, the relations between NPP and child outcomes were examined using semi-partial correlations in which parental education, previous child misbehavior, positive parenting, and maladjusted parenting were successively statistically controlled. The results of these analyses are displayed in Table 5 of your handout. At T3, we also considered whether adolescents' perceptions of their parents as loving and competent accounted for the significant relations between NPP and child outcomes.

We then investigated the existence of relations between the child/adolescent outcomes and the categorical measure of physical punishment, which contains the Green, Yellow, Orange, and Red categories, in order to determine whether child outcomes differed

among categories of parents differentiated by theoretically meaningful levels of frequency and intensity of physical punishment, regardless of whether effects of the dimensional measure of physical punishment had been found. Subsequently, as displayed in Table 6, these relations were tested using a series of ANCOVAs in which the same potentially confounding variables were covaried.

Additionally, within the Green group at Time 1, planned contrasts were conducted between child outcomes of the (five) parent couples who used no physical punishment at all, and those (six) who used a little physical punishment (for example one to three times or three to five times) to see if there was any evidence that the children who were never spanked were better adjusted.

We then asked whether physical punishment moderated the associations between child outcomes and four of the six major parenting types described earlier. We knew from previous analyses of the archival data that the most beneficial child outcomes were linked to Authoritative and Democratic parenting, whereas the most detrimental child outcomes were linked to Rejecting-Neglecting and Authoritarian-Directive parenting. Specifically, did membership in the Orange group where parents used above-average, but still normative, frequency of punishment, decrease the beneficial outcomes associated with Authoritative and Democratic parenting, or increase the harmful outcomes associated with Rejecting/Neglecting and Directive parenting?

Finally, outcomes associated with Verbal Punishment were assessed in order to determine if any negative effects of physical punishment were unique, or instead were common to a nonphysical punishment.

Results and Discussion of Results

I will now summarize and discuss the results of the study that address specific questions. Tables 5 and 6 in your handout present the correlational and categorical parent-child relations respectively as they pertain to these questions.

(Question 1): Our first question pertains to the direction of the uncontrolled correlations between physical punishment and children's competencies.

Prior to excluding families in the Red zone where parents can be said to hit violently, we replicated, as expected, the direction of the parent-child associations found by Straus, and others. Thus, as can be seen in the first column of Table 5, the significant uncontrolled correlations between physical punishment and child outcomes, while small, suggest that frequency/intensity of physical punishment is associated with detrimental child outcomes.

(Question 2): Are the Red zone parents and their children qualitatively different from other study participants?

Indeed they are. Compared to other parents in the study, the 4% to 7% of parents in the Red Zone were much more exploitive and intrusive and much less responsive, planful and consistent in their discipline. Their children were consistently much less competent and more maladjusted than children of parents in the Green or Yellow zones, and the reverse was never true.

(Question 3): We then asked, with the few families in the Red Zone removed, do certain co-occurring conditions that Abelson (1995) calls third variables and impurities account for the relations between normative physical punishment, and children's later detrimental outcomes?

The answer is yes. Controlling for the intervention selection bias introduced by the child's early misbehavior largely accounted for the relations between NPP and later detrimental outcomes. After accounting for earlier child misbehavior, with only one exception, the remaining correlations over .15 between spanking frequency and child/adolescent outcomes approach zero after the measures of positive and maladjusted parenting are covaried.

Time considerations prevent review of the categorical results presented in Table 6, but with a few notable exceptions, due to contrasts between children of Red or Orange zone parents and other children, the same co-occurring conditions account for the significant associations between children's attributes and their parents' physical punishment classification. There are no significant differences between children of parents who spank seldom and those who spank moderately as represented respectively by the Green and Yellow zones.

Somewhat surprisingly, adolescents' perceptions of their parents as "exemplary", assessed by a Q-sort completed by the adolescent for each parent, all by itself fully accounts for the concurrent relation between spanking frequency and adolescent maladjustment.

(Question 4): We then asked -- were preschool children who never experienced physical punishment better adjusted than preschool children who occasionally did?

The answer is clearly no when tested by planned contrasts⁴. In fact, at T1 the reverse tended to be true. At T1, the 5 children in the Green zone who never experienced physical punishment tended to be somewhat less well-adjusted than those other (six) children in the Green zone who experienced occasional but infrequent physical punishment, although contrasts were typically not statistically significant.

(Question 5): We then asked -- is the use of "more frequent than average" physical punishment, represented by the Orange zone, detrimental to children, as we hypothesized it would be?

Orange zone parents spanked their children more frequently than average. There was only moderate support for the hypothesized detrimental effect of spanking in the Orange zone at T1 and T3, and almost no support at T2. However, although not maladjusted, these children were somewhat less competent and well-adjusted than children of parents in the Green zone.

(Question 6): We then asked does physical punishment act as a moderator of the relations between child outcomes and parenting type?

As you can see from Table 7, at all time periods, children from Authoritative, and to a lesser extent Democratic, homes were competent and well-adjusted (Baumrind, 1971, 1991).

In a preliminary analysis we asked, Do these competent parents use the least physical punishment? The answer is clearly “no” with their preschool children. Ninety percent (9 of 10) of Authoritative couples at T1 had scores at or above the mean on the physical punishment scale, and Authoritative or Democratic parents were not disproportionately classified in the Green zone. Thus, the higher competence and lesser maladjustment of the preschool children of the most effective parents was not due to their being spanked infrequently.

Parent types did differ, however, by the likelihood that members would resort to overly severe physical punishment. Thus families classified in the Red zone were disproportionately either Authoritarian-Directive or Rejecting/Neglecting (90% at T1, 75% at T2, 83% at T3), and no Authoritative parent at any time period fell into the Red zone, although one Democratic parent did⁵. Furthermore, both absolute and relative spanking frequency of Authoritative couples decreased rapidly after Time 1 with only 40% at or above the mean at T2, compared to 58% of all other parents, and by T3 with only 17% at or above the mean, compared to 42% of all other parents. Thus by early adolescence, when we in common with other specialists believe physical punishment to be developmentally inappropriate, Authoritative and Democratic parents were significantly less likely than other parents to use physical punishment. Perhaps their firm enforcement policies throughout childhood were successful in achieving a desirable level of behavioral compliance by adolescence.

(Question 7): Finally we asked -- is physical punishment associated with more detrimental child outcomes than verbal punishment?

The answer is clearly no. Total verbal punishment was negatively associated with competence, and positively associated with problem behavior at each age and for most outcomes, typically to a greater degree than the associations between these child outcomes and total physical punishment.⁶ Thus the negative outcomes associated with normative verbal punishment were at least as pronounced as those of normative physical punishment.

To Summarize These Results

Prior to removing the few parents whose use of physical punishment was unusually severe for this population and controlling the methodological artifacts that could account for the associations, frequency of physical punishment was associated with detrimental child outcomes, as anti-spanking advocates such as Straus claim. However, once the Red zone families were removed, there were few significant associations left to explain between child outcomes and dimensional or categorical measures of normative physical punishment. Furthermore, the correlations with detrimental child outcomes of physical punishment did not exceed those of verbal punishment. When alternative explanations, including the adolescents' self-reported favorable perception of their parents, are considered, there are no effects of normative physical punishment on child or adolescent outcomes. The apparent effects of NPP are explained by baseline child

misbehavior and third variables that contribute to a pattern of rejection and overcontrol in which reliance on physical punishment is embedded. The 3 children (all girls) of parents who totally abstained from spanking at all time points, were not more competent by adolescence than those whose parents spanked occasionally. All were prosocial, but two were very low on self-assertiveness and the one who was self-assertive and achievement-oriented manifested severe internalizing and externalizing symptoms.) Unexpectedly, even the presence of above-average frequency of normative physical punishment represented by the Orange zone did not attenuate at all the positive outcomes associated with Authoritative or Democratic parenting. Thus we found no evidence for unique detrimental effects of normative physical punishment.

To my knowledge this is the only study using high quality data in a prospective longitudinal design to assess the effects of normative physical punishment, after controlling for the following methodological artifacts: shared source variance, the intervention selection bias introduced by baseline child misbehavior, and plausible third parenting variables that were associated with both frequency of use of normative physical punishment and detrimental child outcomes. This is one of the few studies to contrast the effects of normative physical punishment with another aversive disciplinary intervention, and to contrast the effects of “no spanking” with those of “low frequency” spanking.

Two important limitations of the study should be noted. First, the population sampled is limited to predominantly middle-class European American families residing in a liberal University community at a particular historical period. Second, failure to find adverse causal outcomes of NPP by early adolescence cannot rule out the possibility that adverse outcomes would emerge in adulthood.

SOCIAL POLICY IMPLICATIONS

I will use the remaining time to editorialize about the social policy implications of the social science research on spanking effects.

In his delightful book Statistics as Principled Argument Robert Abelson (1995) writes, “Every student in the social sciences is to a greater or lesser degree taught to be reluctant to draw causal conclusions from correlations, but it is surprising how causal implications nonetheless sneak insidiously into interpretations of correlations” (p.181). The cock crows and the sun subsequently rises, but does the cock cause the sun to rise? Epidemiologists note in jest the significant correlations between yellow fingers and lung cancer. But are yellow fingers a causal risk factor for lung cancer, or merely an impurity attached to cigarette smoking? Spanking is somewhat associated with certain adverse child outcomes, but is it a causal risk factor for those outcomes? I conclude that it is not. In our study, third variables were well measured and when entered into the equation accounted fully for the apparent detrimental effects of spanking on child outcomes.

Power and Chapieski (1986) qualify their conclusion that physical punishment is an ineffective disciplinary strategy as follows, “It is important to note, however, that reliance on physical punishment, not physical punishment itself, was the critical variable...preliminary analyses showed no significant differences between the occasional-punishment and no-punishment groups” (p.274). By stigmatizing any use of spanking, professional consultants will discourage parents who rely on spanking from learning how to employ its conditional use instead – that is to initially use spanking or another punishment in combination with reasoning primarily as a back-up for time-out or milder tactics, with spanking eventually phased out in favor of greater reliance on reasoning and negotiation. Larzelere (2001) points out that in his study the conditional use of spanking rather than reliance on spanking characterizes effective parents. Similarly, in our study Authoritative parents who used more than average frequency of spanking with their preschoolers, did not rely on this tactic and phased it out in favor of more negotiated strategies of parental control, were outstandingly successful.

Based on the small body of relevant research evidence reported in his recent review of the child outcomes associated with nonabusive customary physical punishment, Larzelere (2000, p.215) offers the following guidelines for effective spanking:

1. Not severe enough to cause more than a moderate level of distress
2. Under control and planned, not impulsive
3. Preferably between ages 2 to 6 and phased out soon after
4. Used in conjunction with reasoning and explanation
5. Used privately
6. Motivated by child-oriented and not parent-oriented concern
7. Used after a single warning to enforce a directive or time-out
8. Used flexibly with recourse to other disciplinary tactics, rather than increasing the intensity of spanking.

Larzelere (2000) documents the critical methodological weaknesses of the studies on spanking effects, and in particular the failure of most studies to account for baseline child misbehavior, which he identifies as a critical “intervention selection bias”. In the methodologically strongest study of the handful of studies Straus cites correctly as more conclusive because they control for baseline child misbehavior, Gunnoe and Mariner (1997) found detrimental effects of spanking in some subgroups of children, but not in others. The other 4 studies Straus cites (Straus and Stewart, 1999) as causally conclusive (Brezina, 1999; Simons, Lin and Gordon, 1998; Straus and Paschall, 1998; Straus, Sugarman and Giles-Sims, 1997) because they include the necessary control for baseline child misbehavior suffer from equally serious methodological flaws, such as shared source variance by reliance on a single reporter or inclusion of abusive parents and/or adolescents in the sample when generalizing to normative spanking of children.

Most investigators (e.g. Deater-Deckard & Dodge, 1997; Ellison, 1996; Gunnoe & Mariner, 1997) other than Straus report that effects of physical punishment vary by ethnicity, child’s age, and parents’ ideology. By contrast with the detrimental outcomes associated with spanking generally found in uncontrolled studies of middle-class intact

white populations, neutral or beneficial outcomes of normative physical punishment are more typically found in working class, conservative Protestant, and ethnic minority families. These group differences may be explained by cognitive mediators: for example, Gunnoe and Mariner report “that the effects of spanking depend on the meaning children ascribe to spanking” (1997, p.768) which in turn depends upon the normative standards of the community, and the extent to which the child perceives the parent as loving and responsive and committed to the child’s welfare. Similarly, Rohner and his colleagues (Rohner, Bourque and Elordi, 1996; Rohner, Kean and Cournoyer, 1991) concluded from several studies of youth in St. Kitts and Georgia that there are no negative correlates of corporal punishment for youth who perceive their parents as loving and fair.⁸

The majority of U.S. adults questioned in a recent survey by Yankelovich (2000) continue to regard it as “appropriate to spank a child as a regular form of punishment” (Question 41), and their position is shared by most children and adolescents. Several studies (e.g., Graziano and Namaste, 1990; Catron and Master, 1993; Siegal and Cowen, 1984; and Siegal and Barkley, 1985) report a high level of acceptance by young adults, including college students, of the use of spanking by their parents during childhood⁹, and respondents generally state that they intend to spank their own children.

Parents in a democratic society rear their offspring with different values and perspectives that ensure desirable diversity in childrearing goals and outcomes. The state has significant interests in the well-being of its youth, but in the absence of compelling evidence that socially approved practices have harmful effects, it promotes children’s welfare by respecting family privacy and parental autonomy in childrearing decisions, thus protecting the supportive and guiding features of family life that contribute to children’s well-being and minimizing unnecessary intrusions into family life that are psychologically threatening to children by undermining their trust in parental authority, even when intended to advance their “best interests”. The ethical problem governing state intervention into family life is to determine when on balance state intervention will yield greater benefit than harm to children.

Similarly, consultants should weigh the potential costs to children against the expected benefits of the advice they give parents. (Baumrind and Thompson, in press). Professional advice that categorically rejects any and all use of a disciplinary practice favored and considered functional by parents is more likely to alienate than educate them (Mosby, Rawls, Meehan, Mays, & Pettinari, 1999). Patterson’s research documents the high rate of parental noncompliance with professional advice that contradicts parents’ own disciplinary preferences based on their personal experience and cultural norms (Patterson & Chamberlain, 1988).

Corrective punishment is intended to enforce short term compliance with parental directives and so to limit oppositional behavior, and prevent coercive cycles from becoming established. Although I do not regard spanking as less humane than other forms of punishment, I am not an advocate of spanking. Evangelicals such as Dr.

James Dobson who advises spanking as an antidote to “stiff-necked rebellion” because “pain is a marvelous purifier” (1970, pp.14, 16) is clearly a pro-spanking advocate. I am not. Spanking need not be impulsive and reactive, but often is by comparison with planned, goal-oriented responses by parents to children’s misbehavior, such as time-out and deprivation of privileges, which require some restraint and forethought. As Straus and Mouradian (1998) have documented, impulsive spanking is associated most strongly with harmful outcomes.

But, no single intervention strategy, including the preferred methods of time-out and explanation meet the needs of a heterogeneous population with diverse childrearing priorities and values, and no one strategy is maximally effective with all children at all times. Kochanska and colleagues (Kochanska, Coy and Murray, 2001) found significant links between committed compliance by young children with their mother’s directives, and children’s fearfulness and shyness. Fearful children are more easily conditioned to inhibit transgression than bold children, who are more likely to defy parental authority. Punishment, in particular physical punishment, is not only functionally superfluous for shy, fearful children (Lepper, 1981), but may be traumatic. Time-out with a barrier backup is the alternative favored by most parent educators with bold, defiant children. However, with defiant children, mothers in Roberts’ studies (Robert and Powers, 1990) preferred a “two-swat” backup rather than a barrier backup to enforce time-out. Whether parents can and will use an alternative backup, such as a barrier, with a defiant child, especially in homes where space and time are limited remains to be studied.

Punishment of any kind is intended to be aversive, and as such is never without costs to be balanced against its benefits. In that vein Joan McCord (1996, 1997) carries Straus’ injunction against physical punishment to its logical conclusion with a claim that all punishment has unintended detrimental effects and should be avoided. Articulate critics of disciplinary spanking, such as McCord and Straus have made us all more mindful of the costs of hasty recourse to aversive discipline, when induction, positive reinforcement and modeling might well accomplish parents’ more encompassing childrearing objectives. Nevertheless, I take exception to arguments premised on the unwarranted claim that causally relevant studies have shown that normative physical punishment harms children and society.

My understanding of scientific argument is that it begins with plausible substantive hypotheses and then proceeds to probe, not prove, the credibility of these hypotheses with empirically derived evidence. What distinguishes a scientific hypothesis from a layperson’s hunch or an ethical judgment is sufficient precision and specification to be refuted or amended in response to empirical evidence. But Straus begins his research with an assumption, not an hypothesis, when he states in the preface to his book Beating the Devil out of Them (1994), “the assumption that guided this research is that corporal punishment, by itself, has harmful psychological side effects for children and hurts society as a whole” (p.xii), and later in his preface when he writes that the problems likely to beset a spanked child “range from attacks on siblings to juvenile delinquency, wife beating, depression, distorted sexual behavior, to lower occupational

success and income” (1994, p. xii). When a scientist begins his or her research with an already formed conclusion, as Straus does, it is likely that the initial bias will be confirmed, not amended or rejected by the ensuing evidence.

The implication that spanking is a proven cause of personal and social pathology is not only scientifically misleading, but also diverts attention from physical abuse, systemic causes of violence associated with injustice and poverty, and neglect of children’s best interests in foster care and child welfare. The results reported here suggest that variations in the complex pattern of childrearing, not whether parents include normative physical punishment among their disciplinary options, accounts for the significant differences in child outcomes.

If the *effectiveness* of a disciplinary practice is the extent to which it has the desired outcome as typically used, and *efficacy* is the power of a practice to produce the desired effect when properly used, then efficacy should concern practitioners (e.g., pediatricians, clinicians, and parent educators) more than effectiveness. By being consistently firm, rational, and responsive and by proactively teaching the child to behave morally, caregivers can minimize the need for spanking or other punishment, as well as render punishment more efficacious.

In this study Authoritative, and to a somewhat lesser extent Democratic, parents were optimally efficacious, whether or not they spanked their children, as almost all did when their children were preschoolers. Although optimal parenting may vary across cultures, it is likely in any culture to have certain of the generic features that characterize authoritative parents. These features include deep and abiding commitment to the parenting role, intimate knowledge of children’s developmental needs; respect for a particular child’s individuality and desires; provision of structure and regimen appropriate to the child’s developmental level; readiness to establish, and disciplinary strategies to enforce, behavioral guidelines; and cognitive stimulation, effective communication, and use of reasoning to ensure children’s understanding of parents’ goals and disciplinary strategies. Clear limits that are firmly enforced during the early years and that occur within the context of a rational-authoritative parent-child relationship should maximize committed compliance and minimize the need for punishment as the child matures.

The take-home message I have tried to convey is this: The certitude with which one conveys one’s findings to the public should not exceed the limitations of one’s science. Public trust in the integrity of evidence-based social policy recommendations is undermined by selective use of weak evidence to support an unqualified assertion that physical punishment harms children and hurts society, and therefore should be criminalized or stigmatized. Methodologically strong studies have not established that normative physical punishment is a causal risk factor for the detrimental child outcomes with which it may be associated. Although a value judgment that spanking is wrong is properly defended by its adherents on ethical grounds, a blanket injunction against disciplinary spanking is not warranted by causally relevant scientific evidence.

FOOTNOTES

¹ Scales were similar across time periods, but with each successive developmental stage additional items were included and worded to match the increasingly differentiated status of the maturing child.

² Note, however, that because no data on the children were available until the preschool years, the possible prior effects of punitive parenting on children's "baseline" behavior could not be measured and controlled.

³ Demandingness refers to the claims parents make on the child to become integrated into the family by their maturity demands, supervision, disciplinary efforts, and willingness to confront the child who disobeys. Responsiveness refers to actions that intentionally foster individuality, self-regulation and self-assertion by being attuned, and that are supportive and acquiescent to the child's special needs and demands.

⁴ Strassberg and colleagues (1994), who are among the very few investigators to report results for children who are not spanked, report that the 4% of children who were not spanked during the year preceding kindergarten were significantly less likely than other children to react aggressively to provocations from peers, although they were not less likely to bully or to aggress instrumentally, and there was no correlation between frequency of spanking and any index of child aggression. Unfortunately, despite its longitudinal design the investigators did not control for base rate measures of child aggression, which leaves the direction of the effect indeterminate. The results could not confirm a parent effect because a child main effect could plausibly account for the associations – that is, socially assertive or aggressive children may provoke physical punishment, and also respond aggressively to provocations from peers.

⁵ However, unlike Red zone membership, Orange zone membership did not increase the detrimental outcomes associated with Rejecting-Neglecting parenting or decrease the effectiveness of Authoritative or Democratic parenting. Generally speaking, within parent type, children of Orange zone parents were not less competent or more maladjusted than other children, although we hypothesized that this would be the case especially for children of Rejecting/Neglecting parents.

⁶ Unlike the effects of removing parents in the Red Zone for physical punishment, doing the same for verbal punishment did not greatly attenuate its detrimental effects.

⁷ The strongest evidence includes three randomized clinical studies with Roberts as an author (Roberts 1982, 1988; Roberts & Powers, 1990) and three sequential analyses (Larzerlere & Merenda, 1994; Larzerlere, Schneider, Larson, & Pike, 1996; Ritchie, 1999)

⁸ Only youth who reported that they were punished physically more than 3 times a week and perceived their parents as rejecting, also reported symptoms of psychological maladjustment on a Personality Assessment Questionnaire.

⁹The documented acceptability of physical punishment to children and youth [Catron and Masters (1993), Siegal and Cowen (1984), and Siegal and Barkley (1985)] may explain the unexpected absence of detrimental outcomes associated with NPP among the adolescents in our study. In Siegel et al.'s studies (Siegal & Cowen, 1984; Siegal & Barclay, 1985), working class Australian children ranging in age from 5 to 17 were asked to judge the acceptability of the use of four different methods of discipline in fictitious vignettes involving a parent and a 4 year old. All expressed approval for both induction (in which the parent reasons with the culprit and points out the harmful consequences of the transgression) and physical punishment (hitting the child) in preference to both permissiveness (not intervening in the belief that the culprit would learn independently) and love withdrawal (disengaging temporarily).

REFERENCES

Abelson, R. P. (1995). Statistics as principled argument. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Baumrind, D. (1971). Current patterns of parental authority. Developmental Psychology Monograph, 4 (1, Pt. 2), 1-103.

Baumrind, D. (1991). The influence of parenting style on adolescent competence and substance abuse. Journal of Early Adolescence, 11(1), 56-94.

Baumrind, D. (1994). The social context of child maltreatment. Family Relations, 43(4), 360-368.

Baumrind, D. (1995). Child maltreatment and optimal caregiving in social contexts. New York and London: Garland Publishing, Inc.

Baumrind, D., & Thompson, R.A. (in press). The ethics of parenting. In M. Bornstein (Ed.), The Handbook of Parenting (2nd ed.).

Brezina, T. (1999). Teenage violence toward parents as an adaptation to family strain: Evidence from a national survey of male adolescents. Youth and Society, 30(4), 416-444.

Catron, T.F., & Masters, J.C. (1993) Mothers' and children's conceptualization of corporal punishment. Child Development, 64, 1815-1828.

Cook, T. D. & Campbell, D.T. (1979). Quasi-experimentation: Design & analysis issues for field setting. Boston: Houghton Mifflin.

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(3), 161-175.

Dobson, J. (1970). Dare to discipline. Wheaton: Tyndale House.

Ellison, C. G. (1996). Conservative Protestantism and the corporal punishment of children; Clarifying the issues. Journal for the Scientific Study of Religion, 35, 1-16.

Friedman, S. & Schonberg, S.K. (1996). Consensus Statements. Supplement to Pediatrics, 98(4), 853. (Also see in same issue their Qualifying Statements [854-856])

and Personal Statement [857-858] – Friedman and Schonberg were organizers of the conference reported in this Supplement.)

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.

Graziano, A.M., Namaste, K. (1990). Parental use of physical force in child discipline: A survey of 679 college students. Journal of Interpersonal Violence, 5, 449-463.

Hill, A. B. (1965). The environment and disease: association or causation? Proc R Soc Med, 58, 295-300.

Kochanska, G., Coy, K.C., and Murray, K.T. (2001). The development of self-regulation in the first four years of life. In Child Development, 72(4), 1091-1111.

Larzelere, R.E. (2000). Child outcomes of non-abusive and customary physical punishment by parents: An updated literature review. Clinical Child and Family Psychology Review, 3(4), 199-221.

Larzelere, R.E. (2001). Combining love and limits in authoritative parenting: A conditional sequence model of disciplinary responses. In J.C. Westman (Ed.), Parenthood in America: Undervalued, Underpaid, Under Siege. Madison, WI: University of Wisconsin Press.

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., 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.

Lepper, M. (1981). Intrinsic and extrinsic motivation in children: Detrimental effects of superfluous social controls. In W.A. Collins (Ed.), Aspects of the development of competence: The Minnesota symposium on child psychology. Vol. 14. Hillsdale, NJ: Erlbaum.

McCord, J. (1996). Unintended consequences of punishment. Pediatrics, 98(4), 832-834.

McCord, J. (1997). On discipline. Psychological Inquiry, 8(3), 215-217.

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 (3), 489-521.

Patterson, G. R., & Chamberlain, P. (1988). Treatment process: A problem at three levels. In L. C. Wynne (Ed.), The state of the art in family therapy research: Controversies and recommendations (pp. 189-223). New York: Family Process Press.

Power, T. G., & Chapieski, M. L. (1986). Childrearing and impulse control in toddlers: A naturalistic investigation. Developmental Psychology, 22, 271-275.

Roberts, M.W. (1982). Resistance to timeout: Some normative data. Behavior Assessment, 4, 239-248.

Roberts, M.W. (1988). Enforcing chair timeouts with room timeouts. Behavior Modification, 12, 353-370.

Roberts, M.W. & Powers, S.W. (1990). Adjusting chair timeout enforcement procedures for oppositional children. Behavior Therapy, 21, 257-271.

Rohner, R.P., Bourque, S.L. & Elordi, C.A. (1996). Children's perceptions of corporal punishment, caretaker acceptance, and psychological adjustment in a poor, biracial southern community. Journal of Marriage and the Family, 58, 842-852.

Rohner, R.P., Kean, K.J., & Cournoyer, D.E. (1991). Effects of corporal punishment, perceived caretaker warmth, and cultural beliefs on the psychological adjustment of children in St. Kitts, West Indies. Journal of Marriage and the Family, 53, 681-693.

Rothman, K. J. & Greenland, S. (1998). Causation and causal inference. In K. J. Rothman & S. Greenland (Eds.), Modern epidemiology. (pp. 7-28). Philadelphia: Lippincott-Raven.

Siegal, M., & Barclay, M. S. (1985). Children's evaluations of fathers' socialization behavior. Developmental Psychology, 21, 1090-1096.

Siegal, M., & Cowen, J. (1984). Appraisals of intervention: The mother's versus the culprit's behavior as determinants of children's evaluations of discipline techniques. Child Development, 55, 1760-1766.

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, 467-479.

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. (1994). Beating the devil out of them: Corporal punishment in American families. Lexington, MA: Lexington/Macmillan Books.

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 longitudinal study. Paper presented at the 14th World Congress of Sociology, Montreal, Canada.

Straus, M.A. & Stewart, J.H. (1999). Corporal punishment by American parental 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 Pediatric Adolescent Medicine, 151, 761-767.

Yankelovich, D. (2000). What grown-ups understand about child development: A national benchmark survey, CIVITAS Initiative, ZERO TO THREE, BRIO Corporation, Researched by DYG, Inc.