Children and Parents Deserve Better Parental Discipline Research: Critiquing the Evidence for Exclusively “Positive” Parenting

Robert E. Larzelere, Marjorie Lindner Gunnoe, Mark W. Roberts & Christopher J. Ferguson


To link to this article: https://doi.org/10.1080/01494929.2016.1145613

Published with license by Taylor & Francis © 2017 Robert E. Larzelere, Marjorie Lindner Gunnoe, Mark W. Roberts, and Christopher J. Ferguson

Accepted author version posted online: 26 Jan 2016.
Published online: 11 Mar 2016.

Article views: 1929

View related articles

View Crossmark data

Citing articles: 4 View citing articles
Children and Parents Deserve Better Parental Discipline Research: Critiquing the Evidence for Exclusively “Positive” Parenting

Robert E. Larzelere, Marjorie Lindner Gunnoe, Mark W. Roberts, and Christopher J. Ferguson

ABSTRACT
This article critiques the scientific evidence for the emerging view in nonclinical parenting research and in popular books that parents should use only positive methods of parenting and rarely resort to any disciplinary consequences. Four methodological fallacies pervade research used to support this viewpoint: the correlational fallacy (inferring causation from correlations), the trumping fallacy (permitting correlational conclusions to trump stronger causal evidence), the extrapolation fallacy (extrapolating favorable comparisons of under-usage versus over-usage to zero usage), and the lumping fallacy (lumping inappropriate and appropriate usages together). Conclusions based on any of these methodological fallacies are premature at best and counterproductive at worst. These fallacies would incorrectly make many medical procedures appear to be harmful, such as radiation treatment. Premature conclusions supporting exclusively positive parenting may partially explain the immigrant paradox in the United States and escalating criminal assaults against minors according to Swedish criminal records (where positive parenting is most prominently advocated). Exclusively positive parenting needs to be supported by stronger research, including randomized trials with oppositional defiant children, before being accepted as definitive. We also need research to understand how the parental management skills featured in effective clinical treatments for young oppositional defiant children generalize to parenting in nonclinical families.

KEYWORDS
child discipline; corporal punishment; parenthood/parenting; statistical methods

Introduction

Positive parenting is the philosophy that parental attempts to influence their children should be limited to warm and supportive guidance. Popular psychologist Laura Markham describes it like this: “Positive parenting—sometimes
called positive discipline, gentle guidance, or loving guidance—is simply guidance that keeps our kids on the right path, offered in a positive way that resists any temptation to be punitive” (Markham, 2015, 3rd para.; 2012). Although we agree that loving guidance is an important element of good parenting, we are concerned by the recent spate of absolute or near-absolute statements in both popular and professional publications opposing any form of disciplinary consequences, including timeout. The assertion that children can never benefit from appropriately applied punitive correction disregards empirical findings that have been foundational to both developmental and clinical child psychology. With respect to developmental psychology, the authoritative parenting style delineated by Baumrind (e.g., Baumrind, 2012, 2013; Baumrind, Larzelere, & Owens, 2010) is widely accepted as the most effective parenting style (Parke & Buriel, 2006; Steinberg, 2001). It combines the positive dimensions of nurturance and give-and-take communication with maturity demands and firm discipline when needed (Baumrind, 2012; Baumrind et al., 2010). As for clinical psychology, all parent-implemented treatments for oppositional defiant disorder and attention deficit hyperactivity disorder (ADHD) in young children that are empirically supported according to the Society of Clinical Child and Adolescent Psychology incorporate punitive measures in the form of timeout and enforcements for cooperation with timeout (Eyberg, Nelson, & Boggs, 2008; Pelham & Fabiano, 2008).

Many advocates of positive parenting seem opposed to any disciplinary consequences. A recent popular book by Siegel and Bryson (2014a) stated that most uses of timeout and intentional ignoring, as well as spanking, are detrimental to the parent–child relationship and therefore harmful to the child. Those authors emphasized their opposition to timeout in two national publications (Siegel & Bryson, 2014b, 2014d), although they later claimed that one of those editorials exaggerated their opposition to it (Siegel & Bryson, 2014c), in response to a letter from the Society of Clinical Child and Adolescent Psychology (2014) to Time magazine.

Many child development scholars also seem reluctant to recommend any disciplinary consequences to parents. In 2011, George Holden led a conference entitled the Global Summit on Ending Corporal Punishment and Promoting Positive Discipline, which featured exclusively positive parenting as the only specified alternative to spanking, citing Durrant (2007). Holden then sponsored a spanking-ban resolution which also featured positive parenting as the only recommended alternative to spanking. This resolution was adopted by the Society for Research in Human Development in 2013 by a membership vote of 15 to 6.

In this essay we argue that these absolute or near-absolute proscriptions of all disciplinary consequences, including timeout and privilege removal, are scientifically premature because of four methodological fallacies that are pervasive in the parenting research used to support all-positive parenting. The methodological fallacies are (1) basing causal conclusions on correlations, (2) ignoring causal
Four methodological fallacies that make absolute conclusions premature

**Correlational fallacy**

The first methodological fallacy is arguably the most well-known tenet of inferential statistics: correlations cannot prove causation. Correlations can be particularly misleading when used to evaluate corrective actions, that is, actions intended to correct a perceived problem (Larzelere & Cox, 2013; Larzelere, Cox, & Swindle, 2015). For example, correlations would make radiation treatment look harmful for cancer patients, because those receiving radiation last year would have more cancer-related symptoms now than those who did not need that treatment (Larzelere & Baumrind, 2010).

Unfortunately, some advocates of positive parenting still place a great deal of emphasis on correlational evidence. The most-cited literature review documenting the dangers of physical discipline is based entirely on unadjusted correlational data (Gershoff, 2002). Unadjusted correlations also make timeout (Gershoff et al., 2010), sending children to their room, and privilege removal appear to have harmful outcomes (Larzelere, Cox, & Smith, 2010). This is not surprising given that longitudinal research has yet to find a corrective action that is correlated with reduced levels of the symptoms it is trying to correct (Larzelere & Cox, 2013).

Such adverse associations are found even in longitudinal studies that control for pre-existing child problems, but the effect sizes are usually tiny. For example, Ferguson’s (2013) meta-analysis reported effect sizes for spanking that explained only one-half of 1% of the variance in externalizing problems still unaccounted for in statistically controlled longitudinal studies. Such tiny effects disappear once researchers remove measurement error (Larzelere et al., 2010a) or control for an additional confound, such as overly frequent spanking (Lansford, Wager, Bates, Pettit, & Dodge, 2012). Moreover, statistically controlled longitudinal studies produce similar small adverse effect sizes for all corrective actions for oppositional defiant disorder, whether by parents...
or professionals. For example, Ritalin users turned out to have more ADHD symptoms 2 years later than nonusers, even after controlling for their initial ADHD symptoms (Larzelere, Ferrer, Kuhn, & Danelia, 2010). Controlling statistically for initial ADHD symptoms reduces the selection bias due to child effects, but it fails to eliminate it entirely because covariate measures are imperfect (Larzelere & Cox, 2013). The remaining residual bias explains why psychotherapy and Ritalin look just as harmful as spanking and nonphysical punishments in such studies (Larzelere et al., 2010b).

A third situation is when tiny effect sizes do accurately represent an unbiased estimate of an average causal effect. Even then an absolute recommendation based on a tiny effect size could easily be the wrong recommendation for a large portion of the sample. To illustrate this, we have created a hypothetical scatterplot corresponding to the tiny effect size ($\beta = .07$) obtained by Ferguson (2013) in his meta-analysis of the average longitudinal association between spanking and externalizing behavior problems in children (Figure 1). If this effect size were an unbiased estimate of an average causal effect, about 56% of children who experienced above-average spanking would become more aggressive than otherwise predicted, but 44% of children who experienced above-average spanking would reduce their aggression more than otherwise expected. These results suggest the need to move away from absolutist proscriptions against traditional disciplinary consequences to redirect researchers’ efforts to discriminate between more versus less effective ways of using each corrective disciplinary action, including spanking and potential replacements for it.

![Figure 1. Scatterplot of a hypothetical adverse outcome regressed upon use of a corrective action, illustrating $r = \beta = .07$ (mean $\beta$, Ferguson, 2013).](image-url)
Trumping fallacy: Ignoring stronger causal evidence

The second methodological fallacy is to ignore published experimental evidence of the causal effects of discipline strategies to improve child outcomes. Although stronger causal evidence would never be ignored in the medical field, it is routinely ignored by those who are philosophically opposed to power assertion, which has been defined by Shaffer and Kipp (2007) as “the use of superior power to control the child’s behavior (including techniques such as forceful commands, physical restraint, spanking, and withdrawal of privileges)” (p. 585).

The most important causal evidence in parenting research comes from randomized clinical trials of intervention strategies for oppositional defiant disorder and other disruptive behavior diagnoses in preadolescent children (see Eyberg et al., 2008; Pelham & Fabiano, 2008). These studies have demonstrated that noncompliance with parental instructions (the defining feature of oppositional defiant disorder) can be effectively reduced by teaching parents to use direct instructions, single warnings, chair timeouts, and timeout enforcement procedures (cf. McMahon, Wells, & Kotler, 2006, pp. 161–172). Each component of this compliance training sequence has been validated through experimental manipulation and observational data of children’s reactions in controlled clinic settings (cf. Roberts, 2008). Specifically, randomized trials have shown that both brief chair timeouts and enforcements for these timeouts are necessary for the program to be effective (Bean & Roberts, 1981; Fee, Matson, & Manikan, 1990; Olson & Roberts, 1987; Roberts, Hatzenbuehler, & Bean, 1981). In interventions with overtly noncompliant, clinic-referred preschool children, enforcement for chair timeouts was accomplished equally well by a brief room isolation or the traditional two-swat spanking (Bean & Roberts, 1981; Day & Roberts, 1983; Roberts, 1988). Some children cooperated more quickly with isolation and others more quickly with spanking, and continuing defiance was overcome by changing to the other enforcement. Follow-up data from the home indicated that the need for enforcements for timeout and for timeouts themselves was reduced to near zero levels within 4 weeks of consistent use for most children (80% of children in Roberts, 1985, Project 2; Roberts & Powers, 1990). Consistent implementation of timeout contingent on defiance to parental warnings in the home for a 2-month period virtually eliminated the need for its own use.

In short, effective clinical treatments train parents how to use timeout and other disciplinary responses skillfully and consistently, which results in rapid decreases in the frequency with which they need to be used, thereby accounting for the correlational superiority of low or even zero use of negative disciplinary consequences after skillful consistent usage (Roberts & Powers, 1990). This produces cooperative children whose parents rarely need to use negative disciplinary consequences, a goal shared by all perspectives on parental discipline. But, at least for clinically defiant children, that goal is achieved by skillful, consistent use of forceful tactics opposed by positive parenting. Positive
parenting advocates appear to dismiss this strong experimental evidence, possibly on the grounds that such data apply only to clinic-referred problem children or a simple lack of awareness of the relevant published literature.

**Extrapolation fallacy: From low to zero usage**

The third methodological fallacy is to overgeneralize (to zero usage) the typical linear associations that favor low usage over high usage of corrective disciplinary actions. This is analogous to completely prohibiting medical treatments (e.g., pharmaceuticals, radiation) because they cause harm at high dosages rather than striving to identify an optimal intermediate dosage. Part of the problem is that the usual linear statistics contrast under-usage versus over-usage of a corrective disciplinary action and thus cannot detect the possibility that an intermediate level of usage might be optimal. Barber and Xia (2013) suggested that this failure to test intermediate usage of behavioral control has hindered cumulative progress in understanding how parents can use behavioral control effectively. This failure gets exacerbated when linear statistical associations are extrapolated from low usage to zero usage, even when zero lifetime usage is rarely measured. To take the example of physical discipline, we know of only five studies that isolated a never-spanked group of children, and even these studies yielded mixed results. Three found that the outcomes of spanked children were never worse and sometimes better than never-spanked children, as long as the spanking was occasional (Power & Chapieski, 1986) or did not continue past 8 or 11 years of age (Ellison, Musick, & Holden, 2011; Gunnoe, 2013). The fourth study indicated that the never-spanked group was associated with fewer externalizing behavior problems concurrently (Straus & Mouradian, 1998), and the fifth study reported longitudinal evidence that the 4% of children who were never spanked were less aggressive on the kindergarten playground (Strassberg, Dodge, Pettit, & Bates, 1994). Of these five studies, only Ellison et al. (2011) controlled for pre-existing differences on the outcome variable. Moreover, none of these studies compared their results with other forms of discipline that might have been used by the parents of never-spanked children, such as timeouts, privilege losses, or reprimands.

**Lumping fallacy: Failure to make important discriminations**

The fourth methodological fallacy is the failure to make necessary discriminations based on how a disciplinary consequence is used and the situations in which it is used. It is imperative that the evaluation of disciplinary consequences be based on precise operational definitions of these actions. Good medical and psychotherapy research specifies the precise way that a treatment is conducted. In contrast, most research on disciplinary consequences lumps a wide range of disciplinary actions together. For example, Baumrind, Larzelere,
and Cowan (2002) reported that 65% of the studies predicting antisocial behavior from physical punishment in Gershoff’s (2002) meta-analysis failed to discriminate customary spanking from overly harsh physical techniques, some of which included beatings with a whip, belt, or stick. More precise definitions are essential for scientific advances. With respect to discriminating the situations in which specific tactics are used, the few correlational studies that have specified the disciplinary situation at all have relied on vignettes of disciplinary episodes. Unfortunately, most vignettes focus on somewhat ambiguous misbehavior (e.g., peer conflicts) rather than the kind of oppositional defiance for which negative consequences such as timeout have been an effective option (e.g., Aronfreed, 1961; Deater-Deckard, Dodge, Bates, & Pettit, 1996). When disciplinary situations have specified “extreme” or dangerous misbehavior, privilege removal and spanking were associated with significantly less aggression 2 months later at preschool than were five other disciplinary tactics, including reasoning (Yarrow, Campbell, & Burton, 1968; effect sizes in Larzelere & Baumrind, 2010).

Possible consequences of methodological fallacies in research

If medical treatments such as radiation therapy were opposed based on the methodological fallacies we have described, more cancer patients would die, especially if the same flaws prevented other cancer treatments from being recognized as effective. Could premature conclusions against evidence-based disciplinary consequences have similar iatrogenic effects? This is one possible explanation for two areas of current concern to social scientists.

The first concern is the well-documented immigrant paradox. The paradox is that newly immigrated youth have more optimal developmental outcomes than do U.S.-born youth, despite their socioeconomic and language disadvantages (Marks, Ejesi, & García Coll, 2014). First-generation American immigrants are 46% less likely to commit antisocial crimes against persons than are other Americans, but they catch up by increasing their likelihood of committing violent crimes by 1.9% for each year in the United States (Vaughn, Salas-Wright, DeLisi, & Maynard, 2014). The immigrant paradox is larger for cultures considered more authoritarian in their parenting (Africa, Asia) than for immigrants from more permissive cultures (Europe). Although many factors influence the acculturation process, part of the explanation may be that American parenting advice through the media is inadvertently undermining parenting strengths in families from non-Western cultures. That is, Euro-American parenting advice may unintentionally be promoting a dysfunctional version of permissive parenting in its well-intentioned opposition to overly authoritarian parenting.

The second concern is the steep increase in criminal assaults by youth in one Scandinavian country that has implemented positive parenting most
vigorously (Bussmann, Erthal, & Schroth, 2012) and the emergence of a dysfunctional type of overly permissive parenting in a neighboring country. Since Sweden banned spanking in 1979, physical abuse of the youngest children and assaults by minors against minors have both increased more than 20-fold, according to criminal records (Larzelere, Swindle, & Johnson, 2013). Reports of children attacking their parents have also increased (Haeuser, 1988). The only other country known to enforce spanking bans as vigorously is Norway, where many clinically referred families displayed a “permissive parenting form of child coercion” where “parents seem simply unable to say no” to their children (Patterson & Fisher, 2002, p. 74). Patterson, an eminent parenting researcher (cf. Patterson, 1982), based these observations on his colleagues’ experiences in training Norwegian therapists to implement their empirically supported Parent Management Training–Oregon Model (Eyberg et al., 2008).

Is there a connection between spanking bans, clinical levels of ineffective parenting, and escalating rates of assaults on and by Scandinavian children? There is some evidence that other disciplinary consequences have fallen into disfavor in Sweden in recent decades in addition to spanking. Janson (2001, Table 13) reported that only 4% of Swedish children thought parents had the right to “threaten or forbid something” in 2000 compared with 39% in 1994 and 1995, with less dramatic decreases in support for grounding or taking away pocket money. Although we suspect that many factors have contributed to the increase in child-related assaults in Scandinavia, we consider it plausible that some of the increase in assaults may be attributable to parents’ increasing reluctance to use any disciplinary consequences, which undermines the use of parental disciplinary skills such as timeout that are especially effective for oppositional defiant children (Marian Forgatch, personal communication, April 18, 2007).

**Conclusions**

We do not question the good intentions of those who advocate a version of positive parenting that excludes all disciplinary consequences. Nor do we question the importance of parents maintaining a positive relationship with their children as much as possible. High levels of support, reasoning, and other specific behaviors encouraged by the advocates of positive parenting are likely important for preventing the emergence of oppositional defiant behavior that leads parents to use punishments and seek help from psychotherapists.

What we are questioning is the putative scientific basis for an overgeneralized opposition to all disciplinary consequences that include any element of aver siveness or power assertion (e.g., timeouts, token fines, privilege losses, physical guidance). An all-positive approach might work well with children who have easy temperaments, but it contradicts the fact that, in addition to reinforcing appropriate behaviors, *all empirically supported parenting interventions for*
oppositional defiant children intentionally train parents in power assertive skills that many positive-parenting advocates oppose. These clinical protocols put a priority on training parents in the mildest disciplinary responses that will be effective. The clinical studies we have already described specified precise ways to use forceful skills to bolster cooperation with milder disciplinary responses. They also demonstrated that the use of forceful skills can be phased out quickly because children learn to cooperate with milder steps in the systematic parental discipline protocol (i.e., instructions and warnings).

The positive parenting movement is a philosophical movement. To attain legitimate scientific credibility, advocates need to demonstrate its effectiveness in research designs that avoid the four methodological fallacies we have explicated. Such research should demonstrate that exclusively positive parenting decreases defiance and fosters the positive outcomes emphasized by its advocates, even in young children with oppositional defiance. Until such research is present, we urge the scientific community to resist absolute or near-absolute prohibitions against the use of disciplinary consequences. Premature acceptance of these prohibitions not only removes from the parental “toolbox” techniques that have been proven effective with some of society’s most at-risk children, it also undermines the capability of parenting research to identify alternative disciplinary tactics that could effectively replace spanking in disciplinary situations where spanking has been a traditional option. For its part, clinical child research needs to explain how their parent-training protocols can facilitate the kind of positive relationships and communication between parents and children that gives all-positive parenting its appeal. Both areas of parenting research need to move beyond the methodological fallacies highlighted in this article if they are going to meet their goal of helping parents find the least punitive but sufficiently effective techniques to maximize children’s potential.

References


Holden, G. W. (2011, June 2–4). Global summit on ending corporal punishment and promoting positive discipline. Conference held in Dallas, TX.


