In This Issue:

- Introduction to This Double Issue
  Al K. DeRoy ............................................. 1

- Spanking By Parents—Some Ideas on
  Measurement and Analysis of a Neglected
  Risk Factor for Serious Mental Health
  Problems—Murray A. Straus .......................... 3

- Bootstrapping the Way to Valid
  Diagnostic Interviews—Lee Robins .............. 9

- Measurement Modeling: A Tool for
  Investigating the Comparative Anatomy
  of Related Instruments—Fred B. Bryant ...... 14

- HaPI Thoughts ...................................... 17

The Behavioral Measurement Letter
Behavioral
BMDS  Measurement
Database
Services
PO Box 110287 • Pittsburgh, PA 15232-0787
Spanking By Parents – Some Ideas on Measurement and Analysis of a Neglected Risk Factor for Serious Mental Health Problems

Murray A. Straus

Both public and professional opinion increasingly are turning against the use of corporal punishment (CP), even occasional use (Straus & Mathur, 1996). In support of this increasingly negative evaluation of CP is a growing body of evidence suggesting that CP has many harmful side effects, such as depression, low educational achievement, and masochistic sex (Straus, 1994). Moreover, it has been found that despite its effectiveness in modifying or suppressing undesirable behaviors in the short run, CP tends to be counterproductive in the longer run (Straus, Sugarman, & Giles-Sims, 1997). The evidence, however, is far from definitive. This combination of inconclusive evidence on the one hand, and growing doubts about the advisability of CP on the other, is probably the major reason for the increase in research on CP in the past five years, a trend that is likely to continue.

This article discusses measurement of CP and analysis of data on CP in the research context. For the purposes of the article, CP is defined as the use of physical force by a parent with the intention of causing the child to experience bodily pain, but not injury, for correction or control (Straus, 1994). Among the most frequent forms of CP are spanking on the buttocks, slapping the hand of a toddler, and slapping the face of a teenager.

Gathering and analyzing data on the seemingly simple phenomena of spanking and other legal forms of CP, and on the effects of CP, turn out to have many subtle aspects and pitfalls. This article covers only some of the many issues involved. However, these and others will be examined in detail in the methodological chapter of my forthcoming book, Corporal Punishment in Social Context.

Measuring Corporal Punishment

Specific Acts Versus General Questions

Lists of Specific Acts. Studies of marital violence and sexual abuse have found that the more specific acts the respondent (either parent or child) is asked about, the higher the prevalence rate found (Straus, 1990a). Presumably, this is because mention of specific acts aids in recall. Similarly, asking about specific acts of CP should assist in determining prevalence. Furthermore, a list of specific acts can be used to create subscales to differentiate ordinary CP from more severe CP, as in the Parent-Child revision of the Conflict Tactics Scales (Straus, Hamby, Finkelhor, Moore, & Runyan, 1998), and to provide data that can be used to examine the prevalence and correlates of each specific act of CP.

General Questions. Despite the finding of higher prevalence rates by listing specific acts, extremely high rates have been found by asking only one question on CP. For example, the National Longitudinal Survey of Youth, the NLSY (Baker, Keck, & Quinlan, 1993) asked mothers of 3-5 year old children about spanking in the previous week. Using only this question, it was found that 64% of the mothers surveyed reported doing so, and that they had spanked an average of 3.2 times in the previous week (Giles-Sims, Straus, & Sugarman, 1995). An important consideration in using a general question rather than a list of specifics is the availability of time or space. Although a single question may underestimate the prevalence of CP, there are many situations in which the time available in an interview or space available in a questionnaire permits the inclusion of only one or two general questions concerning CP.

Terminology for General Questions. When a general question is used rather than a list of specific acts, the words used to indicate CP are particularly important. Parents almost never use terms such as “corporal punishment” or “physical punishment.” Instead they describe what they do as a “swat,” “spanking,” “whooping,” “whack,” or in England, a “smack.” Spanking seems to be the most widely
Spanking (continued from Page 3)

used term in the U.S. Consequently, to be consistent with everyday language, many interview studies, such as the National Longitudinal Survey of Youth (Baker et al., 1993), use “spanking” as a generic term for CP. The spanking question in the NLSY, which itself is part of the Home Observation for Measurement of the Environment (HOME) scales (Caldwell & Bradley, 1984), is: “Sometimes kids mind pretty well and sometimes they don’t. About how many times, if any, have you had to spank your child in the past week?” Although the parents are asked about “spanking” here, experience shows that parents do not usually restrict its meaning to hitting a child on the buttocks. Nevertheless, although the term “spanking” is often used to refer to most forms of CP that are within a socially acceptable range of severity, for example a slap on the hand or a pinch of the arm, some parents limit its meaning to hitting a child on the buttocks. Consequently, it is best to use a phrase such as “spank or slap” or “physical punishment like spanking, slapping, ear twisting, or hitting in some other way” in a general question concerning use of CP.

Referent Period

“Referent period” is defined here as a period of time in which CP may have been used. For preschool age children, the shorter the referent period asked about, the more accurate the data. Thus, the HOME scale uses a one week referent period. However, short referent periods have the disadvantage of making the zero category ambiguous. For example, when “previous week” is the referent period, it is very unlikely that “none” identifies children who do not experience CP because there are previous weeks in which it could have been used. This led critics of the Straus et al. longitudinal study of CP (Straus, Sugarman, & Giles-Sims, 1997) to argue that the improved behavior two years later of the “none” group and the worsened behavior of those spanked once or more in the previous week demonstrated that infrequent CP has long term benefits, and that CP is harmful only when it is done frequently, such as one or more times a week.

The optimum referent period also depends on the age of the child. For children age 10 and over, and especially for teens, using a one-week referent period would underestimate the prevalence of CP because older children are hit much less frequently than toddlers. Among those parents who continue to use CP with teenagers, the mean number of times is at least six per year (Straus & Donnelly, 1994). This is an astoundingly high frequency, but it is still infrequent enough that use of a one-week period would falsely classify most of those teens as not experiencing CP.

Probably the best solution is to use more than one referent period. Thus, for preschool children one could ask about the last week, the last six months, and whether the parent has ever used CP. With teens, the most appropriate referent periods might be the preceding month and preceding year, along with a determination as to whether the parent has ever used CP.

Adult Recall Referent Age

Many studies ask adults, and especially college students, whether their parents used CP. Presumably, they respond about whether the parent had ever used CP. Therefore, it is probably best to ask adults about a specific age referent because this yields age-specific prevalence rates, and more important, resultant data can be analyzed by reference to specified ages. It is also recommended that the referent age not be that when CP is most prevalent (ages 2-4) because most adult respondents will not remember much of what happened to them then. Instead, the youngest referent age used in adult recall studies should be 13 (usually 7th grade), a young enough age for CP to be prevalent (30-40%) still, but old enough so that almost all respondents will be able to remember.
Spanking (continued from Page 4)

Suggested Adult Recall Questions

The following question is similar to the question on CP used in the National Family Violence Surveys and that which provided the data for about half the chapters in Beating the Devil Out of Them (Straus, 1994) and several other papers.

I'd like to ask you about your experiences as a child. Thinking about when you were 13 years old, about how often would you say your mother or stepmother used physical punishment like spanking, slapping, ear twisting, or hitting you in some other way?

(0=Never, 1=Once, 2=Twice, 3=3-5 times, 4=6-10 times, 5=11-20 times, 6=More than 20 times, 9=Did not live with mother/stepmother at the age)

How about your father or stepfather? Again, thinking about when you were 13, about how often would you say he used physical punishment, like spanking, slapping, ear twisting, or hitting you in some other way, that year?

(Choices as above.)

Frequency of Occurrence Response Categories

My research experience indicates that numerical frequency response categories, such as those used in the question above and in the Conflict Tactics Scales (Straus et al., 1996; Straus et al., 1998), are preferable to qualitative categories such as “sometimes” and “frequently.” Although numerical frequency data on CP should not be taken too literally, they have two advantages: (a) Numerical frequency categories avoid subject-to-subject differences in mentally assigning a quantitative range to each qualitative category, and (b) they can be used to compute chronicity (frequency of CP usage by those who use CP) by replacing category values with their midpoints, i.e., replacing code values 3 through 7 with 4, 8, 15, and 25 respectively.

Prevalence and Chronicity

Starting at about five years of age, the frequency distribution for CP becomes more skewed with each additional year of age because a decreasing proportion of parents employ CP as their children age. The skewness is so extreme that, in studies of CP use at various ages, a mean calculated on the entire sample is extremely misleading. An alternative is to create separate measures for prevalence, a dichotomy of whether CP is or is not used, and chronicity, frequency of CP usage by those who use CP.

Research Design and Data Analysis

Age Specificity

As demonstrated below, when a CP study includes a range of ages, both questions and data analysis should be age specific.

Age-Specific Prevalence Rates. There are vast differences in prevalence rates among studies of CP. For example, one survey of a national sample of American parents found that 53% of children experienced CP in 1992 (Darco & Gelles). However, using data from the National Family Violence Surveys conducted in 1975, 1985 and 1994, I concluded that 95% of American children experience CP (Straus et al., 1998; Wauchope & Straus, 1990).

In this case, the huge difference in prevalence estimates is not due to error in either study, but rather to the fact that one included age-specific rates and the other did not. Darco and Gelles’ (1992) rate is the percent of all children in their sample who experienced CP during the preceding 12 months, whereas the 95% rate is the percent of children ages 3 and 4 only who experienced CP. If the objective is to provide information on how many American children experienced CP, the Darco and Gelles rate is misleading, even though accurate. The 95% rate is the better estimate of prevalence of CP among all American children at some time in their lives, because it is almost certain that the 95% rate applied to the older children in the samples when they were toddlers. Thus, if a sample includes a range of ages, statistics on prevalence...
and chronicity should be provided for children of different age groups rather for all children in the sample.

Age-Specific Analyses. The nature and meaning of the same act of CP may be very different for a 4-year-old than for a 14-year-old child. Consequently, when a sample includes children of different ages, statistical testing for interactions with age is very important. For example, if data on a sample of children ages 2-14 is being analyzed using ANOVA, age might be specified as one of the independent variables using categories such as 2-5, 6-9, 10-12, 13-14. For multiple regression analyses, a multiplicative interaction term (age by CP) might be used.

Causal Direction

Perhaps the most difficult methodological problem in research on the effects of CP is posed by the fact that child behavior problems lead parents to spank. Thus the repeated finding that the more CP parents use, the worse the behavior problems of the child does not necessarily show that CP has harmful effects, or even that CP is not effective in reducing misbehavior (as I erroneously argued in the past). Randomized trials are needed that address both causal direction and confounding with other variables. There have been several such experiments with toddlers (e.g., Roberts & Powers, 1990), and these show that CP increases compliance, at least in the short run. However, in my opinion, there is sufficient evidence suggesting long-term harmful effects from CP to make future experiments of this type unethical. Fortunately, there are feasible alternatives.

Control for Level of Intervention. When cross-sectional data is used to investigate possible harmful effects of CP, the analysis is more complex than simply determining causal direction because of "intervention selection bias." This can be illustrated using a hypothetical study of cancer treatment. If the death rate of persons receiving treatment for cancer were compared to the death rate of the general population, it would show a much higher death rate among those treated. Moreover, it probably would also show that the more intensive the treatment, the higher the death rate. Similarly, when studying child misbehavior, even positive disciplinary techniques will seem to show adverse effects. Thus, a study of children in two Minnesota cities (Straus & Mouradian, in press) found that the more CP parents used, the greater their child's score on a measure of antisocial behavior. However, they also found that the more parents used alternatives such as reasoning, time outs, and deprivation of privileges, the greater the antisocial behavior score. These findings reflect the fact that when a child misbehaves, parents try many possible approaches to correcting the misbehavior. Thus, punishment is confounded with amount, or level, of intervention on the part of parents.

One approach to unconfounding use of CP and level of intervention is to create a variable to measure level of intervention. This approach requires data on disciplinary practices employed in addition to CP. The Straus and Mouradian (in press) study discussed above, for example, collected frequency of use data for explanation, time out, deprivation of privileges, and CP. These four variables were summed to create a level of intervention variables. This composite variable was then used as an additional independent variable in an ANOVA and was found to have a much stronger relation to the children's antisocial behavior than any of the other variables. Despite the extremely strong relation between level of intervention and antisocial behavior, however, Straus and Mouradian found a significant net effect for CP because even after adjusting for level of intervention, the differences among the mean scores showed that the more CP used, the higher the antisocial behavior exhibited. Nonetheless, although controlling for level of intervention reduces the chance that the findings are artifacts of intervention selection bias, this strategy is by no means definitive.

Spanking Cessation Experiments and Quasi-Experiments. The prevalence rate of CP is so high that a meaningful experiment can be carried out simply by randomly assigning one group of parents to a "no-spanking" parent education
program and the other to a control group of an alternative parent education program. The treatment families can then be tracked to determine the extent to which CP has been reduced or eliminated, and if so, whether children in the treatment group have lower rates of behavior problems than the children in the control families.

Quasi-experiments are also feasible, such as one I am conducting in two small cities in Minnesota. In this study, one city is the site of a community-wide, no-spanking educational program, and the control community is of similar size and socioeconomic characteristics.

Prospective Studies. A prospective study itself does not control for causal direction. To do so, there must be data on the dependent variable (for example, antisocial behavior) at time 0, time 1, time 2, etc. so that change in the level of the antisocial behavior subsequent to CP can be measured.

Control for Confounding with Physical and Psychological Abuse

Almost every parent who engages in severe assaults classified as “physical abuse” also engages in the legal assaults classified as corporal punishment. Furthermore, there is also an overlap of CP with psychological attacks (Vissing, Straus, Gelles, & Harrop, 1993). Although parents who severely assault a child are only a small proportion of those who use CP, it is possible that even this small proportion could account for the relationship between CP and child behavior problems. A few studies (e.g., Strassburger, Dodge, Pettit, & Bates, 1994; Straus. 1990b) have taken this into consideration and still found a relationship between use of CP and child behavior problems. Nevertheless, it is important to include physical and psychological abuse in studies of CP so that abused children can be eliminated from consideration before the data is analyzed. Analysis can also be done in a way that avoids confounding the effects of CP and those of abuse, for example, by using a typology such as 0=no CP or more severe assaults. 1=CP but no severe assaults. 2=severe assaults. If the Conflict Tactics Scales for Parents and Children, the CTSPC (Straus, et al., 1998), is used to obtain data on CP, its Severe Assault and Psychological Aggression subscales will provide data on physical and psychological abuse as well.

Child Abuse Reporting Obligations

CP by parents is legal in every state of the U.S. Legal CP includes not only spanking and slapping, but also hitting a child with a belt or other object, provided there is no injury that requires medical care. Moreover, most people consider CP to be an appropriate thing for a loving parent to do “when necessary.” Thus, in both the legal and informal norms of American society, CP does not fall under the mandatory abuse reporting laws. In principle, CP can be so chronic that most readers of this article would consider it physical abuse. In practice, child welfare and protection agencies are so understaffed that when such a case is reported, unless there is some special circumstance, it is likely to be filed away without investigation, as are about half of all such reports. Nonetheless, I believe that if information is disclosed during an interview which indicates that a child has been injured and/or is in imminent danger of injury, the child’s welfare takes precedence over promises of confidentiality and, therefore, such information should be reported, even in states where researchers are not mandated to do so.

This article is the result of a research program on corporal punishment supported by National Institute of Mental Health grants R01MH40027 and T32MH15061 and the University of New Hampshire. A publications list will be sent on request (preferably by e-mail to MAS2@CHRISTA.UNH.EDU).

References


Spanking (continued from Page 7)


Murray A. Straus, PhD, is Professor of Sociology and founder and current Co-Director of the Family Research Laboratory at the University of New Hampshire. He has served as President of the National Council on Family Relations (1972-73), Society for the Study of Social Problems (1989-90), and Eastern Sociological Society (1991-92), and has received various awards, the most recent being a Research Career Achievement Award from the American Professional Society on the Abuse of Children (1994). Dr Straus has written or been co-author of more than 200 articles on family relations, research methods and South Asia, and 15 books, including Sociological Analysis (1968), Handbook of Family Measurement Techniques (1990), Stress, Culture, and Aggression (1995), and Understanding Family Violence (1995).

The palest ink is clearer than the best memory.

Chinese Proverb