Toddlers Need Both Positive Parenting and Consistent Consequences from Mothers

Robert E. Larzelere and Sada J. Knowles
Oklahoma State University

Presentation at the Annual Convention of the American Psychological Association, August 6, 2015, Toronto, ON, Canada

Summary for the Media

Parental discipline advice is often polarized in popular parenting resources and in parenting research conclusions. Should parents respond positively to children all of the time, as some experts say? Or should they use firm discipline when necessary, consistent with other experts? This study investigates the immediate and longer-term effects of seven common disciplinary responses to toddler misbehavior, including how their relative effectiveness varies by the type of noncompliance. Mothers varied their disciplinary tactics with the type of toddler noncompliance in a pattern that was usually more effective than following any one parenting expert exclusively.

A voluntary sample of 102 mothers provided detailed descriptions of 5 discipline episodes with their toddlers. Only 8 of them completely avoided the disciplinary responses prohibited by exclusively positive parenting proponents (n = 4) or those prohibited by behavioral parent training (n = 5, with one mother avoiding both sets of prohibited responses).

Offering compromises was the most effective tactic for immediate reductions in noncompliance intensity regardless of the type of noncompliance. The immediate effectiveness of the other disciplinary tactics varied with the noncompliance type. Reasoning was the next most effective response when mothers were responding to mild noncompliance (negotiating, whining), whereas broad power assertive responses (including punishments) were the second most effective response to oppositional noncompliance (defiance, hitting, passive noncompliance). In direct contrast, broad power assertion was the least effective response to mild noncompliance, and reasoning was the least effective response to oppositional noncompliance, with other three disciplinary response categories intermediate in their immediate effectiveness.

Longer term effects revealed a different pattern. The two-month effects of offering compromises varied with the predominant type of toddler noncompliance: Frequent use of offering alternatives made all behavioral problems worse for the most oppositional toddlers but improved those outcomes for negotiating/whining toddlers. (The outcomes included externalizing, internalizing, and total behavior problems.) In contrast, the 2- and 16-month effectiveness of reasoning varied by noncompliance type: Frequent reasoning was more effective at reducing problems for oppositional than for negotiating toddlers. Intermediate use of punishments and warnings (on less than 16% of turns) was effective only for oppositional toddlers: Compared to zero usage, intermediate usage predicted significantly greater reductions in behavior problems over the next two months in oppositional, whereas it had the opposite effect for negotiating toddlers. This benefit for oppositional toddlers did not occur for over-usage of punishments/warnings, which might explain why previous longitudinal analyses have failed to detect the effectiveness of skilled use of punishments and warnings.

The major findings are that offering compromises to toddlers produce the greatest immediate decrease in the severity of noncompliance and has adverse longer-term effects only when used too often with the most oppositional toddlers. For oppositional toddlers, reasoning is effective long-term even though it is the least effective response immediately. The kinds of single warnings and punishments featured in behavioral parent training are effective both immediately and long-term, but only for the most oppositional toddlers and only when used less than 1/6 of the time.

Toddlers Need Both Positive Parenting and Consistent Consequences from Mothers: Immediate, Short-Term, and Long-Term Effectiveness of Disciplinary Tactics by Type of Toddler Noncompliance

Robert E. Larzelere
Sada J. Knowles
Oklahoma State University

Paper Presented at the Annual Convention of the American Psychological Association, August 6, 2015, Toronto, ON, Canada.

robert.larzelere@okstate.edu (405) 744-2053 (office) (405) 564-3616 or -4497 (cell phones)

Abstract

To reconcile the contradictory disciplinary recommendations of positive parenting and behavioral parent training, this study used a convenience sample of 102 mothers and toddlers to investigate the immediate, short-term, and long-term effectiveness of seven disciplinary responses to toddler noncompliance, with unusually thorough methods to approximate unbiased causal evidence. Offering alternatives was the most effective response to noncompliance in reducing noncompliance severity immediately and it had adverse longer term effects only when used too often with the most oppositional toddlers. For immediate effectiveness, reasoning was the second most effective response for parent-oriented noncompliance (negotiating and whining) and power assertion was least effective, but the relative effectiveness of those two responses reversed when responding to oppositional noncompliance (defiance and hitting). To reduce externalizing and total behavior problems during the next two months, the results differed for oppositional vs. negotiating toddlers. Offering alternatives frequently led to reduced externalizing problems in toddlers who usually negotiated and whined, whereas warnings and punishments increased their total behavior problems. Those disciplinary responses had the opposite effects for toddlers who were usually oppositional. For them offering alternatives more than 29% of the time increased their externalizing problems, whereas moderate use of warnings and punishments led to decreases in total behavior problems. Frequent use of reasoning was also effective in reducing subsequent behavior problems for oppositional toddlers, even though it was the least effective disciplinary response in immediately reducing the severity of oppositional noncompliance. Parents need the full range of nonabusive disciplinary responses, but should match them to the type of noncompliance exhibited by toddlers.

Toddlers Need Both Positive Parenting and Consistent Consequences from Mothers: Immediate, Short-Term, and Long-Term Effectiveness of Disciplinary Tactics by Type of Toddler Noncompliance

"Trust yourself. You know more than you think you do." Thus began Spock's (1946) revolutionary book *The Common Sense Book of Baby and Child Care*. Before then the leading book on how to rear children was "dedicated to the first mother who brings up a happy child" (Watson, 1928, p. XXX). It claimed that the new science of childrearing would inevitably lead to a much superior way of raising children than the amateurish efforts of mothers, who Watson considered far too nurturing, not requiring enough of their children: "Never hug and kiss them . . . Give them a pat on the head if they have made an extraordinarily good job of a difficult task. Try it out. . . . You will be utterly ashamed of the mawkish, sentimental way you have been handling it." (pp. 81-82)

It is easy to see why Spock's book provided a breath of fresh air and went on to become the best-selling new title in the United State between 1900 and 1968 (Spock, 1968). Is current parenting advice better than Watson's or Spock's advice in any way? Has it fulfilled the promise of the science of childrearing? Or should parents trust themselves more than parenting experts?

Spock started a huge pendulum swing from the authoritarian views represented by Watson (1928) to the mixed messages of today (Rex Forehand & McKinney, 1993; Hulbert, 2003), although there have always been some "soft" experts emphasizing positive parenting and encouraging children's autonomy and some "hard" experts emphasizing firm discipline and character development (Hulbert, 2003). This study reports on a brief window on how mothers handle discipline espisodes involving with toddlers to see (1) how many mothers restrict their disciplinary responses to those recommended by current parenting experts, and (2) whether their effectiveness depends on their adherance to the advice of current parenting experts.

Those questions cannot be answered without choosing parenting experts, whose advice differs widely both among popular parenting experts and among parenting researchers. This article focuses on two contrasting viewpoints, exclusively positive parenting (Durrant, 2007; Markham, 2012; Siegel & Bryson, 2014) and behavioral parent training (e.g., Barkley, 2013; R. Forehand & Long, 2010; McNeil & Hembree-Kigin, 2010; Webster-Stratton, 1992). Although parenting experts in development science rarely oppose all power assertion explicitly (but see Kochanska, Aksan, & Joy, 2007, p. 233), many of them never recommend any type of power assertion, even when broadly defined 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)" (Shaffer & Kipp, 2007, p. 585). The implicit opposition to all power assertion is more frequently stated as support for positive parenting only, which was the only alternative recommended to replace spanking in a spanking-ban resolution passed by the Society for Research in Human Development (2013). In contrast, behavioral parenting training trains parents to respond to noncompliance with consistent use of mild power assertive tactics, an approach that remains the only parent-implemented treatment that has met scientific standards for treating behavior disorders in young children, such as oppositional defiant disorder and ADHD (Eyberg, Nelson, & Boggs, 2008; Pelham & Fabiano, 2008).

The contrast between these two empirically supported perspectives is striking. The positive parenting perspective opposes any negative consequences imposed by parents, such as time out and privilege removal as well as physical punishment. Based on her observational measures, Kochanska's opposition to power assertion includes all prohibitions as well as "any sense of a clash of will, however subtle" (Kochanska et al., 2007, p. 225). Instead positive parenting advocates recommend gentle disciplinary responses, such as explaining, empathizing, finding mutually acceptable compromises, and preventive actions (e.g., Durrant, 2007). Such exclusively positive protocols have never been shown to be effective in clinical treatments for children with oppositional defiance and other disruptive behavior diagnoses. The empirically supported parent-implemented treatments for those disorders do not train parents to respond to noncompliance with any of the disciplinary responses recommended by the exclusively positive parenting advocates, but instead train parents to respond to noncompliance with a consistent sequence of firm commands, single warnings, time out, and enforcements for time out, all of which are examples of power assertion, broadly defined (Durrant, 2007; Kochanska et al., 2007). These contrasting empirically supported perspectives change our research questions to (1) Are mothers of toddlers consistent with either empirically supported perspective, either exclusively positive parenting or behavioral parent training? (2) Are their discipline practices more effective if they conform to one or the other of these contrasting perspectives? (3) Does the effectiveness of the disciplinary responses recommended by these contrasting perspectives vary by the type of noncompliance in the child? The fact that behavioral parent training has been shown to be effective with clinically referred children, whereas evidence for the exclusively positive parenting perspective comes from nonclinical samples suggests the following: (3a) Do mothers of toddlers use the tactics of behavioral parent training when children are defiant, but use the gentler positive tactics when toddlers show milder types of noncompliance? (3b) Are mothers more effective to the extent they vary their disciplinary tactics to match the type of noncompliance shown by their toddlers?

The latter hypothesis may help correct a possible weakness in both perspectives. They seem to promote a one-size-fits-all perspective, without considering when their recommended disciplinary tactics might be ineffective or counterproductive. In contrast, mothers modify their disciplinary tactics to fit their overall assessment of the disciplinary situation, including the type of noncompliance exhibited at that time by the child (Ritchie, 1999).

To contrast these perspectives fairly, however, it is important to test both immediate and long-term effectiveness. Both perspectives emphasize the distinction between immediate and long-term effects. For example, leading behavior parent training experts argue that parents can inadvertently train children to be antisocial by giving in to child tantrums too often (Patterson, 1982; Roberts, 2008). Parents get immediate peace and quiet, but children learn to get what they want by becoming increasingly aversive. Positive parenting advocates make a similar distinction, especially about spanking, that it is effective in terms of immediate compliance, but has adverse long-term outcomes (e.g., Gershoff, 2002). Therefore this study investigate both immediate effects and longer term effects on behavior problems and effortful control.

There are important substantive and methodological issues related to these research questions. Substantively, some have held that optimal parenting incorporates both positive parenting skills and power assertive tactics. A leading example is authoritative parenting, which combines give-

and-take reasoning with firm control and has been associated with better child outcomes than alternative parenting styles that emphasize one more than the other (Baumrind, Larzelere, & Owens, 2010; Larzelere, Cox, & Mandara, 2013; Steinberg, 2001). Hoffman's (2000) theory of moral internalization holds that inductive reasoning is essential for moral internalization, but that power assertion must be sufficient to optimize the child's attention, without being so severe that it undermines cognitive processing. He also thought that the memory of power assertion, which he called background power assertion, can produce the optimal level of power assertion to maximize the effectiveness of the reasoning component of parental discipline.

Methodologically, behavioral parent training is supported by stronger causal evidence, whereas exclusively positive parenting is supported by correlational research on a broader range of nonclinical children. The correlational bias against all corrective actions makes all disciplinary responses appear to have more adverse effects than they actually do (Larzelere & Cox, 2013), even after controlling for initial differences on the outcome (Larzelere, Cox, & Swindle, 2015; Larzelere, Ferrer, Kuhn, & Danelia, 2010). This correlational bias could explain the apparently harmful long-term effects of traditional power assertive discipline, such as spanking. On the other hand, it is not clear how the parenting skills that are useful for reducing oppositional defiance in young children generalizes to other situations. Does oppositional defiance occur because parents have failed to use positive parenting methods consistently? Or are temperamentally difficult children kept from becoming oppositional and defiant by early consistent use of the skills emphasized in behavioral parent training? In either case, this study focuses on toddlers because that age is most relevant for early prevention of oppositional defiance and related conduct disorders.

Bell's control system model (Bell & Chapman, 1986; Bell & Harper, 1977) can account for both the correlational evidence against power assertive tactics and the fact that parents are more likely to resort to such tactics when initial positive tactics do not work well enough (Larzelere et al., 2013). For example, Larzelere et al. (2013) suggested that authoritative parents rely on positive methods as much as possible, but also resort to consistent use of power assertive methods when an acceptable level of cooperation cannot be achieved with positive methods, as illustrated in Figure 1. The current study controls for more than the usual demographics plus one initial score on the outcome variable to minimize the correlational bias against corrective actions, so that the results come closer to unbiased causal estimates of immediate and longer term outcomes.

The ultimate goal is to integrate the perspectives on parental discipline from developmental science and behavioral parent training. It is a disservice to psychological science and to parents for the two perspectives to continue ignoring each other, when they complement each other in so many ways.

Method

Participants

Participants included 105 mother-toddler pairs recruited through local newspaper ads, university contacts, and local organizations, such as Early Head Start, childcare centers, churches, and businesses. The mean age of mothers at the beginning of the study was 30.4 years, *SD* = 4.8.

Mothers were predominantly Caucasian (80%) and married (79%). Others identified themselves as Native American (7.6%), African American (4.8%), Hispanic (3.8%), Asian (1.9%), or Middle Eastern (1.9%). Additionally, 4.8% of mothers were separated, 2.9% cohabiting, and 13.3% single. The sample was well educated, with only 1.9% reporting less than a high school education, 4.8% high school only, 33.4% some further education (4.8% technical school, 7.6% associate degree, 21% some college), 30.5% a bachelor's degree, and 29.5% additional post-graduate education. The median family income was between $2,500 to $3,000 per month. The toddlers' ages ranged from 17.2 to 30.8 months, M = 23.8, SD = 3.9. The toddlers included 65 boys and 40 girls. Three toddlers had indications of very serious medical conditions or developmental disabilities and were dropped from the analyses, leaving 102 mother-toddler pairs in the final sample.

Procedure

In most cases, mothers brought their toddlers to the university observation center for the initial interview, although 7 interviews occurred in their homes. The mothers completed a demographic questionnaire, the Child Behavior Checklist 1½-5 (CBCL), and one other questionnaire before the interview. After some general questions about how they handled their toddler's misbehavior, they completed the Nurturing and Discipline Practices Questionnaire. Next they completed a modified version of the Child Conflict Index (Frankel & Weiner, 1990) to identify discipline problems during the past 24 hours. Mothers then described the day's most difficult discipline episode in specific detail, using an adaptation of Ritchie's (1999) protocol. Next they provided similar details about a second discipline episode "that could have been problematic, but you kept it from becoming more problematic." Mothers then completed a trauma questionnaire and 36 items from the Early Child Behavior Questionnaire (ECBQ). The interview session then ended with 5 minutes of play time and a 5-minute clean-up task. Research assistants then called the mother as soon as possible to repeat the Child Conflict Index for another 24-hour period, solicit detailed descriptions of the two types of discipline episodes and then get similar details about a discipline episode that occurred during the interview (e.g., during the clean-up task). Mother then completed the CBCL 1½-5 and the ECBQ at two later times, at averages of 2 and 16 months later (Wave 2 and Wave 3, respectively), with response rates of 98% and 94%, respectively.

Measures

Survey Measures. The *Child Behavior Checklist 1½-5* is a widely used measure of behavior problems in preschool children, with coefficient alphas of .89 to .95 for the broadband scales (Externalizing, Internalizing, and Total Behavior Problems: Achenbach & Rescorla, 2000) in the manual and .81 to .93 across the three waves of our data. The three subscales of the *Early Behavior Questionnaire* (Putnam, Gartstein, & Rothbart, 2006) with the highest loadings on their Effortful Control factor was used to measure Effortful Control, following Spinrad et al., (2007). Coefficient alphas for those three subscales averaged .71 to .89 in Putnam et al. (2006), whereas the alpha for Effortful Control was .88 across the three waves in our data.

The *Nurturing and Discipline Practices Questionnaire* is an expanded version of the *Behavioral Indicators of Nurturance* scale (Larzelere, Sather, Schneider, Larson, & Pike, 1998). It asks for the frequencies of 11 nurturing behaviors in the past 2 days and of 23 discipline practices in the

past 7 days. Only the nurturance scales were used in this study, which were maternal warmth (4 items, α = .87) and involvement (7 items, α = .84). The mean difficulty of each of the two days reported on was scored from the Child Conflict Index (CCI). The CCI asked about any misbehavior problems in 14 situations that day (e.g., waking up, making a mess), scored from 0 = no problem to report to 5 = an extremely difficult problem. An overall question asked how difficult the toddler's behavior was in the past 24 hours, on a scale from 1 = no difficulty to 5 = extreme difficulty. The mean difficulty score across all 15 items for both days was used to indicate how difficult the toddler's behavior was on the days reported on (α = .71, M = 1.1).

Codes from the Detailed Episode Descriptions. Coding began with the initial type of noncompliance and the first action the mother used to respond to that noncompliance. Coding the toddler's noncompliance and the mother' next action continued until the mother considered the discipline episode ended even if a temporary compliance occurred in the middle of the episode. Episodes usually ended with child compliance or with the mother giving in.

Some turns were dropped from the analyses. For example, we dropped 104 intermediate turns that were responding to compliance because the current study was investigating mothers' responses to noncompliance. Other turns dropped from the data set included final turns lacking a subsequent child code (e.g., because the episode ended when the mother gave in), repeated cycles of identical mother and child codes, three "other" codes, bribe/reward codes (on 19 turns), and the 16th to 26th turns on the single longest reported episode. Eliminating these turns was necessary to handle missing or less relevant data or to prevent atypical or repetitive turns from biasing the results. The final Level 1 data set for the multilevel modeling included 451 episodes with a total of 1397 turns (M = 3.1 turns per episode).

Mothers' detailed descriptions of the five discipline episodes were transcribed and coded using an expanded version of Ritchie's (1999) codes. Following Kuczynski and Kochanska (1990; Kuczynski, Kochanska, Radke-Yarrow, & Girnius-Brown, 1987), we distinguished among parent-oriented noncompliance (negotiation and whining), parent-avoiding noncompliance (simple refusal and passive noncompliance), and parent-opposing noncompliance (defiance, hitting, and tantrums). In parent-oriented noncompliance, the child is looking to the parent to resolve the disciplinary situation (Kuczynski & Kochanska, 1990), albeit more competently in negotiation than whining. The parent-avoiding types merely ignored the parent, whereas the parent-opposing types of noncompliance actively defied the parent. Compliance was also coded. The major outcome variable for immediate effects was *noncompliance severity*, which varied from 1=compliance, 2 = parent-oriented, 3 = parent-avoiding, and 4 = parent-opposing. Two or three types of compliance or noncompliance were coded on 11.5% of the child turns. For those turns, the maximum score on noncompliance severity was used for the outcome variable in the multilevel modeling of immediate effects. To account for the type of noncompliance being responded to, however, multiple codes were assigned to one code according to the following priority: defiance, hit, tantrum, negotiate, whine, simple refusal, passive noncompliance, and compliance. When investigating the type of noncompliance the mother was responding to, being oppositional took priority over the other noncompliance types, but parent-oriented noncompliance took priority over parent-avoiding noncompliance, to highlight those contrasting types of noncompliance more than parent-avoiding noncompliance. Note that parent-avoiding noncompliance is intermediate between the other two types of noncompliance in intensity

because parent-oriented noncompliance is considered more skillful. In contrast, it gets trumped by the other two types of noncompliance as a contextual moderator of the effects of disciplinary responses because it is the least active of the three types.

In addition to the 10 disciplinary tactics coded in Ritchie (1999), this study added three others to distinguish the following 13 tactics: Reason, Verbal Power Assertion, Warning, Offer Alternative, Give In, Ignore, Physical Power Assertion, Time Out, Privilege Removal, Spank, Affection/Praise, Yell/Shame, and Model. Based on a 20% sample, coders identified the same phrase to code 81.2% of the time and had good agreement when coding the same phrase (kappas = .85 for noncompliance type, .78 for maternal tactic). Multiple tactics were coded on 22.5% of the mothers' turns. They were assigned to single tactics using the following priorities: spank, nonphysical punishments, warning, physical power assertion, yell/shame, offer alternative, reason, give in, affection/praise, model, verbal power assertion, and ignore. Physical punishment took priority because it is opposed by most current perspectives on parental discipline. Skills included in parent management training had the next priority, because they are opposed by the positive-parenting perspective. Physical power assertion (and the one occurrence of yell/shame) took the next priority because it is also opposed by positive-parenting advocates and by some parenting experts in developmental science. Tactics explicitly recommended by the positive-parenting perspective had the next priorities, followed by tactics that were routine for all parents (e.g., verbal power assertion).

Other Coded Variables. Two variables coded from the initial part of the interview and a measure based on follow-up questions asked after each episode were also used. One of the initial interview questions asked mothers "What last-resort action do you use when nothing else seems to work?" A dummy code was created to distinguish the 34 who identified time-out (n = 32) or privilege removal (n = 2) as their last resort tactic. Others identified physical punishment (n = 54), forced compliance (6), yelling (1), or a milder tactic (7) as their last resort. Two coders agreed on the last resort tactic on 90% of a 30% sample of the cases. The dummy code was z-scored for the multilevel analyses.

Second, we coded the mothers' planfulness based on the extent to which they spontaneously said that their disciplinary responses depended upon some aspects about the child and whether they indicated a planned sequence of responses. All mothers were asked, "How do you generally deal with your toddler's misbehavior?" and some were asked, "What is your overall plan in dealing with misbehavior?" Their answers were coded for being either conditional (e.g., it depends upon the type of misbehavior, how tired the toddler is) or sequential (e.g., her disciplinary response might proceed to one or more subsequent steps depending on how the toddler responds to an earlier step). Evidence of either one qualified mothers as "thinking" parents. Kappa was fair for independent coding (κ = .57), but all discrepancies were resolved by consensus.

Finally, the maximum negative affect per episode was coded as the maximum affect reported in answer to three inter-related questions: "How upset were you at the beginning of this episode?", "Did you get more upset?", and, if so, "How upset were you [when you got the most upset]?" The measure used in this study was the maximum upsetness score during each episode on a 5-point scale, from 1 = not upset to 5 = extremely upset (M = 1.88).

Results

Frequency of Disciplinary Tactics by Noncompliance Type

The first question was whether the discipline tactics used by mothers of toddlers conform to either of the two scientifically based parenting perspectives. Table 1 summarizes the percentage use of each discipline tactic by the type of noncompliance that immediately preceded it. Maternal tactics included those opposed by exclusively positive parenting about 29.3% of the time (warnings or punish were 6.9% and physical power assertion 22.4% of the maternal tactics in Table 1). On the other hand, mothers used tactics that were opposed by behavioral parenting experts when responding to misbehavior 34% of the time, including offering alternatives to the child 17.5% of the time, reasoning 14.2%, and giving in for 3.1% of the tactics used.

How many mothers successfully avoided the disciplinary tactics that are proscribed by these two parenting perspectives? Only 4 of 102 mothers completely avoided all power assertion, warnings, and punishment. On the other hand, only 5 mothers always followed the behavioral parent training proscriptions against offering alternatives or reasoning while responding to noncompliance. One of those mothers avoided both sets of prohibited disciplinary tactics, by ignoring misbehavior on most disciplinary turns (3 of 5), and otherwise giving a command (once), or giving in while expressing affection/praise (once).

Did mothers vary their use of tactics by the type of noncompliance shown by the child? Table 1 shows that they were significantly more likely to use positive tactics when their toddlers were negotiating (offer alternatives, giving in) or whining (reasoning, affection/praise), and significantly less likely to use positive tactics for more severe types of noncompliance, including defiance (less reasoning, offering alternatives, affection/praise, and giving in) or hitting (less offering alternatives). In contrast, they were significantly more likely to use physical power assertion and behavioral parent training skills when toddlers were defiant (physical power assertion) or hitting (warnings and punishments), but less likely to use those tactics for negotiating and whining (less physical power assertion and warnings/punishments for both).

Another pattern that emerged was that ignoring and verbal power assertion, the other two common disciplinary tactics, were used for distinct kinds of noncompliance. Ignoring was used significantly more often for whining and tantrums, but less often for negotiating and defiance. In contrast, verbal power assertion was used significantly more often for passive noncompliance, but less often for whining and tantrums. Most of the other significant variations from the unconditional probability of each disciplinary tactic occurred after passive noncompliance. Mothers were then significantly more likely to use verbal power assertion and modeling, but less likely to give in or to use physical power assertion or warnings/punishment.

Immediate Effects by Noncompliance Type

Next we tested the immediate effectiveness of the following maternal disciplinary responses for reducing noncompliance severity: three categories of positive tactics: (1) reasoning; (2) offering alternatives; and (3) a combination of less frequently used tactics, including modeling appropriate behavior, praise/affection, and giving in; (4) broad power assertion (physical power

assertion and warning/punish); (5) ignoring; and (6) verbal power assertion (mostly commands to start or stop some specified behavior). To test whether their relative effectiveness varied by type of noncompliance, the analyses distinguished among the following noncompliance types: (1) negotiating and whining, (2) simple refusal, (3) tantrums, (4) passive noncompliance, (5) initial defiance or hitting, and (6) repeated use of defiance or hitting. The last two categories were distinguished because authoritative parents were expected to respond to initial oppositional noncompliance with positive tactics (unlike authoritarian parents), but switch to broad power assertion if oppositional noncompliance persisted (in contrast to permissive parents). The analyses used broad power assertion as the referent category so that the effects would contrast the immediate effect of each of the other five disciplinary responses compared to the immediate effect of broad power assertion on noncompliance severity in the next turn within the episode.

Improving causal evidence. Several steps were taken to minimize any confounding variables that could bias the estimates of causal influence on noncompliance severity. First, group-mean centering was used in multilevel modeling to eliminate the effect of all confounding variables except those that varied by turns within a discipline episode. The Level-1 effects were therefore pure within-episode effects of the disciplinary responses, indicating whether the noncompliance severity on the next turn was higher or lower following a given disciplinary response compared to other disciplinary responses within the same episode.

Group-mean centering was used in three-level multilevel modeling, with disciplinary turns nested within episodes, which were nested within mother-toddler dyads. Disciplinary tactics were represented by an episode-centered dummy code at Level 1, a dyad-centered proportion of turns on which each tactic was used at level 2, and a grand-mean centered mean of the Level-2 episode proportions at Level 3.

Second, the analyses also controlled for all covariates that predicted noncompliance severity, because Steiner et al. (2010) showed that controlling for potential confounders was more important than the statistical method used to implement statistical control. Covariates were retained if either their main effect or their interaction with one of the three types of positive parenting were at least marginally significant, $p < .10$, in comparison with broad power assertion, the referent disciplinary response category. At each level, possible covariates were added to the model one at a time to test for both main effects and interactions with any of the positive-parenting tactics. Predictors that were at least marginally significant were combined in a final set of predictors at each level and those that remained at least marginally significant were retained. Table 2 lists the covariates in the final model.

Findings. The primary focus was on whether the immediate effects of the different disciplinary tactics on noncompliance severity varied by the immediately preceding type of noncompliance. This led to a multilevel equivalent of a 5 X 5 (Tactic X Noncompliance Type) ANCOVA at Level-1 using dummy codes for five types of noncompliance (with negotiating/whining as the reference category) and episode-mean-centered dummy codes for five tactics (with broad power assertion as the reference category). Distinguishing six disciplinary responses was preferred to make more specific discriminations than are typical in parenting research. On the other hand, the large number of categories exacerbates the notoriously poor statistical power of naturally occurring interactions (McClelland & Judd, 1993). To enhance statistical power, an a priori

interaction comparison was used (Keppel & Wickens, 2004), contrasting the effects of the three positive parenting response categories (reasoning, offering alternatives, and 3 minor positive tactics) with broad power assertion for types of noncompliance oriented toward and against parents (negotiation/whining vs. defiance/hit). Therefore, the a priori interaction comparison tested the combined coefficients for the three positive parenting interaction coefficients versus broad power assertion, contrasting the two types of noncompliance in the a priori interaction contrast.

It turned out that this focused interaction comparison was significant for both repeated defiance/hit and for all defiance/hit turns combined, χ^2 (1, N = 802 relevant turns) = 3.02, $p <$ .05, 1-tailed, and 9.32, $p < .01$, respectively. These interactions indicated that the relative immediate effects of positive responses vs. broad power assertion were different for negotiating and whining vs. defiance and hitting. Because negotiation/whining is the reference category for the centered dummy codes, the main effect for each tactic at Level 1 indicates its effectiveness in predicting lower noncompliance severity immediately following use of that tactic, controlling for everything else in the multilevel model. Compared to broad power assertion, reasoning and offering alternatives both predicted significantly more de-escalation in noncompliance severity immediately after negotiation or whining. The three infrequent positive tactics (modeling, affection/praise, giving in) did not predict de-escalation for boys, but its interaction with gender indicates that those tactics as a group did predict significantly more de-escalation for girls than did broad power assertion, b = -.57 for the Gender X Minor Positives interaction, $p < .001$.

Compared to two of these three positive parenting responses, broad power assertion was more effective in de-escalating noncompliance severity following defiance or hitting, especially after its first occurrence during a discipline episode. Broad power assertion became more effective than either reasoning or minor positive tactics following initial defiance/hitting. Although offering alternatives remained marginally more effective than broad power assertion when responding to repeated defiance and hitting, the magnitude of its advantage in effectiveness over broad power assertion was otherwise reduced, in contrast to its significantly greater effectiveness when responding to negotiation, whining, and temper tantrums.

Two patterns stood out in addition to the a priori predictions. First, offering alternatives was the most effective tactic overall, b = -.52, $p < .001$, compared to broad power assertion when no interactions were included in the analysis. Although three of the five Offer Alternatives X Noncompliance Type interactions were at least marginally significant, the coefficient for offering alternatives never reversed its sign relative to broad power assertion (except for a trivial -.03 in response to simple refusals). In contrast, the other two types of positive responses became significantly or marginally less effective than broad power assertion when responding to passive noncompliance or initial hitting and defiance.

The second unexpected pattern was that the predicted interactions occurred for passive noncompliance as well as for defiance/hitting. Therefore, Figure 1 summarizes the results after combining the preceding types of noncompliance into three categories: (1) negotiate/whine, (2) simple refusal and tantrums, and (3) passive noncompliance, defiance, and hitting. Offering alternatives predicted the lowest noncompliance severity in response to all three types of noncompliance, although its advantage was reduced for defiance/hit/passive and was not

significant for refusal/tantrums. Of the five other discipline categories, broad power assertion was the most effective response for defy/hit/passive, whereas it was the least effective response for negotiate/whine. This focused interaction contrast (Another Response vs. Broad Power Assertion X Negotiate/Whine vs. Defy/Hit/Passive) was significant for Reasoning, Minor Positives, and Verbal Power Assertion and marginally significant for Offering Alternatives. Ignoring was the only response that failed to interact significantly with the type of noncompliance on this focused interaction contrast. Moreover, these interactions indicated that broad power assertion was significantly more effective than the following tactics in responding to defy/hit/passive noncompliance: Reasoning, Minor Positives, and Verbal Power Assertion. In contrast, the interaction effects when comparing negotiate/whine vs. simple refusals and tantrums combined were smaller and not significant, except for a marginally significant interaction involving Reasoning.

Several other interactions indicated that the relative within-episode effectiveness of positive responses vs. broad power assertion varied by several dyad-level factors (see Table 2). Reasoning was more effective for toddlers higher on effortful control and for mothers who were more planful in responding to misbehavior. The effectiveness of offering alternatives interacted significantly with three dyad-level variables. First, the mean behavioral difficulty shown by the toddlers on these two days decreased the effectiveness of offering alternatives (relative to broad power assertion). The second and third interactions indicated opposing moderating effects of maternal responsiveness on the effectiveness of offering alternatives. The quantity of involvement of the mother with the toddler *decreased* the effectiveness of offering alternatives, whereas the quality of maternal warmth increased its effectiveness as expected. Finally, the three minor positive tactics were more effective for girls than for boys.

The other covariates were included to minimize any other confounds that would influence noncompliance severity. Noncompliance severity was generally greater for toddlers who had more behavioral difficulty on those two days and for non-Hispanic Causasian mothers (marginally), and those who used nonphysical punishment as their last-resort tactic. The most problematic episodes and the observed discipline episodes had greater noncompliance severity than the potentially problematic episodes. Maternal negative affect was marginally associated with more severe noncompliance. Whereas noncompliance severity generally decreased during the duration of the episode at Level 1, longer episodes had higher noncompliance severity on average (episode length at Level 2).

Although broad power assertion was significantly more effective within episodes than the other disciplinary tactics in slightly more cells than the reverse (4 cells compared to 3 cells at $p < .05$), it never compared favorably with any of the other disciplinary tactics according to the coefficients at the episode or dyad levels. All disciplinary tactics except for verbal power assertion looked more effective than broad power assertion according to the proportional usage scores at episode Level 2. These Level-2 associations are more likely to be due to child effects than are the within-episode associations at Level 1. That may explain why broad power assertion appears to be more harmful than positive parenting in typical longitudinal findings, because, similar to the Level 2 and 3 findings here, they fail to rule out confounds such as child effects as thoroughly as the multilevel model in this study.

Long Term Effects

Of course, what works in reducing noncompliance severity immediately could be counterproductive for longer term effects.

Improving causal evidence. To compare immediate effects with longer term effects, it is important to minimize confounds such as child effects as much as possible (Larzelere & Cox, 2013). Toward that end, the longitudinal effects of the various disciplinary responses on four major outcome variables also controlled for a wider range of relevant covariates than usual. Most importantly, they all controlled for Wave-1 scores on the outcome variable. They adjusted for those initial differences in two ways, predicting residualized and simple gain scores from Wave 1 to either Wave 2 or Wave 3. Predictions of those two types of gain scores often contradict each other (e.g., Larzelere, Ferrer, et al., 2010), whereas unbiased causal effects will yield robust evidence across both types of gain scores under ideal conditions (e.g., randomized trials with large sample sizes).

In addition, demographic and temperament variables were tested for their ability to predict either type of gain score for each outcome through either Wave 2 or Wave 3. Any variable that was at least marginally significant in predicting one of those gain scores over the given time period was retained for those analyses.

Findings. With those extensive controls, the effects of seven disciplinary responses were analyzed focusing on their interactions with noncompliance type. Noncompliance type was measured as the difference between the proportion of noncompliance codes that were oppositional (defiance, hitting, tantrums) minus the proportion that were parent-oriented (negotiation, whining). The seven disciplinary responses differed slightly from those tested for immediate effects. In particular, broad power assertion was divided into a code for physical power assertion and the codes for skills emphasized in behavioral parent training. The latter included warnings (40.6%), time out (36.4%), privilege removal (4.2%), and spanking (18.8%). (Spanking was used as an enforcement for compliance with time out in these protocols from the late 1960s to the mid-1990s, but is no longer recommended. It was included because it fits there better than with any other category, and has not been replaced by ordinary parents with the forced room isolation that proved to be as effective as the traditional spank enforcement.) The other disciplinary responses were the same as for the analyses of immediate effects.

Altogether there were analyses of 7 disciplinary responses X 4 outcomes X 2 subsequent waves X 2 types of gain scores = 112 analyses. To be able to detect curvilinear relationships, each disciplinary response was divided into low, moderate, and frequent usage, where the extremes consisted of 17 to 25% of the distribution, unless more than 25% of the mothers reported no use of that particular disciplinary response (which occurred for the minor three positive tactics together, ignoring, and punishments/warnings). Dummy codes were used for the moderate and frequent usage categories, which contrasted each of them with low usage as the reference category. Out of those 112 analyses, there were only 3 significant main effects for disciplinary responses, which were replicated for predicting residualized and simple gain scores. Although they are most likely explained as Type I errors, they indicated that both intermediate and frequent use of physical power assertion predicted long-term gains in effortful control (during

the next 16 months, on average), and frequent use of the three minor positive tactics predicted long-term increases in internalizing problems. But the unusually extensive controls apparently reduced the main effects of all disciplinary tactics to non-significance.

Most of the significant findings involved interactions of disciplinary responses with the predominant type of noncompliance shown by the toddlers. Despite the low power for detecting interactions in naturally occurring data (McClelland & Judd, 1993), 8 overall interactions out of 56 were significant, $p < .05$, 2-tailed. Because of the low power for interactions, we report 1-tailed tests, but only for interaction effects. In many cases, a more specific interaction (e.g., with either intermediate or frequent use of a disciplinary tactic, instead of the overall interaction involving both of them as contrasted with low usage) would have met a higher standard of significance if it had been tested by itself, instead of in combination with the other usage level. All the significant overall interactions ($p < .05$, 1- or 2-tailed) occurred for four of the seven disciplinary responses: offering alternatives (3 of 8), punishments/warnings (3), reasoning (3), and ignoring (3). There were no significant interactions involving verbal power assertion, the minor positive tactics, or physical power assertion. Therefore the results from the four disciplinary responses whose effects varied by the predominant type of noncompliance are summarized next.

Note that most of the interactions effects are robust for predictions of both residualized and simple gains scores, a level of robustness rarely documented for disciplinary responses in previous longitudinal studies (e.g., not for spanking, the most replicated disciplinary response: Larzelere et al., 2015; Larzelere, Ferrer, et al., 2010). Although not definitive, this provides stronger evidence of minimal biases in causal estimates than is typical of longitudinal studies.

Whereas offering alternatives was the most effective disciplinary tactic in reducing noncompliance severity immediately, its overuse led to increases in all behavior problems for the most oppositional toddlers (Table 3). The interaction was strongest for predicting short-term gains in externalizing behavior problems during the next two months. Figure 1 illustrates this effect, based on Preacher's (2006) Rweb program for interactions. This graph contrasts the 21% of mothers who offered alternatives for more than 29% of their disciplinary actions versus those who never offered alternatives to their toddler. Such high use of offering alternatives led to significant increases in externalizing behavior problems for the 20% most oppositional toddlers, whereas it led to significant decreases in externalizing behavior problems for the 7% whose noncompliance was mostly negotiation or whining rather than oppositional.

The other significant interactions were due to similar patterns in how the effects of high use of offering alternatives influenced short-term gains in externalizing problems during the next two months. Although the overall Offer Alternatives X Noncompliance Type was significant for short-term gains in internalizing and total problems, the specific interaction contrasting frequent vs. no use was only marginally significant, $p < .10$, 2-tailed for total problems, 1-tailed for internalizing problems. Although the specific interactions were of borderline significance, both specific interactions became significant if the specific interaction for moderate use was dropped from the analysis (equivalent to fixing the latter effect to zero). Note also that these two overall significant interactions were due not only to the specific interaction of high usage with noncompliance type, but due also to the opposite direction of the interaction contrasting

moderate use of offering alternatives. In contrast to offering alternatives too frequently, a moderate use of offering alternatives tended to predict somewhat lower internalizing and total problems during the subsequent two months for oppositional toddlers. Note also that moderate use predicted longer term reductions in externalizing and total behavior problems during the next 16 months for them, although the evidence was only marginally significant.

Overall, offering alternatives had the most optimal immediate effects and also had beneficial longitudinal effects when used either (1) less than 29% of the time or (2) with non-oppositional toddlers. The only adverse effect of offering alternatives showed up with increased behavior problems during the next two months when that tactic was relied on too often with the most oppositional toddlers.

So how should parents respond to oppositional toddlers? Part of the answer is indicated by the interaction of punishments and warnings with noncompliance type. It was the only other discipline response that had three clearly significant interactions with noncompliance type in predicting subsequent gains in behavior problems, $p < .05$, 2-tailed. Consistent with behavioral parent training, punishments/warnings interacted with the type of noncompliance such that they reduced behavior problems in the most oppositional toddlers, but only if used less than 16% of the time. Figure 3 illustrates the strongest interaction, which was for short-term effects of those tactics on total behavior problems. Punishments and warnings significantly reduced total behavior problems for the 12% most oppositional toddlers, but they significantly increased total behavior problems for the 35% whose noncompliance was mostly directed toward the mother (negotiating and whining). Significant overall interactions indicated similar patterns for longer term changes in total behavior problems and for short-term changes in externalizing behavior problems. In contrast, overly frequent use of punishment and warnings predicted marginal increases in total behavior problems and externalizing problems, $p < .10$, 1-tailed for all short-term effects. A similar interaction pattern was marginally significant for long-term changes in externalizing problems as well, $p < .10$, 1-tailed.

In contrast to being the least effective disciplinary response for immediate effects with oppositional toddlers, frequent use of reasoning reduced subsequent externalizing problems long-term, as indicated by a significant Reasoning X Noncompliance Type interaction. The interaction was also marginally significant short-term, $p < .05$, 1-tailed, such that using reasoning at least 25% of the time significantly decreased externalizing problems over the subsequent two months for the 11% most oppositional toddlers, $p < .05$, 2-tailed. Reasoning also had marginally significant interactions for short- and long-term changes in total behavior problems, $p < .10$, 1-tailed.

The only other significant interactions involved ignoring. Frequent use of ignoring showed marginal improvements in total problems (short-term only) and in effortful control (short- and long-term) for the most oppositional toddlers, whereas moderate use of ignoring led to marginal decreases in effortful control for them during both periods of time.

Discussion

By investigating how the effectiveness of disciplinary responses vary by the type of noncompliance in toddlers, this study showed how to reconcile the contradictory recommendations of positive parenting and behavioral parent training with each other. In general, positive parenting responses are sufficiently effective when toddlers' noncompliance is oriented toward parents, in the form of negotiations and whining. For those children, offering alternatives whenever possible reduces their noncompliance severity immediately and also reduces behavior problems over the next two months. Reasoning also produces movement toward compliance immediately in such children, whereas power assertion, including warnings and threats are the least effective disciplinary responses in terms of immediate effects.

In contrast, the types of power assertive skills that are taught in behavioral parent training are the second most effective response for the most oppositional toddlers immediately and also reduce behavior problems subsequently if used on less than one-sixth of disciplinary turns in this small sample of somewhat difficult discipline episodes. What else should be used with the most oppositional toddlers? The results indicate that offering alternatives works with them also, as long as that response is not used too much (i.e., used on less than 29% of disciplinary turns). Reasoning has both positive and negative effects for the most oppositional toddlers. On one hand, it is the least effective disciplinary response to oppositional noncompliance in terms of immediate reductions in noncompliance severity. On the other hand, it is effective at reducing externalizing behavior problems long-term for the most oppositional toddlers when used frequently (i.e., on at least ¼ of disciplinary turns). Frequent use of ignoring also has marginally significant effects in reducing total behavior problems and improving effortful control in oppositional toddlers.

Whereas the results for the most oppositional toddlers confirm the parenting skills used to respond to noncompliance in behavioral parent training, the results also show how the correlational evidence relied on by positive parenting advocates seems to support their viewpoint. For the most part, the correlational evidence supporting positive parenting is consistent with this study's stronger causal evidence for toddlers when their noncompliance is oriented toward the parent (negotiating and whining). But the correlational evidence is misleading for toddlers whose noncompliance is oppositional. It is especially misleading when the correlational evidence in support of positive parenting is used to support absolute conclusions that parents should never (or hardly ever) use any power assertive discipline responses under any circumstances.

This study incorporated an unusually thorough set of innovations to move beyond correlational evidence to approximate unbiased causal evidence more closely. First, the study aimed to get information on the same number of discipline episodes from each mother-toddler dyad. This prevented differential frequencies of episodes from biasing the results. Second, it controlled statistically for initial differences on each outcome variable in two ways, as residualized and simple gain scores. Third, it controlled for variables most likely to influence parental selection of disciplinary responses, such as temperament indicators and the difficulty of the toddler's behavior on the two days of the disciplinary episodes. Fourth, it used episode-mean centering to isolate within-episode immediate effects, which controls for all episode-invariant variables, whether measured or not. Fifth, it included any demographic variables that predicted either type of gain in the outcome variables. Sixth, it incorporated a short longitudinal interval to

approximate a more realistic causal lag in addition to a more typical interval between waves, because parents want to see improvement in a matter of days or weeks, not years. The success of this combination of innovations is indicated by the unusual consistency of the results across analyses of residualized and simple gain scores in the longitudinal outcome variables. Such robust effects of spanking have never been documented, and the few attempts to do so have failed (e.g., Larzelere, Cox, & Smith, 2010; Larzelere, Ferrer, et al., 2010), despite the fact that spanking has been investigated more than any other disciplinary response.

These efforts would still have failed to identify conditions under which the tools used to respond to noncompliance in behavioral parent training improve toddlers' behavior over the subsequent two to 16 months, except for the fact that this study investigated curvilinear effects of the most common disciplinary responses. The usual linear effects in statistics would have contrasted over-usage of punishments and warning vs. no use, which would have favored no use of such responses even for oppositional toddlers. Only moderate use of punishment and warnings proved to be effective in reducing subsequent behavior problems in the most oppositional toddlers. As Barber and Xia (2013) observed, research on effective behavioral control has been hindered by the failure to investigate curvilinear effects instead of implicitly contrasting over-usage with under-usage.

There are several reasons why moderate use of punishment and warnings are more effective with oppositional toddlers than either under-usage or over-usage. First, this study indicates that other disciplinary responses are also important for reducing behavior problems in oppositional toddlers. Offering alternatives whenever possible is effective immediately with all toddlers and only has longer-term adverse effects on oppositional children's behavior problems if used on more than 29% of disciplinary turns. To our surprise, frequent use of reasoning decreases behavior problems subsequently with oppositional toddlers, even though it is the least effective response for immediate reduction of noncompliance severity for oppositional noncompliance. We thought that the compliance of the most oppositional children had to improve to typical levels before they could be positively influenced by reasoning. Otherwise we thought they would just tune parents out. Longitudinal analyses of Canadian data also showed that the combination of frequent use of reasoning and frequent use of nonphysical punishment was a particularly effective combination for the most antisocial children (Larzelere, Ferrer, & Kuhn, 2006).

The effectiveness with which warnings and punishment are used would lead them to be used moderately as opposed to frequently. Behavioral parent training's effectiveness may be largely due to the fact that it trains parents to use a consistent set of skillful responses to noncompliance. As Roberts (1990) has shown, consistent use of a sequence of a clear command, a single warning, time out, and an enforcement for time out results in a rapid decrease in the frequency of the latter tactics, as defiant children learn to cooperate with verbal directives. In contrast, unskillful and inconsistent use of multiple warnings and punishments can lead to overly frequent use of such disciplinary responses.

Putting this together, behavioral parent training may need to consider how parents should transition from these contingent responses to misbehavior to incorporating offering alternatives and reasoning into their disciplinary repertoire (Larzelere & Kuhn, 2005).

Offering alternatives was the most effective disciplinary response for immediate improvement for all toddlers and had adverse longer term effects only if used too frequently with oppositional toddlers. It may be effective because it permits a toddler to express their newfound independence in a way that modifies the parent's initial request. It avoids a direct conflict between what the parent and the toddler want the toddler to do at that time. But it does so by helping toddlers move beyond independence to interdependence as they work together with their parents toward cooperative resolutions of differing desires (Mahler, Pine, & Bergman, 2000).

So parents should trust themselves rather than adhere to extreme views of parental discipline that restrict their disciplinary options unnecessarily. The near-absolute opposition to all power assertion by positive parenting advocates will prevent parents from using disciplinary responses that have proven to be effective with oppositional-defiant young children (Eyberg et al., 2008). On the other hand, the opposition to positive disciplinary tactics when responding to misbehavior needs to be reconsidered by behavioral parent training experts. Parents need the full range of nonabusive disciplinary responses, although power assertive responses are less necessary for the most cooperative children.

References

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Barber, B. K., & Xia, M. (2013). The centrality of control to parenting and its effects. In R. E. Larzelere, A. S. Morris & A. W. Harrist (Eds.), *Authoritative parenting: Synthesizing nurturance and discipline for optimal child development* (pp. 61-87). Washington, DC: American Psychological Association.

Barkley, R. A. (2013). *Defiant children: A clinician's manual for assessment and parent training* (3rd ed.). New York: Guilford.

Baumrind, D., Larzelere, R. E., & Owens, E. B. (2010). Effects of preschool parents' power assertive patterns and practices on adolescent development. *Parenting: Science and Practice, 10*, 157-201.

Bell, R. Q., & Chapman, M. (1986). Child effects in studies using experimental or brief longitudinal approaches to socialization. *Developmental Psychology, 22*, 595-603.

Bell, R. Q., & Harper, L. V. (1977). *Child effects on adults*. Hillsdale, NJ: Erlbaum.

Durrant, J. E. (2007). *Positive discipline: What it is and how to do it*. Bangkok, Thailand: Save the Children Sweden.

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology, 37*, 215-237.

Forehand, R., & Long, N. (2010). *Parenting the strong-willed child* (3rd ed.). New York: McGraw-Hill.

Forehand, R., & McKinney, B. (1993). Historical Overview of Child Discipline in the United States: Implications for Mental Health Clinicians and Researchers. *Journal of Child & Family Studies, 2*(3), 221-228.

Frankel, F., & Weiner, H. (1990). The child conflict index: Factor analysis, reliability, and validity for clinic-referred and nonreferred children. *Journal of Clinical Child Psychology, 19*, 239-248.

Gershoff, E. T. (2002). Corporal punishment by parents and associated child behaviors and experiences: A meta-analytic and theoretical review. *Psychological Bulletin, 128*, 539-579.

Hoffman, M. L. (2000). *Empathy and moral development*. Cambridge, UK: Cambridge University Press.

Hulbert, A. (2003). *Raising America: Experts, parents, and a century of advice about children*.

Keppel, G., & Wickens, T. D. (2004). *Design and analysis* (4th ed.). Upper Saddle River, NJ: Pearson Education International.

Kochanska, G., Aksan, N., & Joy, M. E. (2007). Children's fearfulness as a moderator of parenting in early socialization: Two longitudinal studies. *Developmental Psychology, 43*(1), 222-237.

Kuczynski, L., & Kochanska, G. (1990). Development of children's noncompliance strategies from toddlerhood to age 5. *Developmental Psychology, 26*(3), 398-408.

Kuczynski, L., Kochanska, G., Radke-Yarrow, M., & Girnius-Brown, O. (1987). A developmental interpretation of young children's noncompliance. *Developmental Psychology, 23*, 799-806.

Larzelere, R. E., & Cox, R. B., Jr. (2013). Making valid causal inferences about corrective actions by parents from longitudinal data. *Journal of Family Theory & Review, 5*, 282-299.

Larzelere, R. E., Cox, R. B., Jr., & Mandara, J. (2013). Responding to misbehavior in young children: How authoritative parents enhance reasoning with firm control. In R. E. Larzelere, A. S. Morris & A. W. Harrist (Eds.), *Authoritative parenting: Synthesizing nurturance and discipline for optimal child development* (pp. 89-111). Washington, DC: American Psychological Association.

Larzelere, R. E., Cox, R. B., Jr., & Smith, G. L. (2010). Do nonphysical punishments reduce antisocial behavior more than spanking? A comparison using the strongest previous causal evidence against spanking. *BMC Pediatrics, 10*(10).

Larzelere, R. E., Cox, R. B., Jr., & Swindle, T. M. (2015). Many replications do not causal inferences make: The need for critical replications to test competing explanations of non-randomized studies. *Perspectives on Psychological Science, 10*, 380-389.

Larzelere, R. E., Ferrer, E., & Kuhn, B. R. (2006, August). *Antisocial trajectories by balance of disciplinary reasoning and nonphysical punishment.* Paper presented at the annual conference of the American Psychological Association, New Orleans, LA.

Larzelere, R. E., Ferrer, E., Kuhn, B. R., & Danelia, K. (2010). Differences in causal estimates from longitudinal analyses of residualized versus simple gain scores: Contrasting controls for selection and regression artifacts. *International Journal of Behavioral Development, 34*(2), 180-189.

Larzelere, R. E., & Kuhn, B. R. (2005). Enhancing behavioral parent training with an extended discipline ladder. *The Behavior Therapist, 28*(5), 105-106, 108.

Larzelere, R. E., Sather, P. R., Schneider, W. N., Larson, D. B., & Pike, P. L. (1998). Punishment enhances reasoning's effectiveness as a disciplinary response to toddlers. *Journal of Marriage and the Family, 60*, 388-403.

Mahler, M. S., Pine, F., & Bergman, A. (2000). *The psychological birth of the human infant: Symbiosis and individuation*. New York: Basic Books.

Markham, L. (2012). *Peaceful parent, happy kids: How to stop yelling and start connecting*. New York: Penguin Group.

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin, 114*, 376-390.

McNeil, C. B., & Hembree-Kigin, T. L. (2010). *Parent-child interaction therapy* (2nd ed.). New York: Springer.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia Press.

Pelham, W. E., Jr., & Fabiano, G. A. (2008). Evidence-based psychosocial treatments for attention-deficit/hyperactivity disorder. *Journal of Clinical Child and Adolescent Psychology, 37*, 184-214.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31*(4), 437-448.

Putnam, S. P., Gartstein, M. A., & Rothbart, M. K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant Behavior & Development, 29*, 386-401.

Ritchie, K. L. (1999). Maternal behaviors and cognitions during discipline episodes: A comparison of power bouts and single acts of noncompliance. *Developmental Psychology, 35*, 580-589.

Roberts, M. W. (2008). Parent training. In M. Hersen & A. M. Gross (Eds.), *Handbook of clinical psychology. Vol. 2: Children and adolescents.* (pp. 653-693). Hoboken, NJ: Wiley.

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

Shaffer, D. R., & Kipp, K. (2007). *Developmental psychology: Childhood & adolescence* (7th ed.). Belmont, CA: Thomson.

Siegel, D. J., & Bryson, T. P. (2014). *No-drama discipline*. Los Angeles: Mind Your Brain, Inc.

Society for Research in Human Development. (2013). *Resolution on corporal punishment (CP) of children*. Retrieved May 25, 2015, from http://www.srhdonline.org/reolution-on-corporal-punishment.html

Spinrad, T. L., Eisenberg, N., Gaertner, B., Popp, T., Smith, C. L., Kupfer, A., et al. (2007). Relations of maternal socialization and toddlers' effortful control to children's adjustment and social competence. *Developmental Psychology, 43*(5), 1170-1186.

Spock, B. (1946). *The common sense book of baby and child care*. New York: Duell, Sloan and Pearce.

Steinberg, L. (2001). We know some things: Parent-adolescent relationships in retrospect and prospect. *Journal of Research on Adolescence, 11*, 1-19.

Steiner, P. M., Cook, T. D., Shadish, W. R., & Clark, M. H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods, 15*, 250-267.

Watson, J. B. (1928). *Psychological care of infant and child*. New York: W. W. Norton & Co.

Webster-Stratton, C. (1992). *The incredible years*. Toronto, Ontario: Umbrella Press.

Table 1

Unconditional and Conditional Probabilities of Disciplinary Tactics by the Immediately Preceding Type of Noncompliance

	Given the Following Type of Noncompliance							
Tactic Group	Nego-tiate	Whine	Simple Refusal	Passively Noncomply	Tantrum	Hit	Defiance	Overall
Reason	19.4%	19.5%*	20.8%[a]	12.9%	14.8%	17.7%	8.4%**	14.2%
Offer Alternative	34.7***	19.8	17.7	18.5	17.6	4.8*	12.1**	17.5%
Affection/Praise	0.8	8.2***	6.2[a]	1.7	3.5	1.6	1.4*	3.2%
Modeling	0.0[a]	1.2	4.2	7.3***	0.0[a]	0.0	2.1	2.6%
Give In	12.9***	3.9	2.1	1.0*	5.6[a]	0.0	0.9*	3.1%
Verbal Power Assertion	18.5	8.6***	16.7	33.6***	9.2**	21.0	21.9	19.8%
Ignore	0.8***	25.3***	5.2	7.7	19.7***	4.8	4.4***	10.2%
Phys Power Assertion	11.3**	12.5***	13.5[a]	14.0**	19.0	24.2	40.0***	22.4%
Warning/Punish	1.6*	1.2***	13.5*	3.1*	10.6[a]	25.8***	8.8	6.9%
% Noncompliance Type	8.9%	18.4%	6.9%	20.5%	10.2%	4.4%	30.8%	100.0%[b]

Note. Based on 1397 turns within 451 discipline episodes for 102 mother-toddler dyads.

[a]$p < .10$, compared to Overall (unconditional) probability of that disciplinary tactic.

[b]After dropping turns with immediately preceding compliance, bribe/rewards, and other types of tactics.

*$p < .05$, compared to Overall (unconditional) probability of that disciplinary tactic. **$p < .01$. ***$p < .001$.

Table 2
Multilevel Model of Immediate Effects of Disciplinary Tactics on Noncompliance Severity by Preceding Noncompliance Types

	b
Level 3 Covariates (Main Effects)	
Child characteristics	
Female gender[a]	.02
Effortful control[b]	-.04
Mean difficulty[b]	.09*
Maternal characteristics	
White race[a]	.16[e]
Warmth[b]	-.06
Involvement[b]	.01
Planfulness[ab]	-.04
NonPhys Punish LR[b]	.07*
Mean Probability of Tactics (across episodes)	
Reasoning[b]	-.04
Offer alternatives[b]	-.45
3 Other positives[b]	-1.01**
Ignore[b]	-.52
Verbal power assert[b]	.08
Level 2 Covariates (Main Effects)	
Episode characteristics	
Episode type (Potentially problematic reference category	
Most problematic[a]	.42***
Observed episode[a]	.61***
Length (total turns)[b]	.15***
Maternal neg affect[b]	.05[e]
Mean probability of tactics (within episode)	
Reasoning[c]	-.39*
Offer alternatives[c]	-.58***
3 Other positives[c]	-.53**
Ignore[c]	-.50*
Verbal power assert[c]	-.17
Level 1 Covariates (Main Effects and Interactions)	
Turn number[d]	-.19***
Preceding Noncompliance Type (Negotiation/Whining reference category)	
Defiance/Hit Repeat[a]	.18
Defiance/Hit Initial[a]	.22*
Passive noncomply[a]	.14
Tantrum[a]	.41***
Simple refusal[a]	.02
Disciplinary Tactic	*Intercept* *Interaction* bs *(add to Intercept for Effect*

(Broad Power Assertion reference category) *After Each Noncompliance Type)*

	(Effect given Negotiate or Whine)	*Simple Refusal*	*Tantrum*	*Passive Noncomply*	*Defy/Hit Initial*	*Defy/Hit Repeated*
Reasoning[ac]	-.55*	.64	.61	1.32***/*	1.00**/[e]	1.17*[/ns]
X Planfulness[ab]	-.64*					
X Effortful control[b]	-.26*					
Offer alternative[ac]	-.95***	.98[e/ns]	.05[ns/*]	.76*[/ns]	.58[e/ns]	.32[ns/e]
X Warmth[b]	-.35*					
X Involvement[b]	.32*					
X Mean difficulty[b]	.29*					
3 other positives[ac]	-.43	.87[e/ns]	-.18[ns/ns]	1.15*/[e]	1.43**/**	.43
X Female gender[a]	-.57*					
Ignore[ac]	-.22	.12	.34	.77[e/ns]	.72[e/ns]	-.06
Verbal power assert[ac]	-.25	.39	.35	.79*/*	.66[a/*]	.76[e/ns]

[a] Dummy variable. [b]Grand-mean centered. [c]Group-mean centered. [d]Centered at the 1st turn.
[e]$p < .10$.
*$p < .05$, **$p < .01$. ***$p < .001$. For interaction *b*s above, symbols before the slash indicate the significance of the interaction, whereas those after the slash indicate the significance of the simple effect of that tactic given that noncompliance type (add the intercept to the interaction coefficient to get the simple effect for that cell, e.g., -.55 +.64 = .09 for the simple effect of Reasoning when responding to Simple Refusal, compared to Broad Power Assertion). [ns] = not significant, i.e., $p > .10$.

Table 3

Standardized Regression Coefficients Predicting Gains in CBCL Broadband Scales and Effortful Control from Intermediate and Frequent Offering Alternatives by Predominant Noncompliance Type Controlling for Selected Covariates

	W2|W1[a]	W2-W1[b]	W3|W1[a]	W3-W1[b]
		Externalizing Problems		
Step 1: Baseline: Change in R^2	.65***	.10*	.56***	.17**
W1 Externalizing	.79***		.71***	
Demographic covariates				
Occupational prestige	.06	.13	.17*	.27**
Child age			-.10	-.18[d]
White race	.08	.14		
Number of children	-.16**	-.25*		
Temperament covariates				
Negative emotionality			-.04	-.25*
Lethargic			.21**	.23*
Step 2: Main effects: Change in R^2	.01	.03	.01	.01
Prop(Oppos – Negot&Whine)	-.01	-.02	-.02	-.06
Offering alternatives (zero use as reference category)				
Used from 1 to 29%	.13	.19	-.09	-.12
Used > 29%	.02	.03	-.02	-.06
Step 3 : Interactions: Change in R^2	.04**	.10**	.02[c]	.05[d]
Oppos-Negot X Moderate	.01[e]	.01[e]	-.32[de]	-.46[de]
Oppos-Negot X Frequent	.23**	.36**[e]	-.01	.01[e]
		Internalizing Problems		
Step 1: Baseline: Change in R^2	.58***	.20***	.37***	.13**
W1 Internalizing	.69***		.42***	
Demographic covariates				
Occupational prestige			.14	.24*
White race	.13[d]	.19*		
Number of children	-.15*	-.17[d]		
Temperament covariates				
ODD symptoms	.25**	.27*	.24*	.16
Negative emotionality			.07	-.31**
Low surgency	-.20*	-.40***		
Step 2: Main effects: Change in R^2	.01	.01	.01	.00
Prop(Oppos – Negot&Whine)	-.06	-.04	-.05	.01
Offering alternatives (zero use as reference category)				
Used from 1 to 29%	.00	-.05	.05	.02
Used > 29%	.08	.06	.09	.04
Step 3 : Interactions: Change in R^2	.04*	.06*	.01	.01

Oppos-Negot X Moderate	-.14[e]	-.15[e]	-.10	-.12
Oppos-Negot X Frequent	.15[ce]	.20[ce]	.03	.05
	Total Behavior Problems			
Step 1: Baseline: Change in R^2	.67***	.18**	.58***	.20**
W1 Total Behavior Problems	.75***		.54**	
Demographic covariates				
Occupational prestige	.01	.02	.16*	.28**
Child age			-.15*	-.21*
White race	.12[d]	.20*		
Number of children	-.15*	-.22*		
Temperament covariates				
ODD symptoms	.17[d]	.11	.25*	.12
Negative emotionality			-.06	-.36**
Lethargic			.17*	.16
Low surgency	-.15[d]	-.32**		
Step 2: Main effects: Change in R^2	.00	.00	.00	.00
Prop(Oppos – Negot&Whine)	-.02	-.01	-.03	-.02
Offering alternatives (zero use as reference category)				
Used from 1 to 29%	.04	.01	-.04	-.07
Used > 29%	.02	.00	.00	-.02
Step 3 : Interactions: Change in R^2	.03*	.06*	.02	.03[c]
Oppos-Negot X Moderate	-.09[e]	-.12[e]	-.29[ce]	-.39[ce]
Oppos-Negot X Frequent	.15[de]	.23[de]	-.01	.01
	Effortful Control			
Step 1: Baseline: Change in R^2	.58***	.04	.46***	.11*
W1 Effortful control	.73***		.53***	
Demographic covariates				
Child age	-.09	-.16		
Family income			-.17[d]	-.23*
Married			.21*	.32**
Temperament covariate				
Negative emotionality	-.12[d]	-.14	-.27**	-.17[d]
Step 2: Main effects: Change in R^2	.01	.01	.02	.03
Prop(Oppos – Negot&Whine)	-.04	-.06	-.13	-.14
Offering alternatives (zero use as reference category)				
Used from 1 to 29%	.03	.04	.10	.09
Used > 29%	.08	.10	.04	-.01
Step 3 : Interactions: Change in R^2	.01	.03	.00	.01
Oppos-Negot X Moderate	-.18	-.29	-.08	-.17
Oppos-Negot X Frequent	-.14[c]	-.22[c]	-.07	-.15

[a]Predicting behavior problems or effortful control at Wave 2 or 3 controlling for Wave-1 scores on the same variable.

[b]Predicting simple gains in behavior problems or effortful control from Wave 1 to either Wave 2 or to Wave 2. [c]$p < .10$, 1-tailed (reported for interactions only). [d]$p < .10$, 2-tailed ($p < .05$, 1-tailed). [e]This interaction would be "more" significant if entered by itself, without the other interaction, e.g., , $p < .05$ instead of $p < .10$.
*$p < .05$. ** $p < .01$. *** $p < .001$, all 2-tailed tests.

Table 4

Standardized Regression Coefficients Predicting Gains in CBCL Broadband Scales and Effortful Control from Intermediate and Frequent Punishments and Warnings by Predominant Noncompliance Type Controlling for Selected Covariates

	W2|W1[a]	W2-W1[b]	W3|W1[a]	W3-W1[b]
		Externalizing Problems		
Step 1: Baseline: Change in R^2	.65***	.10*	.56***	.17**
Step 2: Main effects: Change in R^2	.01	.03	.00	.00
Prop(Oppos – Negot&Whine)	-.03	-.05	-.02	-.06
Punishment and warnings (zero use as reference category)				
Used from 1 to 16%	.11	.17	.00	.01
Used > 16%	.00	-.03	.01	-.01
Step 3 : Interactions: Change in R^2	.04**	.10**	.02[c]	.04[c]
Oppos-Negot X Moderate	-.24*[e]	-.39**	-.13[e]	-.20[e]
Oppos-Negot X Frequent	.14[de]	.23[de]	.16[ce]	.20[c]
		Internalizing Problems		
Step 1: Baseline: Change in R^2 (see Table 3 for details)	.58***	.20***	.37***	.13**
Step 2: Main effects: Change in R^2	.01	.00	.01	.01
Prop(Oppos – Negot&Whine)	-.07	-.05	-.06	.01
Punishment and warnings (zero use as reference category)				
Used from 1 to 16%	.01	-.01	.03	-.01
Used > 16%	.04	.03	-.07	-.10
Step 3 : Interactions: Change in R^2	.02	.03	.02	.02
Oppos-Negot X Moderate	-.17[ce]	-.21[c]	-.23[d]	-.22[c]
Oppos-Negot X Frequent	.10[e]	.11	.01	.01
		Total Behavior Problems		
Step 1: Baseline: Change in R^2 (see Table 3 for details)	.67***	.18**	.58***	.20**
Step 2: Main effects: Change in R^2	.01	.00	.00	.00
Prop(Oppos – Negot&Whine)	-.03	-.01	-.03	-.02
Punishment and warnings (zero use as reference category)				
Used from 1 to 16%	.01	.01	.01	.01
Used > 16%	-.02	-.05	-.01	-.02
Step 3 : Interactions: Change in R^2	.04**	.10**	.04*	.06*
Oppos-Negot X Moderate	-.27**	-.43**	-.25*	-.35*
Oppos-Negot X Frequent	.12[ce]	.19[ce]	.11[e]	.13[e]
		Effortful Control		
Step 1: Baseline: Change in R^2	.58***	.04	.46***	.11*

(see Table 3 for details)				
Step 2: Main effects: Change in R^2	.01	.03	.02	.02
Prop(Oppos – Negot&Whine)	-.03	-.04	-.13	-.13
Punishment and warnings (zero use as reference category)				
Used from 1 to 16%	-.11	-.16	.02	.01
Used > 16%	-.00	-.00	-.02	-.04
Step 3 : Interactions: Change in R^2	.01	.01	.01	.01
Oppos-Negot X Moderate	.01	.01	.07	.09
Oppos-Negot X Frequent	-.11	-.16	-.10	-.12

[a]Predicting behavior problems or effortful control at Wave 2 or 3 controlling for Wave-1 scores on the same variable.
[b]Predicting simple gains in behavior problems or effortful control from Wave 1 to either Wave 2 or to Wave 2.
[c]$p < .10$, 1-tailed (reported for interactions only). [d]$p < .10$, 2-tailed ($p < .05$, 1-tailed). [e]This interaction would be "more" significant if entered by itself, without the other interaction, e.g., , $p < .05$ instead of $p < .10$.
*$p < .05$. ** $p < .01$. *** $p < .001$, all 2-tailed tests.

Table 5

Standardized Regression Coefficients Predicting Gains in CBCL Broadband Scales and Effortful Control from Intermediate and Frequent Reasoning by Predominant Noncompliance Type, Controlling for Selected Covariates

	W2|W1[a]	W2-W1[b]	W3|W1[a]	W3-W1[b]
		Externalizing Problems		
Step 1: Baseline: Change in R^2 (see Table 3 for details)	.65***	.10*	.56***	.17**
Step 2: Main effects: Change in R^2	.00	.01	.01	.01
Prop(Oppos – Negot&Whine)	-.01	-.03	-.02	-.06
Reasoning (zero use as reference category)				
Used from 1 to 24%	.01	.01	-.09	.13
Used > 24%	-.05	-.07	-.06	.09
Step 3 : Interactions: Change in R^2	.02[d]	.05[d]	.03*	.06*
Oppos-Negot X Moderate	-.04	-.05	.12[e]	.15[e]
Oppos-Negot X Frequent	-.19*	-.31*	-.16[ce]	-.24[de]
		Internalizing Problems		
Step 1: Baseline: Change in R^2 (see Table 3 for details)	.58***	.20***	.37***	.13**
Step 2: Main effects: Change in R^2	.01	.01	.04	.04
Prop(Oppos – Negot&Whine)	-.06	-.04	-.08	-.02
Reasoning (zero use as reference category)				
Used from 1 to 24%	-.09	-.12	.10	.16
Used > 24%	-.04	-.06	-.12	-.07
Step 3 : Interactions: Change in R^2	.00	.01	.02	.02
Oppos-Negot X Moderate	-.02	-.01	.21[ce]	.18[e]
Oppos-Negot X Frequent	-.08	-.13	-.02	-.05
		Total Behavior Problems		
Step 1: Baseline: Change in R^2 (see Table 3 for details)	.67***	.18**	.58***	.20**
Step 2: Main effects: Change in R^2	.00	.00	.01	.01
Prop(Oppos – Negot&Whine)	-.03	-.02	-.04	-.02
Reasoning (zero use as reference category)				
Used from 1 to 24%	-.03	-.05	.08	.14
Used > 24%	-.04	-.07	-.04	.03
Step 3 : Interactions: Change in R^2	.02[c]	.04[d]	.03[d]	.07*
Oppos-Negot X Moderate	-.00	.01	.14[e]	.19[e]
Oppos-Negot X Frequent	-.16[de]	-.27[de]	-.14[ce]	-.22[ce]
		Effortful Control		
Step 1: Baseline: Change in R^2 (see Table 3 for details)	.58***	.04	.46***	.11*

Step 2: Main effects: Change in R^2	.00	.01	.02	.02
Prop(Oppos – Negot&Whine)	-.05	-.07	-.12	-.14
Reasoning (zero use as reference category)				
Used from 1 to 24%	.02	.04	-.03	-.04
Used > 24%	.03	.03	.06	-.03
Step 3 : Interactions: Change in R^2	.01	.03	.01	.01
Oppos-Negot X Moderate	-.19[d]	-.27[c]	.05	.15
Oppos-Negot X Frequent	-.12	-.18	.13	.16

[a]Predicting behavior problems or effortful control at Wave 2 or 3 controlling for Wave-1 scores on the same variable.
[b]Predicting simple gains in behavior problems or effortful control from Wave 1 to either Wave 2 or to Wave 2.
[c]$p < .10$, 1-tailed (reported for interactions only). [d]$p < .10$, 2-tailed ($p < .05$, 1-tailed). [e]This interaction would be "more" significant if entered by itself, without the other interaction, e.g., , $p < .05$ instead of $p < .10$.
*$p < .05$. ** $p < .01$. *** $p < .001$, all 2-tailed tests.

Table 6

Standardized Regression Coefficients Predicting Gains in CBCL Broadband Scales and Effortful Control from Intermediate and Frequent Ignoring by Predominant Noncompliance Type, Controlling for Selected Covariates

	W2|W1[a]	W2-W1[b]	W3|W1[a]	W3-W1[b]
		Externalizing Problems		
Step 1: Baseline: Change in R^2 (see Table 3 for details)	.67***	.18**	.58***	.20**
Step 2: Main effects: Change in R^2	.02	.02	.00	.01
Prop(Oppos – Negot&Whine)	-.00	-.01	-.02	-.06
Ignoring (zero use as reference category)				
Used from 1 to 19%	.09	.14	.04	.07
Used > 19%	.02	.01	.06	.02
Step 3: Interactions: Change in R^2	.04[c]	.01	.00	.00
Oppos-Negot X Moderate	.01	.02	.02	-.02
Oppos-Negot X Frequent	-.16[de]	-.26[de]	-.06	-.07
		Internalizing Problems		
Step 1: Baseline: Change in R^2 (see Table 3 for details)	.58***	.20***	.37***	.13**
Step 2: Main effects: Change in R^2	.01	.01	.01	.04
Prop(Oppos – Negot&Whine)	-.06	-.05	-.05	.01
Ignoring (zero use as reference category)				
Used from 1 to 19%	-.06	-.12	.07	.01
Used > 19%	-.03	-.06	-.06	-.19[d]
Step 3: Interactions: Change in R^2	.01	.02	.01	.02
Oppos-Negot X Moderate	.10[e]	.15[e]	-.10	-.14
Oppos-Negot X Frequent	-.07	-.10	-.15	-.18
		Total Behavior Problems		
Step 1: Baseline: Change in R^2 (see Table 3 for details)	.67***	.18**	.58***	.20**
Step 2: Main effects: Change in R^2	.00	.00	.01	.01
Prop(Oppos – Negot&Whine)	-.02	-.01	-.03	-.02
Ignoring (zero use as reference category)				
Used from 1 to 19%	.03	.02	.07	.07
Used > 19%	.00	-.01	.01	-.07
Step 3 : Interactions: Change in R^2	.02[d]	.06*	.01	.02
Oppos-Negot X Moderate	.06[e]	.12[e]	-.03	-.05
Oppos-Negot X Frequent	-.17*	-.27*	-.13[e]	-.20[c]
		Effortful Control		
Step 1: Baseline: Change in R^2	.58***	.04	.46***	.11*

Step 2: Main effects: Change in R^2	.01	.03	.02	.02
Prop(Oppos – Negot&Whine)	-.05	-.07	-.13	-.14
Ignoring (zero use as reference category)				
Used from 1 to 19%	-.07	-.08	-.05	.01
Used > 19%	-.00	.02	-.02	.05
Step 3: Interactions: Change in R^2	.04*	.06[d]	.03[d]	.04[c]
Oppos-Negot X Moderate	-.13[ce]	-.17[e]	-.14[e]	-.07
Oppos-Negot X Frequent	.14[ce]	.21[ce]	.16[ce]	.24[d]

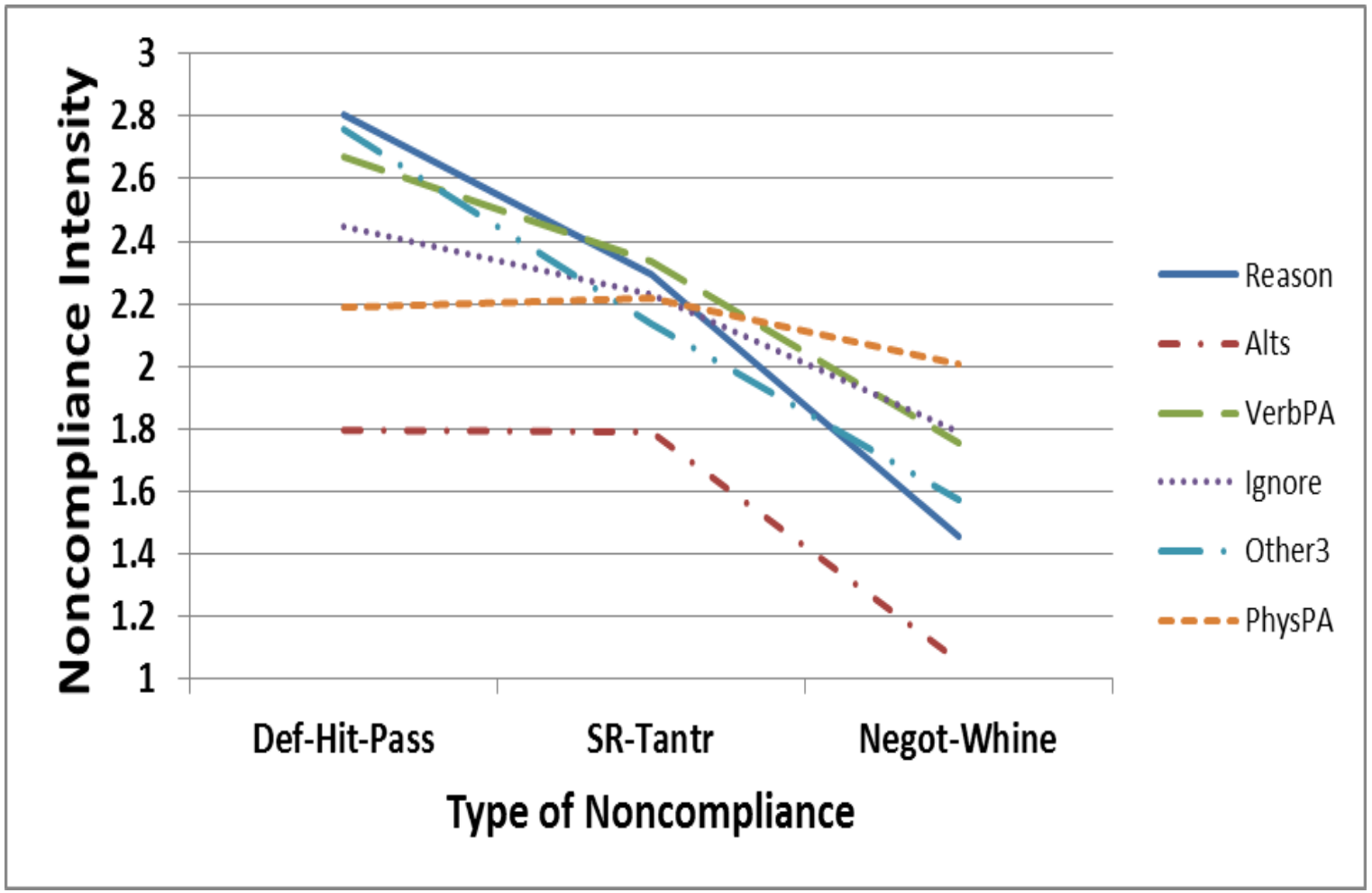
[a]Predicting behavior problems or effortful control at Wave 2 or 3 controlling for Wave-1 scores on the same variable.
[b]Predicting simple gains in behavior problems or effortful control from Wave 1 to either Wave 2 or to Wave 2.
[c]$p < .10$, 1-tailed (reported for interactions only). [d]$p < .10$, 2-tailed ($p < .05$, 1-tailed). [e]This interaction would be "more" significant if entered by itself, without the other interaction, e.g., $p < .05$ instead of $p < .10$.
*$p < .05$. ** $p < .01$. *** $p < .001$, all 2-tailed tests.

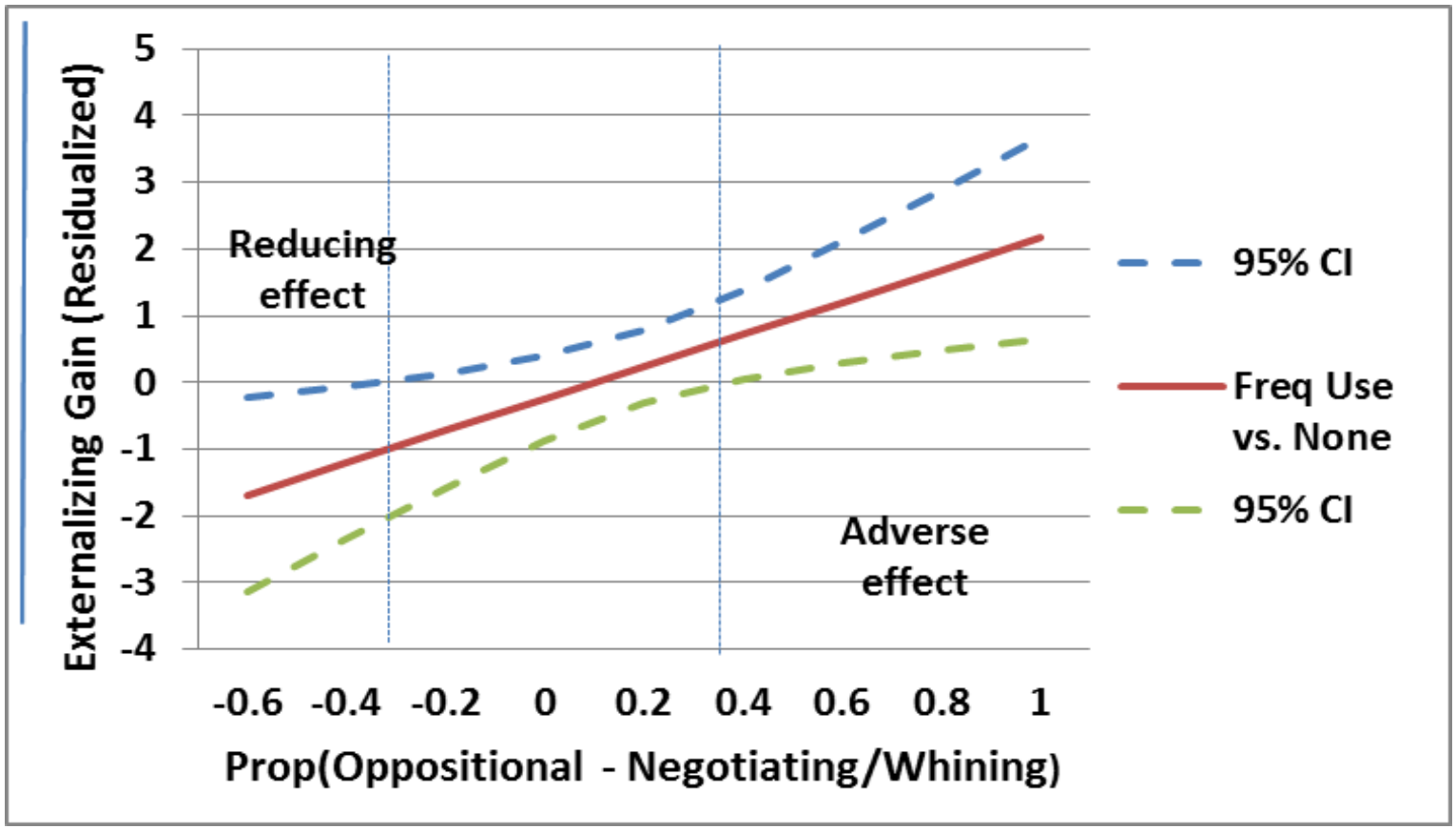
Figure Titles

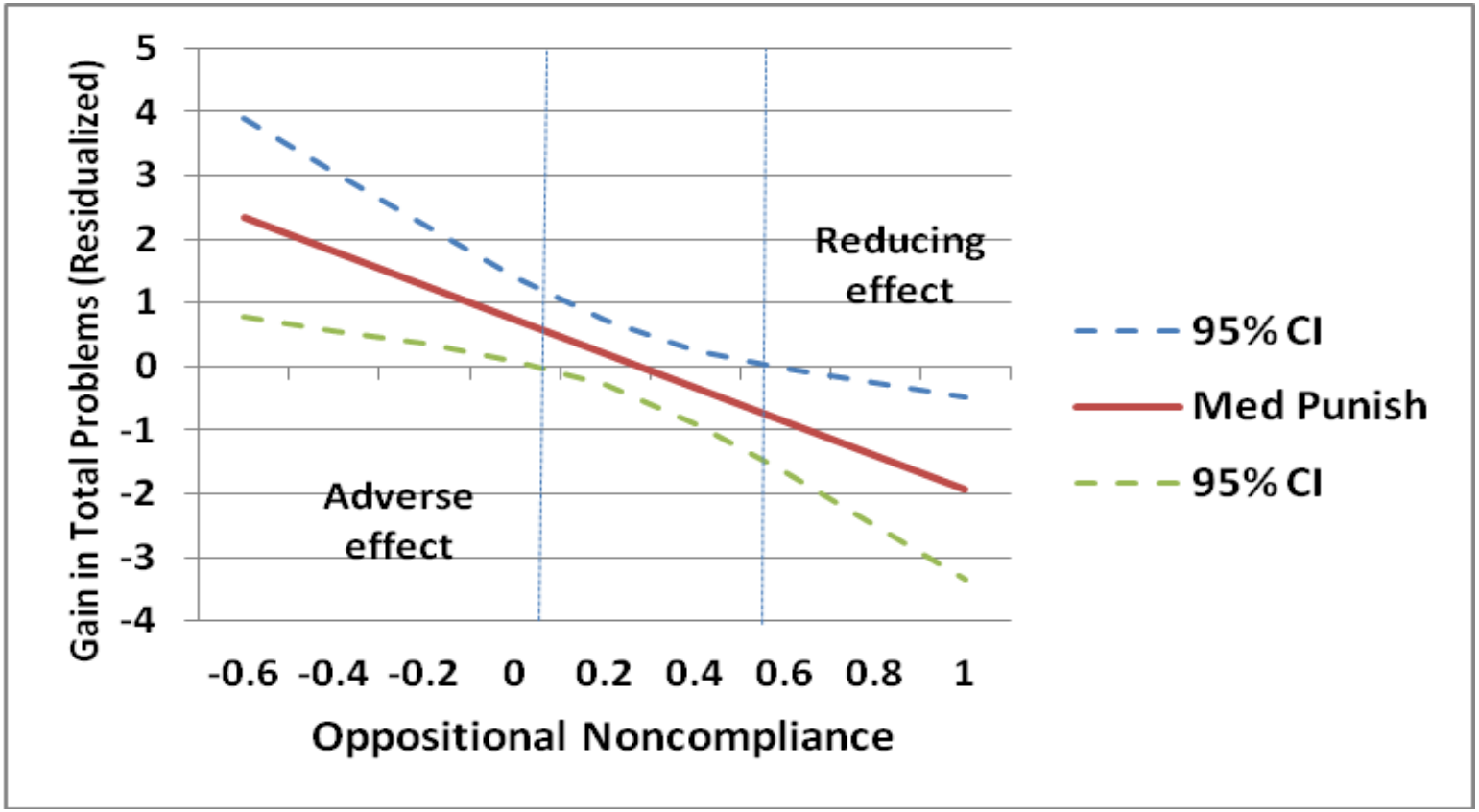
Figure 1. Noncompliance intensity immediately following six disciplinary responses to three types of noncompliance.

Figure 2. The slope coefficient for offering alternatives frequently (> 29% of disciplinary turns) on residualized gains in externalizing behavior problems over the next two months compared to never offering alternatives by the proportional occurrence of oppositional vs. parent-oriented noncompliance.

Figure 3. The slope coefficient for moderate use of warnings and punishments (< 16% of turns) on residualized gains in total behavior problems over the next two months compared to never using those tactics by the proportional occurrence of oppositional vs. parent-oriented noncompliance


Noncompliance Intensity
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
Def-Hit-Pass
SR-Tantr
Negot-Whine
Type of Noncompliance
Reason
Alts
VerbPA
Ignore
Other3
PhysPA


Externalizing Gain (Residualized)
5
4
3
2
1
0
-1
-2
-3
-4
Reducing effect
Adverse effect
95% CI
Freq Use vs. None
95% CI
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
Prop(Oppositional - Negotiating/Whining)


Gain in Total Problems (Residualized)
5
4
3
2
1
0
-1
-2
-3
-4
Reducing effect
Adverse effect
-0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1
Oppositional Noncompliance
95% CI
Med Punish
95% CI