Punishment Revisited—Science, Values, and the Right Question:
Comment on Gershoff (2002)

Ross D. Parke
University of California, Riverside

In this comment on E. T. Gershoff (2002), the author notes the historical phases of punishment research. Punishment as a disciplinary tactic is best viewed as a packaged variable and therefore needs to be investigated in the context of other socialization practices. The role of parental values in this debate about punishment utilization and effectiveness merits more consideration. New directions in punishment research are also noted. These include the need for a family-systems perspective, a family-typology approach, a transactional model of punishment, the use of innovation observation and self-report methods, and more culturally sensitive paradigms.

Just as Rip Van Winkle in Washington Irving’s classic tale found that the world had changed after his 20-year nap, so I have found after a nearly 20-year break from my active study of punishment that society’s as well as psychologists’ understanding of the issue has changed. As a long-time student of this issue, I welcome this opportunity to reflect on the progress that has been made and on the progress that has not been made.

As this new synthesis (Gershoff, 2002), clearly has documented, a great deal more is known about the effects of punishment on children than was several decades ago. At the same time, it is important to inquire as to whether researchers are asking the right question. I present a brief overview of the recent history of the phases of inquiry into this issue in order to show that attempts to treat punishment as a separate variable are bound to failure. Instead, the inherent packaged nature of parental discipline renders the attempt to answer the question about the effects of corporal punishment on children a misguided one, or at least an inquiry that requires a different set of methodological approaches than typically used in the past. Finally, the issue of the ever-present interplay between science and values needs to receive more attention in the discussion of this topic.

A Historical Note

Research on the topic of punishment has gone through a series of phases over the past half century. Several phases can be distinguished: First, there was a descriptive phase (1940s–1960s), in which the main shape of the answer to the question about punishment effectiveness was provided. This was followed by an experimental phase in which the issue of punishment effectiveness was dismantled by experimental analysis (1960s–1970s). In the most recent phase, (1980–present) investigators turned to an examination of what were a major focus of Gershoff’s (2002) review, the processes that could account for the effects of physical punishment on children. Coincident with the process phase, a fourth phase emerged, the politicization of punishment phase, which alerted us to the underlying value-laden decisions surrounding the use of punishment as a disciplinary strategy. This politicization phase has continued to the present, and serves as a backdrop to the scientific debate about punishment of children (Straus, 1994).

Descriptive Phase

As Gershoff (2002) indicated, the classic studies of child rearing by the Gluecks, the McCords, and Sears and his colleagues established the links between punishment as a disciplinary tactic and children’s social and moral development which subsequent work has confirmed. This first phase—the descriptive phase—was characterized by several positive features, namely careful interviews of parents about their disciplinary practices and in some cases independent assessments of child outcomes (Bandura & Walters, 1959). These earlier descriptions involved real parents disciplining their own children in real-life settings. The concern about ecological validity that plagued later work was not an issue in this earlier work. However, sampling was not always very sophisticated or, at a minimum, was very restricted. As I note below, the samples were usually White, often middle class, and clearly of unknown representativeness. Developmental psychology had not yet discovered sociology with its greater sampling prowess. Unfortunately, in some cases the parent was the reporter of both the child-rearing techniques and the child outcomes (Sears, Maccoby, & Levin, 1957), a situation that continues to plague even some contemporary studies. Another problem that plagued earlier work was the confounding of a variety of aspects of punishment; thus, it was difficult to conclude whether the effects attributed to punishment were due to punishment alone or to the disciplinary practices which often covary with the use of punishment in real-life contexts. Take the early classic research of Sears et al. (1957) as an example.

Parental punishment often confounded intensity, frequency, consistency, and the nature of the relationship between parent and child. From this early work, it was extremely difficult to sort out the aspects of the package of variables which, in fact, accounted for the effectiveness or ineffectiveness of a particular cluster of disciplinary strategies.
Experimental Phase

Recognition of the multifaceted nature of the disciplinary package led to a second phase—the experimental analyses of punishment. In the 1960s, under the dual guidance of social learning theory (Bandura & Walters, 1963) and of experimental work on aversive conditioning with animals (Solomon, 1964), a series of lab-based analogue studies were conducted (Aronfreed, 1969; Deur & Parke, 1970; Parke, 1969; Parke & Walters, 1967). These studies were useful in sorting out the often covarying aspects of punishment. Using laboratory paradigms of forbidden toys or suppression of Bobo doll-directed hitting behavior, investigators showed clearly that punishment—in the form of an anxiety-inducing noxious buzzer—could in fact effectively suppress responses such as touching an attractive but forbidden toy or inhibit aggressive responses such as punching Bobo. Moreover and more important, researchers (Aronfreed, 1969; Parke, 1969, 1977) found that punishment was dependent on a variety of factors, including the timing of punishment (early vs. late), the intensity (high vs. low), the consistency (intra- vs. interagent), and the nature of the relationship between the child and the agent of punishment (warm, close vs. cold, distant).

In addition, the presence or absence of and the complexity of the verbal rationale accompanying the noxious event was important in demonstrating that punishment effectiveness was increased by the use of a verbal rationale (Cheyne & Walters, 1969; Parke, 1969). Perhaps more important was the discovery that these punishment parameters interacted with one another in achieving their suppression effects. For example, if punishment was well timed and delivered early in the response sequence, a lower level of intensity of punishment was sufficient to achieve suppression (Parke, 1969). Similarly, if the child had a warm nurturant relationship with the punitive agent, less intense punishment would suffice. Perhaps most dramatic were the demonstrations that the addition of a verbal rationale modified the operation of an array of other factors including timing, intensity, and the nature of the relationship (Parke, 1969). In each case, the importance of these accompanying variables was less when a verbal rationale was provided. Moreover, these earlier studies clearly showed that punishment was more effective in producing short-term suppression and that producing long-term internalization required the use of a rationale (Parke, 1969). Later work has shown that shifts in attributions and other cognitive processes undergird the movement from suppression to internalization (Perry & Perry, 1983).

This earlier work achieved several goals, including the task of unpackaging the set of disciplinary practices, which were often confounded in earlier interview studies. Second, these studies clearly demonstrated the effectiveness of punishment but at the same time underscored the complex set of contingencies involved in achieving varying levels of suppression. Punishment, then and now, is a many splendored variable. At the same time, these studies were limited in a variety of ways, most obviously by their reliance on laboratory contexts and the resulting concerns about the ecological validity of the findings. As Alfred Baldwin (1967) noted in his astute critique of the body of lab-based experimental studies inspired by social learning theory: “These studies created a mythology of childhood whereby effects achieved in the laboratory are assumed to be explanations of events that occur in real-life contexts. These studies confuse necessary and sufficient causality” (p. 41).

This kind of critique, as well as Bronfenbrenner’s (1974) indictment of the era’s studies of socialization as “the science of behavior of children in strange situations with strange adults” (p. 3), led to a reevaluation of strategies and brought into question the value of these kinds of studies. In defense of this era, however, it should be noted that if the aim is to analyze the operational effectiveness of various punishment parameters and to permit relatively clear conclusions about these effects, this approach is valid and valuable. If, on the other hand, the aim is to provide a more ecologically valid account of how socialization—including punishment—operates in more naturalistic contexts, alternative approaches are needed.

In addition to the lack of ecological validity, there are ethical problems with experimental studies of punishment. Society’s sense of ethical standards has shifted over the last several decades because of a variety of changes in its appreciation of children’s rights, the decreased acceptance of the use of deception, the need to more sensitively balance cost–benefit ratios, and the institutionalization of the review process. If one accepts as valid the conclusion that negative side effects, such as increased aggression are associated with the use of punishment (Gershoff, 2002), it is questionable whether experimental procedures involving punishment analogues are ethical. At the same time, it is acknowledged that the kinds of aversive events to which children in these studies are typically exposed are well within the range of normal experience for most children. Moreover, exposure to repeated patterns of punishment is more likely to produce these negative side effects than a single exposure to an unpleasant noise in a single laboratory session (Aronfreed, 1969). In any case, researchers clearly moved forward in this period, but their conclusions were still limited, especially in terms of generalizability to naturalistic contexts.

Process Phase

Since the 1980s, investigators have focused on the processes that may account for punishment effectiveness—a history well summarized by Gershoff (2002) in her synthesis. In this phase, the complexity of the processes that can account for the effectiveness or ineffectiveness of punishment are more clearly appreciated than in earlier eras. This phase represents an extension of the prior lab–experimental era, but with a wider range of processes placed under the research microscope and with more attention paid to ecological validity issues.

What Is the Right Question?

The problem with asking about the effects of punishment per se is that punishment is not a single variable but in reality is situated in a context of other practices—an insight recognized by the early researchers of the modern era. To borrow a term from the anthropologist Whiting (1975), punishment is a packaged variable that requires “unwrapping” to isolate the components that account for its effectiveness. At the same time, there is a risk of destroying or at least distorting the set of interrelated processes one hopes to understand.

If one takes seriously the view that socialization strategies represent packaged variables and adopts a typological approach,
one needs to ask whether he or she is asking the right question. The term right question has multiple meanings. First, is it possible to achieve a clear answer to the question? If not, is it scientifically defensible to frame a question for which it is improbable that a clear answer will be forthcoming? Second, if punishment is likely to covary with other practices, isolation of punishment from its naturally occurring context may not be an advisable strategy. Third, researchers need to distinguish audiences for whom they are seeking answers. If the goal is to inform the scientific community, experimental analyses of punishment may be a valid scientific approach to help understanding the parameters that control the operation of punishment. However, it is unlikely that researchers can generate experimental paradigms that would sufficiently capture the complexities of how punishment is used in situ. Perhaps it is time to cease asking about the impact of punishment, per se, and reframe our question to be: What are the effects of packages of parental disciplinary tactics—of which punishment may be only one component—on children’s short- and long-term development?

If the goal is to inform policymakers, a different approach may be needed that focuses on alternative strategies to punishment. The problem of effectiveness needs to be framed in terms of comparisons with other strategies. Are some of the alternatives less likely to produce the negative side effects associated with physical punishment (Gershoff, 2002; Parke, 1977)? Advice to parents and policymakers is likely to be effective only if recommendations about the alternatives to punishment are included and if both groups know the comparative costs and benefits of different strategies.

The Politicization of Punishment and the Issue of Values

As a backdrop to the ongoing scientific debate about the effectiveness of punishment, research has entered a fourth phase that I call the politicization of punishment phase. This phase overlaps temporally with the process phase but has a different agenda. In part because of arguments that the use of physical punishment may lead to physical abuse (Parke & Collmer, 1975), some commentators have argued that punishment should be dropped from the parental disciplinary repertoire (e.g., Straus, 1994). One of the major services provided by Gershoff’s (2002) analyses is her call for a clear distinction between the known effects of punishment—both in terms of its inhibitory value as well as its negative side effects—and the parental values concerning the appropriateness of punishment as a socialization tool. Science is never independent of values in spite of best efforts to treat it as exclusive of such issues. Decisions about disciplinary tactics are based not only on effectiveness but on parental views about the ethics and humanity of using physical punishment as an approach. A better understanding of changing views of children in terms of their decision-making abilities, their rights to a certain level of treatment, their sense of child agency—and of parents’ views of the extent to which the socialization process is a negotiated one will lead to greater insight into how punishment and other tactics are chosen and how the appropriateness of these tactics are evaluated by parents.

The issue of values concerns not just parents’ views of the appropriateness of punishment but the value perspectives of a variety of other people and institutions. First, are the same attitudes toward punishment shared by all disciplinary partners who are part of the socialization team (e.g., mother vs. father, extended family members such as grandparents, child care workers)? Second, what are the community attitudes toward the appropriateness of physical punishment? As Coleman’s (1988) concept of social capital notes, the extent to which community members share similar attitudes and values increases the effectiveness of parental practices. Similarly, the extent to which school and family share similar values will determine punishment effectiveness. Cultural factors need consideration as well. In cultures in which punishment is accepted and its practice is routinized, the impact of punishment will be different than in cultures in which punishment is not sanctioned. Historical period matters, too. When punishment was commonly accepted and used instrumentally rather than emotionally, the effectiveness of punishment probably was higher, the side effects lower, and the links with child abuse less clear (Parke & Collmer, 1975). Similarly, the child’s reaction to punishment is probably different in an era when punishment is normative. Parental reactions to punishment use by a partner or other socializing agent is also likely to vary as a function of the normativeness of the practice.

New Directions in Punishment Research

The Need for a Family-Systems Perspective

As Gershoff (2002) recognized, parental use of punishment needs to be placed in a family-systems perspective (Parke & Buriel, 1998). This perspective suggests that various levels of analysis require more attention, including dyadic, triadic, and family levels of analyses. At the dyadic level, the parent–child relationship requires more attention in recognition that parental socialization patterns are a negotiated dyadic process. The parent–parent relationship needs closer scrutiny as well. In fact, one of the determinants of punishment effectiveness is interpersonal consistency, and punishment effectiveness is lessened when there is disagreement between parents (Parke, 1977). The degree of conflict within the parent–parent relationship is probably a determinant of the degree of punitiveness directed toward the child, in part as a result of displaced anger or a shift in the threshold for responsiveness to child misbehavior (Cummings & Davies, 1994; Fincham, 1998).

At the triadic level, little is known about the impact of witnessed punishment. Not all children in the family are likely to elicit the same degree of punishment either because of differences in rates of misbehavior, temperamental and self-regulatory variations, or varying levels of skill in avoiding punishment (Grusec, Dix, & Mills, 1982; Kochanska, 1997; Parke, 1977). What impact does witnessed punishment of a sibling have on children’s outcomes? Little is known about this issue, nor do researchers have an understanding of children’s perceptions of fairness about the level of and frequency of punishment they receive relative to a sibling. Earlier work by Daniels and Plomin (1985) concerning siblings’ perception of differential treatment could serve as a guide for addressing this question. Finally, family-level variables need to be examined as well. Children in families with a clearly articulated philosophy concerning the appropriateness or inappropriateness of punishment are likely to respond differently either to being punished or to witnessed punishment. Similarly, the level of cohesive-
Hayden et al., 1998) but rarely have been applied to this issue. Finally, the interplay across the levels needs to be considered as well. Do parents who are inconsistent also have the highest levels of interparental conflict and lowest levels of family cohesion?

Beyond Variable-Oriented Approaches: A Family-Typology-Oriented Approach

Just as Magnusson and Stattin (1998) argued to capture the reality of individual lives, a person-oriented approach is preferred to the variable-oriented approach, which has been common in psychology in general and in the study of punishment, in particular. Several recent studies of parenting have used a person-oriented approach (e.g., Jain, Belsky, & Crnic, 1996; Magnusson & Stattin, 1998) and could provide a model for a new phase of punishment research. In this regard, it is time to revisit and update Baumrind’s (1967) classic studies of parental child-rearing typologies, now informed by new knowledge of processes that can be used to characterize child-rearing practices in terms of emotional and cognitive as well as social processes (e.g., Gottman, Katz, & Hooven, 1997). Moreover, this typological or person-oriented approach can be applied to not just mothers, but fathers, dyads, and family units to describe typologies of discipline at various levels of analysis.

Toward a Transactional Model of Parental Discipline

As Kuczynski (in press) has argued, as a field, psychology has moved beyond the question of parent versus child effects and recognizes the mutual and transactional nature of the parent–child relationship. Gershoff’s (2002) review provided persuasive evidence that parent and child characteristics are important determinants of both the utilization of punishment and the subsequent effectiveness of punishment. Transactional models (Sameroff, 1975) are best suited to capture the dynamic nature of punishment effects. Implicit in a transactional approach is the importance of recognizing and measuring developmental changes in children’s reactions to incidents likely to yield empathetic responses. A focus on the parent’s own mood and emotional state as well as the child’s reaction to parental discipline would move research forward (Dix, 1991). Another approach merits more use, namely the use of longer term observations of families. As Gottman (1994) has shown, the observation of couples in a laboratory for a period of days has yielded important insights into the daily tempo of marital interactions. Similarly, Yarrow et al. (1973) effectively used extended lab visits in which mothers and children were observed when confronted with a range of normal challenges such as meal preparations, distracting phone calls, and clean-up tasks. More use of these kinds of approaches will greatly aid efforts to isolate not only the real-life elicitors but the relative short-term effectiveness of punishment and related disciplinary practices as well. No single method has ever been sufficient; it is not likely to be so in this case either.

Clearly, experimental approaches have a central role to play in studies of socialization, but there is also a need to recognize that hybrid experiential designs are available, which in part overcome some of the problems associated with lab-based experimental studies (Parke, 1979). Manipulations in the lab can be supplemented by measurement of the effects of punishment in more naturalistic contexts. Alternatively, children who are punished in more naturalistic contexts (e.g., home) can be assessed in the lab to determine the effectiveness of the punishment manipulation. Finally, field experimental approaches in which discipline is administered in naturalistic settings and measured in the same or other naturalistic settings can be used.

The new generation of experimental work (e.g., Cummings & Davies, 1994; Kochanska, 1997) is providing tests of the mediating and moderating processes that have been proposed to account for the differential effectiveness of punishment. For example, the advances in the understanding of emotional development, including emotional understanding, emotional expressiveness, and especially emotional regulation, (e.g., Denham, 1998) can inform this new generation of studies. Basic assumptions concerning the importance of emotional arousal for punishment effectiveness can be evaluated by the use of psychophysiological techniques and recent advances in brain scanning techniques, such as magnetic resonance imaging. Similarly, the genetics of punishment is poorly understood, although Kochanska’s (1997) studies of temperament and punitiveness are a step toward a clearer understanding. Twin studies, adoption studies, and within-family sibling designs are all approaches that merit more use in future studies.

Beyond Quantitative Approaches: The Need for Culturally Sensitive Approaches

Just as questionnaires, beepers, and experiments are not sufficient by themselves, it is unlikely that quantitative approaches alone are going to be adequate, especially if one takes seriously Gershoff’s (2002) call for a more ethnically and racially representative picture of the use and impact of punishment. To assume that the same attitudes toward the act of punishment, the elicitors of punishment, and the consequences of punishment are similar across cultural and ethnically diverse groups is not warranted. Two steps are needed. First, better descriptive data are needed not only
on the use of punishment by parents of different racial and ethnic groups but on their attitudes toward punishment. National surveys in which a range of ethnic groups are adequately represented is a first step. This approach, however, will provide little insight into processes of punishment or the use of punishment in the lives of different groups. Instead, other approaches are needed that include ethnographic methods. Burton and Price-Spratlen (1999) have provided richly textured data concerning the flow of life in African American families; Miller and Sperry’s (1987) ethnographic accounts of anger socialization in lower class African American families are another example. Focus groups offer another promising strategy to help refine the questions and generate hypotheses that may be unique to different ethnic groups. These approaches are especially critical as research in this area moves into relatively uncharted ethnic territory (Parke et al., in press). Without the use of these approaches, the study of punishment is likely to apply paradigms developed on majority samples without adequate understanding of the cultural appropriateness of these strategies.

Conclusion

In sum, Gershoff (2002) has provided a synthesis that will serve as a useful guide to research efforts in this area of long-standing interest. At the same time, recognition that punishment is best viewed as a packaged variable that needs to be located in the matrix of other socialization practices will improve understanding by refocusing the framing of the question. Finally, more attention to the interplay between science and values will be critical for understanding not only the effects of punishment but for understanding when and why parents choose to use punishment as a socialization tactic.

References


Parke, R. D., & Walters, R. H. (1967). Some factors determining the


Received November 14, 2001
Accepted November 16, 2001

Low Publication Prices for APA Members and Affiliates

Keeping you up-to-date. All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the American Psychologist and APA Monitor. High School Teacher and International Affiliates receive subscriptions to the APA Monitor, and they may subscribe to the American Psychologist at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

Essential resources. APA members and affiliates receive special rates for purchases of APA books, including the Publication Manual of the American Psychological Association, and on dozens of new topical books each year.

Other benefits of membership. Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

More information. Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.