



Contents lists available at ScienceDirect

Child Abuse & Neglect



Comparing physically abusive, neglectful, and non-maltreating parents during interactions with their children: A meta-analysis of observational studies

Steven R. Wilson^{a,*}, Jessica J. Rack^a, Xiaowei Shi^b, Alda M. Norris^a

^a Department of Communication, Purdue University, 100 North University Street, West Lafayette, IN 47907-2098, USA

^b Department of Psychology, Iowa State University, Ames, IA, USA

ARTICLE INFO

Article history:

Received 21 March 2007

Received in revised form 20 December 2007

Accepted 31 January 2008

Available online xxx

Keywords:

Parent–child interaction

Observational research

Meta-analysis

Aversive behavior

Positive behavior

Involvement

ABSTRACT

Objective: To clarify the nature and extent of differences in the ways that physically abusive, neglectful, and non-maltreating parents communicate during interactions with their children.

Method: A meta-analysis was conducted of 33 observational studies comparing parent–child interactions in families where parents have a documented history of physical abuse or neglect vs. where parents have no history of child maltreatment. Parental behaviors were grouped into three clusters (positivity, aversiveness, and involvement) for comparison across studies.

Results: When comparing maltreating (physically abusive or neglectful) vs. non-maltreating parents, mean weighted effect sizes for the three behavioral clusters range from $d = .46$ to $.62$. Physically abusive parents are distinguished from non-maltreating parents more so than neglectful parents in terms of aversive behavior, whereas the reverse is true for involvement. Publication date, parent and child age, and task structure moderate the magnitude, though not direction, of differences.

Conclusion: Parents with a documented history of child physical abuse or child neglect also are distinguished from non-maltreating parents by the levels of aversiveness, positivity, and involvement they display during interactions that constitute the parent–child relationship.

Practice implications: Researchers and practitioners need to carefully consider sample size, length and setting of observation, and interaction tasks when using observational methods.

© 2008 Elsevier Ltd. All rights reserved.

Introduction

Child abuse and neglect are not isolated incidents, but rather occur as part of ongoing patterns of interaction that constitute the parent–child relationship (Urquiza & Timmer, 2002). Because parent–child interaction is part of the complex etiology of child maltreatment (Belsky, 1993), observational studies have the potential to clarify risk factors and inform intervention and prevention efforts. Observational methods provide important data about parenting in maltreating samples since reports from maltreating parents do not always correspond with ratings from independent observers (Lau, Valerie, McCarty, & Weitz, 2006; Reid, Kavanagh, & Baldwin, 1987) and can be influenced by social desirability bias (Bennett, Sullivan, & Lewis, 2006). For these reasons, more than 30 observational studies have compared maltreating and non-maltreating parents during interactions with their children. Drawing conclusions from this literature currently is difficult, however, as many studies are under-powered and a host of potential moderating conditions vary across studies. This manuscript uses meta-analysis to

* Corresponding author.

clarify the nature and extent of differences in how physically abusive, neglectful, and non-maltreating parents communicate with their children.

Several theories regarding the etiology of child maltreatment implicate parent–child interaction. According to social-interactive views (Reid, 1986; Urquiza & Timmer, 2002), child physical abuse typically arises out of interactions in which parents respond to perceived child misbehavior. Abusive parents are thought to be inconsistent and ineffective in gaining their children's compliance; hence they resort to aversive behaviors (e.g., yelling, physical negative touch) more quickly and reciprocate their child's aversive responses more than non-maltreating parents. Information-processing models (Bugental et al., 2002; Milner, 2003), which highlight perceptual biases that typify maltreating parents (e.g., over-attributing hostile intent to even very young children), also highlight aversive parental behavior.

Attachment theory suggests that physically abusive and neglectful parents will display less positive or responsive behavior compared to non-maltreating parents (Crittenden & Ainsworth, 1989; Morton & Browne, 1998). Given its focus on how working models of self and other influence subsequent relationships, the theory suggests that individuals who experienced maladaptive histories as children are at risk for carrying insecure patterns of attachment into adulthood. Due to attachment-related dynamics (e.g., ambivalence about relationships), maltreating parents would be less likely to display sensitivity (e.g., accessibility, responsiveness) when interacting with their own children. Although these theories suggest that maltreating parents will enact different behaviors than non-maltreating parents, observational studies offer only mixed support. In a narrative literature review, Cerezo (1997) concluded that "When studies about abusive family interaction have included parental behavior, the expected differences have not always been found" (p. 222).

Some observational studies do find statistically significant differences between maltreating and non-maltreating parents on a broad range of behaviors (e.g., Alessandri, 1992; Bousha & Twentyman, 1984; Lau et al., 2006; Oldershaw, Walters, & Hall, 1986; Oldershaw, Walters, & Hall, 1989). For example, Bousha and Twentyman compared $n = 12$ abusive, neglectful, and non-maltreating mothers during home visits which, on average, lasted 4.5 h per family. Frequency of maternal behaviors were recorded for 10 different categories including displays of positive affect, verbal and physical aggression, and initiating topics of conversation. Statistically significant differences were detected between physically abusive vs. non-maltreating mothers in terms of the frequencies with which they enacted 9 of the 10 behaviors; neglectful and non-maltreating mothers differed significantly on 7 behaviors.

Other observational studies, however, have detected significant differences between maltreating vs. non-maltreating parents only for some types of behavior (e.g., Kavanagh, Youngblade, Reid, & Fagot, 1988; Lahey, Conger, Atekson, & Treiber, 1984; Lyons-Ruth, Connell, Zoll, & Stahl, 1987). Kavanagh et al. compared 22 physically abusive vs. 23 non-maltreating mothers as they completed a 12-min playtime and cleanup with their child. The authors detected significant differences between the two groups for positive behaviors but not for aversive behaviors. A third group of studies report few or no significant differences between maltreating and non-maltreating parents (e.g., Bennett et al., 2006; Koenig, Cicchetti, & Rogosch, 2000; Schindler & Arkowitz, 1986). Koenig et al. (2000) compared the control strategies used by 20 physically abusive, 26 neglectful, and 43 comparison mothers during a cleanup period following a free-play session in their lab. The three groups did not differ significantly in their use of any of the maternal control strategies.

Statistical power and moderating factors offer possible explanations for these inconsistent findings. Observational studies of parent–child interaction are labor intensive; hence, sample sizes often are small. For example, Kavanagh et al.'s (1988) comparison of physically abusive vs. non-maltreating mothers involved a sample of $N = 43$, which is by no means atypical. If one assumes that the actual behavioral differences between maltreating vs. non-maltreating parents are what Cohen (1988) would describe as "medium" (i.e., $d = .50$), then statistical power to detect a significant effect was only .51. By capitalizing on much larger combined sample sizes, meta-analysis should help clarify the actual magnitude of differences that exist between abusive, neglectful, and non-maltreating parents.

A second possible explanation is that study, parent, or child characteristics qualify general conclusions about parental behavior. For example, maltreating parents may try to enact what they believe to be socially appropriate behavior when they are being observed, and hence differences between maltreating and non-maltreating parents may not be observed when the observation period is short (Bennett et al., 2006). Aside from observation length, factors such as maltreatment type (abuse vs. neglect), parent sex (mothers vs. fathers), parent race/ethnicity, child age and sex, type of comparison group (community, clinical), observation setting (home vs. lab) and the structure of tasks completed also vary across studies. By assessing whether moderating factors are associated with obtained effect sizes, meta-analysis offers an empirical test of whether general conclusions about behavioral differences between abusive, neglectful, and non-maltreating parents are warranted (Hunter & Schmidt, 2004; Rosenthal, 1991).

Although useful, meta-analyzing this literature presents its own challenges. One major obstacle is that no single coding system has been used to analyze parental behaviors; indeed, there are nearly as many coding systems as there are observational studies. Coding systems vary in terms of what parenting behaviors are scored; for example, some studies code "praise" whereas others do not. Even when the "same" coding system is employed (e.g., the Dyadic Parent–Child Interaction Coding System; Eyeberg et al., 2005), researchers vary in whether they code all or only some of the behaviors in the system as well as whether they report frequencies for individual behaviors or for groups of behaviors.

Meta-analysts could deal with this situation by focusing on one concrete behavior present in many coding systems. For example, Wilson, Shi, Tirmenstein, Norris, and Rack (2006) meta-analyzed studies comparing maltreating and non-maltreating parents' use of physical negative touch (PNT; e.g., slapping, spanking, or taking a toy roughly from a child). Although PNT is part of many coding systems, the authors could analyze only 12 effect sizes from 8 studies because some

studies did not code PNT, whereas other studies coded PNT but then collapsed it along with other behaviors (e.g., verbal aggression, yelling) into a larger cluster of “aversive” behavior without reporting separate data on PNT.

The present paper compares maltreating and non-maltreating parents in terms of three broad clusters of behavior: aversiveness, positivity, and involvement. These clusters can be gleaned from virtually any coding system, and existing theories would lead us to expect differences in them. The first cluster, *aversive behavior*, involves parental actions that communicate negative affect (e.g., anger, disapproval) and thus are likely experienced as unpleasant by the child. Examples of specific behavioral indicators drawn from the literature include physical negative touch (Burgess & Conger, 1978) disapproval, humiliation, and threats (Oldershaw et al., 1989), and spurning and terrorizing (Brassard, Hart, & Hardy, 1993). The second cluster, *positive behavior*, includes actions displaying positive affect, thus demonstrating liking, approval, and/or support. Examples of specific parental positive behaviors include approval, laughter, and positive physical touch (Oldershaw et al., 1989), encouragement and positive affect (Wasserman, Green, & Allen, 1983), and verbal praise (Timmer, Borrego, & Urquiza, 2002). The third cluster is composed of behavior signaling *involvement*, defined as the degree to which a parent enacts behaviors that indicate responsiveness, cooperation, or interest in the child or the child’s agenda (Burgoon & Newton, 1991; Cegala, Savage, Brunner, & Conrad, 1982). Parental behaviors that signal high involvement include questions (Schindler & Arkowitz, 1986), eye contact (Schaeffer, 1983), and positive responsiveness (Fagan & Dore, 1993); those that signal low parental involvement include ignoring (Wasserman et al., 1983), detachment or disengagement (Haskett, Ahern, Ward, & Allaire, 2006), and non-attending (Kavanagh et al., 1988).

Research hypotheses and questions

Social-interactional and information-processing models suggest that maltreating parents should be distinguishable from comparison parents in terms of aversive behavior; hence, it is predicted that *parents with a history of child maltreatment (physical abuse or neglect) will display more aversive behavior than non-maltreating parents when interacting with their children* (Hypothesis 1). Attachment theory suggests that maltreating parents should fail to display affection and responsiveness, and thus it is predicted that *parents with a history of child maltreatment will display less positive behavior and involvement than non-maltreating parents when interacting with their children* (Hypothesis 2). Two research questions also are posed. RQ1: Which behavioral clusters (aversiveness, positivity, and/or involvement) most accurately distinguish physically abusive from non-maltreating parents? What about neglectful from non-maltreating parents? RQ2: Do differences between maltreating and non-maltreating parents vary depending on family or study characteristics?

Methods

This section describes three phases of the meta-analysis: (a) searching the literature based on inclusion criteria, (b) deriving effect size estimates and characteristics from included studies, and (c) calculating cumulative measures of effect size and testing for moderators.

Inclusion criteria and literature search

Given the research questions, studies were included only if they had the following:

1. A sample of parents with a documented history of child maltreatment. Parents in the maltreatment group: (a) were referred to the researchers directly by Child Protective Services (CPS) agencies (or similar child welfare agencies in non-US studies), (b) were identified by the researchers as maltreating based on a check of official CPS records, or (c) self-reported CPS involvement. Studies comparing parents who scored high vs. low on the Child Abuse Potential Inventory (CAPI) (e.g., Dolz, Cerezo, & Milner, 1997; Wilson, Morgan, Hayes, Bylund, & Herman, 2004) were not included.
2. A comparison sample of parents with no history of child maltreatment, matched with the maltreating sample on socio-demographic factors (e.g., parent education, child age). Thus, studies that coded parent–child interaction before and after treatment but with no comparison group (e.g., Wolfe & Sandler, 1981) were not included. Many studies did not explain the exact procedures used to verify that comparison parents had no history of child maltreatment, but those that did checked official CPS records, interviewed parents, and/or administered the CAPI. A few studies (e.g., Lahey et al., 1984; Reid, Taplin, & Lorber, 1981) reported data on two comparison groups (one matched to the maltreatment group, another not matched), in which case only data from the matched comparison group were used.
3. Observational data on parent–child interaction. For example, studies analyzing self-reported parenting behaviors using the Conflict Tactics Scales (Straus, Hamby, Finkelhor, Moore, & Runyan, 1998), diaries (e.g., Trickett & Kuczynski, 1986), or daily telephone reports (e.g., Whipple & Wilson, 1996) were not included. No observational study of parents with a documented history of child sexual abuse was identified; hence, this form of maltreatment was not included in the meta-analysis.
4. Coding or rating and analysis of parental behavior relevant to at least one of the three clusters (aversiveness, positivity, involvement). Hence, studies that observed parent–child interactions but coded only children’s behavior (e.g., Reid et al., 1987) or analyzed only children’s behavior without comparing parents (e.g., Crittenden, 1992) were not included in this report. Howes, Cicchetti, Toth, and Rogosch (2000) was excluded because coding was done with the entire “family” as the

unit of analysis. Crittenden (1982) was excluded because mothers' predominate interaction style was classified in a way that combined behaviors from different clusters (e.g., an "abusive" style included both aversive and inappropriate positive behaviors).

5. Sufficient information from which to derive effect size estimates (see below).

Several search strategies were employed to identify studies meeting these inclusion criteria. First, online databases (Communication Abstracts, Dissertation Abstracts, Family & Society Studies Worldwide, Linguistics and Language Behavior Abstracts, PsycINFO, Social Sciences Abstracts, Women's Studies International) were searched by using all combinations of either child abuse, child neglect, or child maltreatment along with either communication, observation, interaction, or parenting (e.g., child abuse AND communication, child abuse AND observation). Searches included all possible entries (e.g., journal articles, book chapters, dissertations) up through 2006. Second, when studies meeting the inclusion criteria were identified, their reference lists were checked for citations to earlier studies. Reference lists from narrative reviews (e.g., Cerezo, 1997) also were checked. Third, all subsequent citations to two early studies (Bousha & Twentymann, 1984; Burgess & Conger, 1978) were identified using the Social Science Citation Index and checked for relevance. Fourth, the table of contents and abstracts for Child Abuse & Neglect were searched from 1994–2006 to ensure relevant studies were not missed. Finally, scholars who had published extensively on child maltreatment and family interaction were contacted, sent the list of studies found to date, and asked to identify relevant studies—including unpublished data—that might have been overlooked. A total of 33 studies meeting these inclusion criteria were identified and included in the meta-analysis (see Table 1).

Deriving effect size estimates and coding study characteristics

Effect size estimates. The coding scheme from each of the 33 included studies was assessed to determine which categories should be used to derive effect size estimates for aversiveness, positivity, and involvement. Table 2 presents a complete list of behaviors from each of the 33 studies that fall into these three clusters. Not every study coded behaviors relevant to all three clusters, but each reports data relevant to at least one cluster. When a study reported data on more than one behavior relevant to the same cluster, separate effect sizes were computed for each behavior and then averaged. Effect size estimates always were computed such that the value of d would be positive if the mean scores for maltreating (abusive or neglectful) vs. non-maltreating parents were in the expected direction, and hence negative if the group means were opposite of the expected direction.

Several strategies were used to estimate effect sizes when studies failed to report information needed to calculate d directly. When a study reported the exact value of an inferential statistic, that statistic was converted to d using procedures described in Hunter and Schmidt (2004) and Lipsey and Wilson (2001). If a study described a comparison as "statistically significant" at the $p < .05$ or $p < .01$ level without reporting the exact value of an inferential test, the critical value for the test was identified and converted to d . Many studies failed to report inferential tests or descriptive statistics when comparisons were not statistically significant. For cases in which the direction of difference was known, two separate estimates were calculated: (a) a *maximum* estimate, obtained by identifying the critical value for the test at $p < .05$ and (b) a *minimum* estimate, which treated the difference between groups as zero. Because the true effect size would fall between these two estimates, the overall meta-analyses were computed twice (once using minimum and once using maximum estimates). As readers will see, the mean weighted effect sizes across studies are similar regardless of which procedure was used, and hence only *minimum* estimates are reported for moderator analyses. Finally, in cases where even the direction of difference was not known, the authors were contacted to see if they could provide the relevant descriptive statistics. If a study was based on a dissertation, the dissertation report was searched for relevant information. Because an effect size could not be derived using any of these methods, the following studies were excluded: Camras et al. (1990), Cerezo, D'Ocon, and Dolz (1996), Coster, Gersten, Beeghly, and Cicchetti (1989), DiLalla and Crittenden (1990), Herrenkohl, Herrenkohl, Toedter, and Yanushefski (1984), and Nastasi and Hill (1982).

In most cases, the "study" was treated as the unit of analysis and hence only one effect size per study is reported for each of the three behavioral clusters (aversiveness, positivity, and involvement). There were two instances where a single study provided multiple effect size estimates for the same cluster. When a single study included data from abusive, neglectful, and non-maltreating parents, two effect size estimates were derived from that study: one comparing abusive vs. non-maltreating parents and a second comparing neglectful vs. non-maltreating parents. In addition, when a study reported data on fathers as well as mothers, separate estimates were derived for maltreating vs. non-maltreating fathers and for maltreating vs. non-maltreating mothers. Although some may object that the obtained effect sizes are not independent in these cases, we deemed it more important to retain data on both types of maltreatment (or on fathers as well as mothers) rather than discard one group. In addition, analyses use independent data once moderators analyses are conducted (e.g., when separate analyses of abusive vs. non-maltreating and neglectful vs. non-maltreating parents are conducted, data from studies with both abusive and neglectful groups then contribute only one effect size to each analysis).

Study characteristics. For each study, maltreating parents were classified into one of three types: physically abusive, neglectful, or mixed. "Physically abusive" means that parents were recruited into the study because they had a documented history of child physical abuse. Because CPS agencies may give greater priority to "active" forms of abuse (physical or sexual abuse) than to "passive" forms (neglect or emotional maltreatment, see Lau et al., 2005), it is possible some of these parents also

Table 1
 Demographic information for included studies.

Author/publication year	Study characteristics						Parent characteristics			Child characteristics	
	Groups studied ^a	N	Country	Setting	Obs. length (min)	Task structure ^b	Sex ^c	Age	% Minority (USA only) ^d	Age	% Boys
Alessandri (1992)	M vs. C	30	USA	Lab	40	MMT	F	27.91	87	4.42	53
Alessandri and Lewis (1996)	M vs. C	84	USA	Lab	25	HS	F	27.3	100	4.38	50
Aragona and Eyberg (1981)	N vs. BC	18	USA	Lab	10	MMT	F	27.5	0	5.58	50
Bennett et al. (2006)	A vs. C	100	USA	Lab	20	HS	F	30.8	81	5.03	55
	N vs. C	108	USA	Lab	20	HS	F	30.8	81	5.03	55
Borrego, Timmer, Urquiza, and Follett (2004)	A vs. BC	30	USA	Lab	15	MMT	F	29.6	30	4.15	60
Bousha and Twentyman (1984)	A vs. C	24	USA	Home	270	US	F	N/A	17	4.45	N/A
	N vs. C	24	USA	Home	270	US	F	N/A	21	4.25	N/A
Brassard et al. (1993)	M vs. C	49	USA	Home	15	HS	F	N/A	54	6.67	55
Burgess and Conger (1978)	A vs. C	36	USA	Home	240	HS	F	29.8	N/A	6.50	N/A
	A vs. C	36	USA	Home	240	HS	M	29.8	N/A	6.50	N/A
	N vs. C	36	USA	Home	240	HS	F	30.6	N/A	6.50	N/A
	N vs. C	36	USA	Home	240	HS	M	30.6	N/A	6.50	N/A
Cerezo and D'Ocon (1995)	M vs. C	30	Spain	Home	360	US	F	38.5	N/A	8.10	67
Cerezo and D'Ocon (1999)	M vs. C	50	Spain	Home	364	US	F	N/A	N/A	8.36	72
Edwards, Shipman, and Brown (2005)	N vs. C	48	USA	Home	12	MS	F	N/A	63	9.17	44
Fagan and Dore (1993)	N vs. C	27	USA	Home	30	US	F	26.7	72	2.29	55
Givens (1978)	A vs. C	14	USA	Lab	12.5 ^e	HS	F	N/A	N/A	2.25	N/A
Haskett et al. (2006)	A vs. C	185	USA	Lab	30	HS	Both	34.4	68	7.0	52
Kavanagh et al. (1988)	A vs. C	45	USA	Lab	12	MS	Both	N/A	0	3.11	N/A
Koenig et al. (2000)	A vs. C	63	USA	Lab	2.59	MS	F	N/A	19	3.61	52
	N vs. C	69	USA	Lab	2.59	MS	F	N/A	19	3.61	52
Lahey et al. (1984)	A vs. C	16	USA	Home	270	US	F	27	N/A	5.67	63
Lau et al. (2006)	A vs. BC	205	USA	Lab	13	MMT	F	N/A	57	11.46	59
Lorber, Felton, and Reid (1984)	A vs. BC	18	USA	Home	60	US	F	N/A	N/A	5.96	56
Lyons-Ruth et al. (1987)	M vs. C	38	USA	Home	40	US	F	25.5	18	1	54
Mash, Johnston, and Kovitz (1983)	A vs. C	36	USA	Lab	35	MMT	F	N/A	N/A	4.79	67
Oldershaw et al. (1986)	A vs. C	20	Canada	Treat	40	US	F	24.5	N/A	3	50
Oldershaw et al. (1989)	A vs. C	116	Canada	Treat	40	US	F	27	N/A	3.75	48
Reid et al. (1981)	A vs. BC	87	USA	Home	60	US	F	N/A	N/A	8.3	89
	A vs. BC	62	USA	Home	60	US	M	N/A	N/A	8.3	89
Schaeffer (1983)	M vs. C	26	USA	Daycare	12.5 ^e	US	F	N/A	N/A	2.8	50
Schindler and Arkowitz (1986)	A vs. C	13	USA	Lab	35	MMT	F	29.5	22	4.8	91
Shipman and Zeman (1999)	A vs. C	44	USA	Home	12	MS	F	N/A	20	9.0	68
Silber, Bermann, Henderson, & Lehman (1993)	A vs. C	32	USA	Lab	10	HS	F	N/A	0	10.5	12
	A vs. C	32	USA	Lab	10	HS	M	N/A	0	10.5	12
Timmer et al. (2002)	A vs. BC	30	USA	Lab	15	MMT	F	30	N/A	4.2	63
Valentino, Cicchetti, Toth, and Rogosch (2006)	A vs. C	83	USA	Lab	25	MMT	F	26	27	1.5	44
	N vs. C	99	USA	Lab	25	MMT	F	26.5	23	1.5	44
Wasserman et al. (1983)	A vs. C	24	USA	Lab	5	US	F	N/A	N/A	1.2	N/A
Webster-Stratton (1985)	A vs. BC	40	USA	Home	60	US	F	30	N/A	4.8	73
Whipple and Webster-Stratton (1991)	A vs. BC	121	USA	Home	30	US	F	32.5	04	4.9	69
	A vs. BC	109	USA	Home	30	US	M	35.2	04	4.9	69

^a A = physically abusive parents, N = neglectful parents, M = mixed maltreatment history. C = comparison parents (community), BC = comparison parents (clinical).

^b HS = highly structured; MS = moderately structured; US = unstructured. MMT = study employed multiple tasks with different levels of task structure.

^c F = female (mothers); M = male (Fathers), both = study gathered data from mothers and fathers, but results are combined together.

^d Because most studies reported data on race/ethnicity for parents, we place that data here. Some studies, however, reported data on race/ethnicity for children or families; those data are shown here too.

^e Study observation length was estimated from study procedures.

Table 2

Behaviors included in the four behavioral clusters and effect sizes for each cluster from each study for parents.

Study	Groups compared ^a	Aversive behaviors	ES	Positive behaviors	ES	Involvement	ES
Alessandri (1992)	M vs. C	Negative affective tone	3.16	Positive affective tone	2.59	High involvement Questions/instruction Demonstrate	2.16
Alessandri and Lewis (1996)	M vs. C	Negative behaviors	2.13	Positive behaviors	1.22		N/A
Aragona and Eyberg (1981)	N vs. BP	Critical statement	-.25	Praise	.42	Describe verbal acknowledge	.80
Bennett et al. (2006)	A vs. C N vs. C	Negative parenting	0 .12	Positive parenting	-.23 -.12		N/A N/A
Borrego et al. (2004)	A vs. BC	Negative behavior	.38	Verbal praise positive behavior	1.08		N/A
Bousha and Twentyman (1984)	A vs. C N vs. C	Verbal aggression Vocal negative Physical aggression	1.66 .37	Verbal affect Nonverbal affect	2.17 2.19	Initiation Social interaction	3.35 4.42
Brassard et al. (1993)	M vs. C	Spurning/verbal Aggression Terrorizing	.73	Mutual pleasure Positive touch Supportive presence	.79	Deny emotional Responsiveness (-) Body harmonics	.45
Burgess and Conger (1978)	A vs. C Moms A vs. C Dads N vs. C Moms N vs. C Dads	Negative behavior	.73 0 or -.72 .73 0 or .72	Positive behavior Parental compliance	.88 0 or .72 .88	Verbal rate Physical behavior rate	.37 or .73 N/A .37 or .73 N/A
Cerezo and D'Ocon (1995)	M vs. C	Negative interaction	0 or .76	Positive interaction	0 or .76	Neutral social approach	0 or .76
Cerezo and D'Ocon (1999)	M vs. C	Mother aversive	.36	N/A		N/A	
Edwards et al. (2005)	N vs. C		N/A		N/A	Emotional understanding	2.13
Fagan and Dore (1993)	N vs. C		N/A		N/A	Positive responsiveness	.72
Givens (1978)	A vs. C	Rough instrumental Touch Grab away Overloud, overlow voice Hand on Hips/Chin/Point Retract Lips	1.53	Oversoft voice Overhigh pitch Shoulder movement Lateral head tilt Smile	1.70		N/A
Haskett et al. (2006)	A vs. C	Negative regard Intrusiveness	.22	Positive regard	.40	Detached/disengaged (-)	.24
Kavanagh et al. (1988)	A vs. C	N/A		Positive parenting	.68	Nonattend/onlooker (-)	0 or .60
Koenig et al. (2000)	A vs. C N vs. C	Negative evaluation Physical negative touch Slightly negative affection Irritable/anger	.22 .13	Positive evaluation Positive touch Tender/affection Joy	.35 .21	Empathy	-.40 0
Lahey et al. (1984)	A vs. C	Verbal negative Physical negative	.84	Verbal positive Physical positive Positive affect	.42		N/A
Lau et al. (2006)	A vs. BC	Hostility/emotional control	.15	Support/happiness	.19		N/A
Lorber et al. (1984)	A vs. BC	Total aversive behavior	1.74		N/A		N/A
Lyons-Ruth et al. (1987)	M vs. C	Hostile/intrusive	.90		N/A	Low involvement (-)	.26

Study	Groups compared ^a	Aversive behaviors	ES	Positive behaviors	ES	Involvement	ES
Mash et al. (1983)	A vs. C	N/A		Praise	0 or .69	Question Interaction No response (-)	0 or .23
Oldershaw et al. (1986)	A vs. C	Threat Physical negative touch Humiliation Disapproval Negative demand Negative affect	1.23	Approval Laugh Positive physical Positive affect (initial) Positive affect (follow-up)	1.39	Cooperation Modeling	.76
Oldershaw et al. (1989)	A vs. C	Disapproval Humiliation Negative physical Threat % Denied requests % Negative affect (initial) % Negative affect (follow-up)	.81	Approval Positive physical Laugh % Positive affect	1.22	Ignore (-) Cooperation	1.19
Reid et al. (1981)	A vs. BC Moms A vs. BC Dads	Total aversive behavior Negative physical behavior Command negative	.68 .40		N/A N/A		N/A N/A
Schaeffer (1983)	M vs. C	Negative behavior	1.26	Positive behavior	.41	Caretake Verbalize Physical involvement Look at Greet/farewell	.45
Schindler and Arkowitz (1986)	A vs. C	Mother disapproval Threat Noncomply	.22	Mother approval	1.10	Mother questions	1.33
Shipman and Zeman (1999)	A vs. C		N/A		N/A	Emotional understanding	1.00
Silber et al. (1993)	A vs. C Moms A vs. C Dads	Criticism/protest	0 or .48 .49	Agreement	0 or -.48 .49	Information	0 or .48 0 or .48
Timmer et al. (2002)	A vs. BC	Negative behaviors	.36	Verbal praise Positive behavior	1.04	Descriptions Questions Total verbalization	.58
Valentino et al. (2006)	A vs. C N vs. C	Limit setting	-.37 -.28		N/A N/A	Point/reposition Demonstration	.09 .09
Wasserman et al. (1983)	A vs. C	Negatives Distraction Negative affect	.52 or .95	Encouragement Positive affect	.60 or 1.04	Attention management Verbal teaching Physical teaching Initiating response Ignoring (-)	.23 or .58
Webster-Stratton (1985)	A vs. BC	Physical negative Criticism	.92	Praise Physical positive	-.28	Descriptive/reflective comments	-.53
Whipple and Webster-Stratton (1991)	A vs. BC Moms A vs. BC Dads	Spanking Criticism Physical negative touch	.48 .59		N/A N/A		N/A N/A

Please cite this article in press as: Wilson, S. R., et al. Comparing physically abusive, neglectful, and non-maltreating parents during interactions with their children: A meta-analysis of observational studies. *Child Abuse Negl* (2008), doi:10.1016/j.chiabu.2008.01.003

Table 2 (Continued).

Meta-analysis summary	Aversive	Positive	Involvement
# effect sizes (<i>K</i>)	38	31	28
Total # parents (<i>N</i>)	2193	1607	1317
Mean weighted <i>d</i> (min/max): Ave (<i>d</i>)	.46/.48	.53/.57	.53/.62
Confidence interval (<i>CI</i> ₉₅) for Ave (<i>d</i>)	.27–.65/.28–.68	.32–.74/.36–.78	.21–.85/.31–.93
Observed variance in <i>d</i> values: Var (<i>d</i>)	.365/.378	.370/.369	.756/.713
Expected variance in <i>d</i> (sampling error): Var (<i>e</i>)	.07	.08	.09
% Obs. Var. due to sampling error: [Var (<i>e</i>)/Var (<i>d</i>)] × 100	19.8%/19.1%	21.9%/22.1%	11.5%/12.7%

Note: N/A = Not available. Studies with two effect sizes reflect cases in which not enough information was provided to calculate an exact effect size (e.g., a report stated that a mean difference was non-significant and gave means and *N* but not standard deviations nor an exact *p* value). Two procedures were used to estimate effect sizes in such cases. The first estimate (min) assumes the effect size was zero; the second estimate (max) reports the largest effect size that would have been non-significant.

^a A = physically abusive parents, N = neglectful parents, M = mixed maltreatment history. C = comparison parents (community), BC = comparison parents (clinical).

may have neglected their children, but their classification with the CPS agency was physical abuse. “Neglect” means that these parents all had a documented history of child neglect. “Mixed” means that the “maltreatment” group includes some parents with a documented history of child physical abuse but also other parents with a documented history of neglect (or in two cases, either child physical abuse or emotional maltreatment); hence, “mixed” indicates that parents were included based on varied but not necessarily multiple forms of maltreatment.

Aside from maltreatment type, the following information was recorded for each study: type of comparison group (matched community sample of non-maltreating parents, matched sample of non-maltreating parents whose children were receiving mental health counseling), method used to control for differences in socio-demographic variables (case–control design, statistical covariates), parent gender (mother or father), mean parent age, percentage of parents in the sample who were ethnic/racial minorities (for USA studies only), mean child age, % of children who were boys, sample size, publication date, country where the study was conducted, setting (lab, home, daycare), length of parent–child interaction observed (in minutes), and degree of structure in the interaction task (see Table 1).

To assess interaction task structure, the second and third authors independently classified each of the 33 studies using a four-category scheme: highly structured, moderately structured, unstructured, and multiple tasks with different levels of structure. In a highly structured task, the researcher explicitly tells participants what tasks to do, what order to do them in, what objects to use, how to use those objects, and what the end result of the activity should look like. In an unstructured task, the researcher does not provide rules about what task to perform, what order to do activities in, or what should be the result of the activity. The researcher may provide a standard set of toys or objects but does not dictate how they should be played with, in what order, or with what outcome. In moderately structured tasks, the researcher might tell the participants what to do but not how to do it (e.g., cleanup, play with blocks) or how to do it but not what to do (e.g., teach the child something, follow the child’s lead). Finally, some studies used multiple tasks with differing levels of structure (e.g., one highly structured task and one unstructured task). Inter-coder reliability was $Kappa = .80$; disagreements were resolved via discussion between these authors.

Statistical analyses of effect sizes

Analyses of parental aversiveness, positivity, and involvement took place in three steps. First, the mean effect size (weighted by sample size) and variance across studies was computed for each behavioral cluster using the Hunter and Schmidt (2004) Meta-Analysis Program. Second, the amount of observed variability in effect size estimates was compared to the amount of variability expected by chance. The program calculates an estimate of the amount effect sizes would vary due to sampling error. Hunter and Schmidt recommend testing for moderators when sampling error accounts for less than 75% of the observed variance in effect size estimates across studies (p. 401). Third, categorical (nominal-level) and continuous (interval-level) moderators were tested in cases where effect sizes were heterogeneous. Testing for a categorical moderator involves subdividing the effect size estimates based on a meaningful moderator (e.g., studies conducted in the home vs. the lab); calculating the category means, variances, and sampling errors separately for each category; and determining if the unexplained variance is accounted for. Testing for a continuous moderator (e.g., length of observation in minutes) can be done by correlating effect sizes with corresponding scores on the continuous variables, where n = the number of effect sizes.

One issue that had to be decided was whether to remove “outliers.” Because effect sizes are weighted by sample size (which gives less weight to a very large effect size when it occurs in a small sample study) and because there are a sizeable number (between 28 and 38) of effect sizes for each cluster when all studies are included (which reduces the impact of any one effect size), extreme effect sizes were not removed in meta-analyses involving all studies. When conducting moderator analyses, however, studies are broken into smaller subgroups, in which case one extreme effect size can have a substantial impact on the mean and observed variability of effect sizes for that group. In such cases, moderator analyses were conducted both with and without an extreme effect size value ($d > 2.0$) being included. Moderators are discussed only if they remained when extreme effect sizes were removed.

Results

Descriptive statistics for studies in meta-analysis

Tables 1 and 2 present 43 comparisons of maltreating vs. comparison parents from 33 studies. Across these 43 comparisons, the average sample size is $N = 56$. Most studies (27 of 33 studies, 82%) observed only mothers interacting with children, four (12%) conducted separate observations of mothers and fathers, and two (6%) combined observations of mothers or fathers into a single group. Slightly more than half of the studies (18 of 33, 55%) compared physically abusive vs. non-maltreating parents, 3 (9%) compared neglectful and non-maltreating parents, 5 (15%) compared three groups (abusive, neglectful, non-maltreating), and 7 (21%) compared a mixed maltreatment group vs. a non-maltreating group. Children’s mean age across studies varies from 1 to 11 years.

Regarding research design, studies are split fairly evenly between those that involve home observation (14 of 33, 42%) and those conducted in a laboratory or treatment facility (18 of 33, 55%); one study observed in a daycare. Most studies (29 of 33, 88%) use a case–control design, where non-maltreating parents as a group were matched with maltreating parents on socio-demographic characteristics (e.g., parent age/ethnicity/education; child age/sex; family size/income). Four studies

(12%) were not able to match maltreating and non-maltreating parents and hence included socio-demographic variables as covariates in comparisons of the two groups. The length of parent–child interaction actually observed varies dramatically across studies, from a low of 3 min to a high of 6 h per family. Finally, 7 of the 33 studies (21%) had parents and children complete highly structured tasks, 5 (15%) involved moderately structured tasks, 14 (42%) involved unstructured tasks, and 7 (21%) had parents and children complete multiple tasks with different levels of structure.

Research hypotheses 1 and 2

Hypothesis 1 predicted that parents with a history of child maltreatment (physical abuse, neglect, or a mixed group) would display more aversive behavior than parents with no maltreatment history while interacting with their children. The first column in Table 2 shows 38 effect sizes from 28 studies ($N=2193$ parents) that compare maltreating and non-maltreating parents on aversive behavior; summary statistics appear at the bottom of this table. The mean weighted d across the 38 effect sizes is .46 (minimum estimates) and .48 (maximum estimates). Consistent with Hypothesis 1, maltreating parents on average are about one-half a standard deviation higher than comparison parents in terms of rate of aversive behaviors. Observed variance in the 38 effect sizes is .37 for minimum estimates and .38 for maximum estimates. Variance due to chance is .07 in either case; hence, sampling error accounts for only about 20% of the observed variance in effect sizes, which suggests the presence of moderating factors. A significance test (Q) also was computed to assess homogeneity of effect sizes (see Hunter & Schmidt, 2004, p. 419). The test was significant both for minimum estimates, $Q(37)=192.06, p<.001$, and for maximum estimates, $Q(37)=198.51, p<.001$, reinforcing the high probability of moderating factors. Although the data are consistent with Hypothesis 1, the degree to which maltreating parents display more aversiveness than non-maltreating parents also appears to depend on other factors.

Hypothesis 2 predicted that non-maltreating parents would display more positivity and involvement than maltreating parents when interacting with their children. The second column in Table 2 reports 31 effect sizes from 24 studies ($N=1607$ parents) that compare maltreating vs. non-maltreating parents in terms of positive behavior. The mean weighted d across these 31 effect sizes is .53 (minimum estimates) and .57 (maximum estimates). Observed variance across these 31 effect sizes is .37 using either minimum estimates or maximum estimates. The amount of variance expected due to chance is .08 in either case; hence, sampling error accounts for less than 25% of the observed variance in effect sizes, suggesting the presence of moderator factors. Performing the homogeneity test on the 31 effect sizes resulted in $Q(30)=141.32, p<.001$ (minimum estimates), or $Q(30)=140.37, p<.001$ (maximum estimates). Although these findings support Hypothesis 2 with regard to positivity, the degree to which non-maltreating parents display more positivity than maltreating parents while interacting with their children depends on other factors.

The third column in Table 2 shows 28 effect sizes from 23 studies that compare maltreating and non-maltreating parents ($N=1317$) on behaviors signaling involvement. The mean weighted d across the 28 effect sizes is .53 (minimum estimates) and .62 (maximum estimates). Again consistent with Hypothesis 2, non-maltreating parents on average are between one-half and two-thirds of a standard deviation higher than maltreating parents in terms of displaying involvement during parent–child interactions. Zero does not fall within the CI_{95} for either minimum or maximum estimates, but these intervals are larger than those for parental aversiveness and positivity because: (a) observed variability in effect sizes is larger for involvement than for the other two clusters and (b) there are fewer studies in the involvement cluster. Observed variance in the 28 effect sizes is .75 for minimum estimates and .71 for maximum estimates. Variance due to chance is .09 in either case; hence, sampling error accounts for less than 15% of the observed variance in effect sizes, which suggests the presence of moderating factors. The significant test for homogeneity of effect sizes is significant both for minimum estimates, $Q(27)=228.46, p<.001$, and for maximum estimates, $Q(27)=212.98, p<.001$, again suggesting that moderating factors are present.

Research question 1: Maltreatment type

RQ1 asked whether some behavioral clusters do a better job of distinguishing physically abusive from non-maltreating parents whereas others better distinguish neglectful from non-maltreating parents. Because analyses using minimum and maximum estimates produced similar answers, only minimum estimates are reported in this and subsequent moderator analyses. Studies that involved a “mixed” maltreatment group (i.e., parents with a history of abuse or neglect were combined; see Table 1) are not included in this section.

Table 3 presents results from the subgroup analysis based on maltreatment type. Regarding aversiveness, the mean weighted effect size for 24 effects from studies comparing physically abusive vs. comparison parents ($d=.40$) is much larger than for 7 from studies comparing neglectful vs. comparison parents ($d=.06$). Sampling error accounts for more than 40% of observed variance in effect sizes for abusive parents and more than 90% for neglectful parents, suggesting that breaking the studies out by maltreatment type reduces within-group variance in effect sizes as compared to looking only at maltreating vs. non-maltreating parents. Maltreatment type has a strong impact on parental aversiveness with children, which indicates that aversive behavior is more useful in distinguishing physically abusive as opposed to neglectful parents from non-maltreating parents.

Regarding positivity, the mean weighted effect size for the 20 effects from studies comparing physically abusive vs. comparison parents is $d=.47$, whereas the mean for the 6 effect sizes from studies comparing neglectful vs. comparison

Table 3
 Moderator analysis by parent maltreatment type (minimum estimates).

Statistic	Aversiveness		Positivity		Involvement	
	A vs. C	N vs. C	A vs. C	N vs. C	A vs. C	N vs. C
Mean weighted (<i>d</i>): Ave (<i>d</i>)	.40	.06	.47	.40	.40	.80
Observed Variance: Var (<i>d</i>)	.1564	.0801	.2752	.3896	.4875	1.4656
Var. due to sampling error: Var (<i>e</i>)	.0663	.0728	.0761	.0854	.0827	.0956
% Obs. Var. due to sampling error (%)	42.5	90.9	27.7	21.9	19.3	6.5
# of parents (<i>N</i>)	1496	390	1097	291	651	321
# of effect sizes (<i>K</i>)	24	7	20	6	16	7

Note: A = physically abusive parents, N = neglectful parents, C = non-maltreating comparison parents. % Obs. Var. due to sampling error = [Var (*e*)/Var (*d*)] × 100.

parents is $d = .40$. Although the effect size is slightly larger for physically abusive parents, it is close to a medium effect for both groups. Equally important, sampling error still only accounts for about 25% of the observed variance in both sets of effect sizes (see Table 3), which is similar to the overall analysis of maltreating vs. non-maltreating parents reported above (see Table 2). Maltreatment type has little impact on rates of parental positivity during interactions with children, which suggests that positive behavior can be used to distinguish either physically abusive or neglectful parents from non-maltreating parents to a similar degree.

Regarding involvement, the mean effect size for the 7 effects from studies comparing neglectful vs. non-maltreating parents ($d = .80$) is twice as large as for the 16 effects from studies comparing physically abusive vs. non-maltreating parents ($d = .40$). Breaking studies out by maltreatment type, however, does not reduce the percentage of observed variance due to sampling error (see Tables 2 and 3), suggesting that results for both maltreatment types still vary substantially depending on other factors. The mean effect sizes suggest that lack of involvement during parent–child interactions better distinguishes neglectful than physically abusive parents from non-maltreating parents. Both the heterogeneous effects within each group, plus the fact that estimates for neglectful parents are based on only seven comparisons ($N = 321$ parents), suggests some caution in this conclusion is warranted.

Research question 2: Family and study characteristics

RQ2 asked to what degree conclusions drawn so far might be moderated by characteristics of the research participants or methods. Beyond maltreatment type, the following potential moderating factors were assessed: mean parent age, parent sex (mothers, fathers), percentage of parents who were racial/ethnic minorities (USA studies only), mean child age, percentage of children who were male, type of control group (community, clinical), publication date, research setting (home, lab), observation length, and task structure. Other possible moderators (e.g., USA vs. international studies) could not be examined because of the small number of studies for some subgroups (there were only four international studies).

For each of the three clusters, the 2–3 moderators that had the biggest impact on effect sizes are discussed. Tests of statistical significance are reported for associations between continuous moderator variables and effect sizes; however, there is limited power to detect significant associations for anything except large moderating factors because only 33 observational studies have been conducted. For this reason, a more liberal criterion for statistical significance ($p < .10$) was used when conducting moderator analyses.

Aversiveness. As noted already, maltreatment type is a moderator for aversiveness, such that those studies comparing physically abusive vs. non-maltreating parents obtained larger effects than those comparing neglectful vs. non-maltreating parents. Publication year also is inversely associated with the magnitude of obtained effect size for history of parent maltreatment on aversive behavior, $r(36) = -.31, p = .06$. When only studies that compare physically abusive vs. non-maltreating parents are included, the association between publication year and effect size for parental aversiveness is stronger, $r(22) = -.60, p < .01$.

To explore this finding, studies were divided into three groups corresponding to decades. Because only two studies were published prior to 1980, studies published from 1978 to 1989 were grouped into decade one (roughly the 1980s). Studies published from 1990 to 1999 were grouped as decade two (1990s) and those published from 2000 to 2006 were grouped as decade three (post 2000). Subgroup analyses were conducted by decade, both for all studies as well as for those focused specifically on physically abusive parents. When all maltreatment types were assessed, the mean weighted effect size across 19 effects from studies conducted in the 1980s was $d = .69$; sampling error accounted for 67% of observed variance in these 19 effects. For studies published in the 1990s, the mean across 9 effects was $d = .86$, though sampling error accounted for only 10% of the observed variance. For studies published during 2000–2006, however, the mean across 10 effects was only $d = .07$ and sampling error accounted for all (99%) of the observed variance. When analyses were limited to studies comparing physically abusive vs. non-maltreating parents, the mean effect sizes were $d = .75$ (13 effects), $.47$ (4 effects), and $.11$ (7 effects) for studies published in the 1980s, 1990s, and post 2000 respectively, with sampling error accounting for 73%, 100%, and 100% of observed variability in studies from these three decades. Using the Hunter and Schmidt (2004) 75% rule, studies

Table 4

Changes in study characteristics across three decades.

Decade	% Lab studies		Mean Obs. length (min) ^a		Mean sample size (N) ^a	
	All studies	A vs. C studies	All studies	A vs. C studies	All studies	A vs. C studies
1978–1989	40%	46%	112 (112)	103 (114)	35 (25)	43 (30)
1990–1999	36%	33%	84 (138)	18 (10)	55 (34)	65 (40)
2000–2006	82%	86%	17 (9)	17 (9)	99 (57)	99 (70)
Significance test	$\chi^2(2) = 6.07^{**}$	$\chi^2(2) = 4.19$	$F(2,37) = 2.64^*$	$F(2,22) = 3.46^{**}$	$F(2,40) = 8.41^{**}$	$F(2,23) = 3.49^{**}$
Effect size	$V = .38$	$V = .40$	$\eta^2 = .13$	$\eta^2 = .24$	$\eta^2 = .29$	$\eta^2 = .24$

* $p < .10$, ** $p < .05$.^a Mean scores outside of parentheses; standard deviations within parentheses.

comparing physically abusive vs. non-maltreating parents in terms of aversive behavior can be considered homogenous once they are split by decade.

To interpret this pattern, differences in participants and methods over the three decades of research were assessed. Table 4 displays the average number of participants (N), the average length of observation in minutes, and the percentage of lab studies for each decade, both for studies of all maltreatment type and for studies that compare physically abusive vs. non-maltreating parents. The average sample size (N) increased substantially over the three decades. Due perhaps to the labor-intensive nature of observational research, the average length of time each family was observed decreased over time as did the percentage of studies conducted in the home. These trends are related: sample size and observation length are inversely associated across studies, $r(38) = -.31$, $p = .05$, and observation length on average is longer in home studies ($M = 141$ min, $SD = 131$, $n = 20$ effect sizes) than in lab studies ($M = 20$ min, $SD = 13$, $n = 20$ effect sizes), $t(38) = 4.11$, $p < .01$, $d = 1.30$.

The methodological features of observation setting (home, lab) and length also are associated with the magnitude of effects obtained for parent aversiveness. When analyzing studies of all maltreatment types (abuse, neglect, or mixed vs. comparison), the mean weighted effect size for the 17 effects from home studies is $d = .58$, whereas the mean for 20 effects from lab studies is $d = .37$. Sampling error accounts for 72% of the observed variability in effect sizes from home studies but only 12% for lab studies. For comparisons of physically abusive vs. non-maltreating parents, the mean d for 10 effects from home studies is .63, whereas the mean d for 14 effects from lab studies is only .26. Sampling error accounts for 64% of the observed variability in effect sizes from home studies and 49% from lab studies, indicating that splitting studies based on setting (home, lab) produces mean effect size differences across groups and reduces variability within groups.

Observation length is not related to the magnitude of effect size for parental aversive behavior across studies of all maltreatment types, $r(34) = .01$, ns , but it is when the analysis is limited to studies comparing physically abusive vs. non-maltreating parents, $r(20) = .40$, $p = .06$. To clarify this, observation length was divided into three groups: studies comparing physically abusive vs. non-maltreating parents for 15 min or less (8 effects), for 16–59 min (8 effects), and for 60 min or longer (8 effects). Mean weighed effect sizes were $d = .26$, .34, and .71 respectively for short, medium, and long studies. Sampling error accounted for 100% of the observed variance in effect sizes for short studies, 32% for medium-length studies, and 56% for long studies. For comparisons of physically abusive vs. non-maltreating parents in terms of aversive behavior, studies which observed families for an hour or more obtained much larger effects than those which observed less than 60 min.

Finally, the impact of publication date on effect sizes for parental aversiveness appears due, in large part, to methodological features. For studies involving all maltreatment types, the association between publication year and effect size drops from $r(36) = -.31$, $p = .06$ to partial $r(29) = -.13$, ns , once effects of setting (home, lab), observation length and sample size on parental aversiveness are controlled. For studies comparing physically abusive to non-maltreating parents, the association between publication year and effect size drops from $r(21) = -.60$, $p < .01$ to partial $r(15) = -.29$, ns once the effects of these same methodological features are controlled. In sum, studies observing physically abusive parents and those conducted prior to 2000 obtained the largest effect sizes for parental aversiveness.

Positivity. Many studies did not report parents' average age (see Table 1), but for those that did there was a strong inverse association between mean parent age and size of effect obtained, $r(16) = -.48$, $p < .05$. To clarify this finding, studies were broken into those in which the average parent age was 30 years or older vs. younger than 30 and conducted meta-analyses separately for each group. For the 10 effects from studies in which the average parent age was <30 , the mean weighted $d = 1.09$ and sampling error accounted for 37% of the observed variance in effect sizes. For the eight effects from studies in which the average parent age = 30, the mean weighted $d = .20$ and sampling error accounted for 36% of the observed variance in effect sizes. Parent age is a strong moderator of the magnitude of difference in positive behavior between maltreating vs. non-maltreating parents, such that studies with younger parents obtained much larger differences than those with older parents. Splitting studies by parent age also increased the percentage of observed variance due to sampling error, though there is still considerable heterogeneity within both age groups.

Studies with older parents also involved older children; the association between mean parent and child age across studies is $r(22) = .69, p < .001$. Hence, it makes sense that mean child age also is inversely associated with obtained effect sizes for positive behaviors, $r(28) = -.30, p < .10$. Studies were split into those in which the mean child age was less than 5 years (16 effect sizes) vs. those in which children on average were 5 or older (14 effect sizes). For parents interacting with younger children, the mean weighted d for maltreating vs. non-maltreating parents' positivity was .89, though sampling error accounted for only 22% of the variance in these effect sizes. For parents interacting with older children, the mean d was only .30, and sampling error explained 50% of the variance in these effect sizes. In sum, positive behavior did a better job of distinguishing maltreating vs. non-maltreating parents for studies in which both parents and children on average were younger as opposed to older.

Involvement. As noted already, studies comparing neglectful vs. non-maltreating parents obtained larger differences in involvement than those comparing physically abusive and non-maltreating parents. Task structure is a second moderator. The mean weighted effect size for involvement differences between maltreating and non-maltreating parents was $d = 1.10$ in five studies with unstructured tasks, $d = .55$ in five studies with moderately structured task but only $d = .25$ in six studies using highly structured tasks (studies using multiple tasks with different levels of structure were not included). Sampling error accounted for 100% of the observed variance in effect sizes from studies with highly structured tasks, but only 11% for studies with moderately structured tasks and 7% for those with unstructured tasks, indicating that effect sizes from the latter two groups of studies still were heterogeneous. In sum, studies observing neglectful parents and those employing unstructured tasks obtained the largest effect sizes for parental involvement.

Discussion

This paper reports a meta-analysis of observational studies that compare maltreating and non-maltreating parents in terms of their aversiveness, positivity, and involvement during interactions with their children. Although more than 30 observational studies have been published, drawing conclusions from this literature has been difficult: some studies find statistically significant differences between maltreating and non-maltreating parents for all three types of behavior, some find differences in some behaviors but not others, and some find no differences. Taking stock of this situation, Bennett et al. (2006) concluded that “consensus on how to best observe parents to detect the potential for maltreatment has yet to emerge” (p. 72). By combining sample sizes across studies and formally testing for moderating factors, the current meta-analysis offers insights about this issue.

Across studies, differences between maltreating and non-maltreating parents are medium in size ($d = .46$ to $.62$) for all three types of behavior. These differences, however, vary depending on family and study characteristics. Type of maltreatment history (physically abusive vs. neglectful) had some impact on the degree of difference between maltreating and non-maltreating parents. Aversive behaviors better distinguished physically abusive from non-maltreating parents whereas involvement better distinguished neglectful from non-maltreating parents. Positive behaviors were equally useful in distinguishing either maltreatment group from non-maltreating parents. These findings support the decision to look at three separate clusters of parenting behavior, especially the potential role of involvement in helping understand the etiology of neglect. Neglect is evident not just in a parent's failure to meet a child's basic needs (e.g., clothing, supervision) but also in a more subtle failure to display attentiveness and responsiveness. Lack of involvement may reflect a parent's own mental models about self and relationships and simultaneously communicate implicit messages about relationships to the neglected child (Crittenden & Ainsworth, 1989).

Publication year was a moderator for aversive behavior. Studies published post 2000 found much smaller differences between maltreating and non-maltreating parents compared to those published before that time. Researchers have become increasingly aware of the importance of statistical power, which probably explains why average sample size has increased from approximately $N = 40$ in studies from the 1980s to $N = 100$ in studies since 2000. To increase sample size, however, researchers are making other methodological choices such as observing families for 30 min or less and conducting studies in the lab. Such interactions typically are pleasant and undemanding, making it unlikely that anything other than small differences between maltreating and non-maltreating parents will occur. This meta-analysis shows that maltreating (especially physically abusive) parents do exhibit substantially more aversive behavior than non-maltreating parents when families are observed for longer periods of time (especially at home). These are, of course, exactly the conditions that maltreated children experience on a day-to-day basis. To understand how aversive parental behavior can create or sustain dangerous patterns of interaction, researchers need to observe under conditions where stressors such as child noncompliance are likely to occur. Methodological choices also have practical implications. For example, a parenting program that uses observational tasks as an evaluation tool is unlikely to detect reductions in parental aversive behavior if families are observed only for short periods of time (e.g., <15 min) even if the program really is producing desired changes.

Parent and child age were moderators for positive behavior, in that studies with younger parents (<30 years) and children (<5 years) detected much larger differences between maltreating and non-maltreating parents compared to studies with older parents and children. Young parental age may accentuate differences between maltreating and non-maltreating parents for several reasons. Younger parents with a history of child maltreatment may: (a) have especially poor parenting skills, (b) perceive greater life stress, which in turn influences their perceptions of child behavior (as information-processing models would suggest, see Milner, 2003), or (c) have younger children who cope differently than older children (as attachment theory

would suggest, see Crittenden, 1992). Older maltreating parents also could have greater knowledge as to what constitutes “positive” parenting and hence exhibit such behavior when being observed by researchers. In any case, the fact that younger maltreating parents display much less positive behavior than younger comparison parents reinforces the importance of home visiting programs that target “at risk” (young, low-income, unmarried) mothers from pregnancy through their child’s early years and offer social support, health services, and information about empathic, sensitive parenting (Olds, 2006).

Task structure was a moderator for parental involvement, such that studies using unstructured tasks (e.g., free play without direction) found larger differences between maltreating and non-maltreating parents than did those using only highly structured tasks (e.g., teach your child X). Unstructured activities may allow for a greater observable range of parental involvement than structured tasks since the latter tell all parents what to do in completing the activity. Because behavior always is, in part, a function of the immediate context, decisions regarding what to have parents and children do during observational research must be made as mindfully as those regarding where and for how long to observe. Evaluation studies using observational methods would be well advised to have parents and children complete unstructured as opposed to only structured tasks.

No study is without limitations, the current meta-analysis included. The number of observational studies with data on neglectful parents—or fathers with any maltreatment history—is small; hence conclusions about these groups should be viewed with caution. Very few studies were conducted outside of the USA, which suggests potential boundary conditions on our conclusions. The studies included in this meta-analysis classify parents as maltreating vs. non-maltreating based on CPS involvement, a procedure with known limitations (Lau et al., 2005). Studies also compare maltreating and non-maltreating parents in terms of behavioral frequencies, whereas social-interactional views (Reid, 1986; Urquiza & Timmer, 2002) suggest that these groups also should differ in the duration, timing, and/or sequencing of their responses to child behavior. These limitations suggest important avenues for future research. Despite these limits, this meta-analysis provides a clearer picture of the nature and extent of differences in the ways that physically abusive, neglectful, and non-maltreating parents interact with their children, highlights challenges in doing observational research with families, and offers practical insights regarding intervention and evaluation.

References¹

- *Alessandri, S. M. (1992). Mother–child interactional correlates of maltreated and nonmaltreated children’s play behavior. *Development and Psychopathology*, 4, 257–270.
- *Alessandri, S., & Lewis, M. (1996). Differences in pride and shame in maltreated and nonmaltreated preschoolers. *Child Development*, 67, 1857–1869.
- *Aragona, J. A., & Eyberg, S. M. (1981). Neglected children: Mothers’ report of child behavior problems and observed verbal behavior. *Child Development*, 52, 596–602.
- Belsky, J. (1993). Etiology of child maltreatment: A developmental-ecological analysis. *Psychological Bulletin*, 114, 413–434.
- *Bennett, D. S., Sullivan, M. W., & Lewis, M. (2006). Relations of parental report and observation of parenting to maltreatment history. *Child Maltreatment*, 11, 63–75.
- *Borrego, J., Timmer, S. G., Urquiza, A. J., & Follett, W. C. (2004). Physically abusive mothers’ responses following episodes of child noncompliance and compliance. *Journal of Consulting and Clinical Psychology*, 72, 897–903.
- *Bousha, D. M., & Twentyman, C. T. (1984). Mother–child interactional style in abuse, neglect, and control groups: Naturalistic observations in the home. *Journal of Abnormal Psychology*, 93, 106–114.
- *Brassard, M. R., Hart, S. N., & Hardy, D. B. (1993). The psychological maltreatment rating scales. *Child Abuse & Neglect*, 17, 715–729.
- Bugental, D. B., Ellerson, P. C., Lin, E. K., Rainey, B., Kokotovic, A., & O’Hara, N. (2002). A cognitive approach to child abuse prevention. *Journal of Family Psychology*, 16, 243–258.
- *Burgess, R. L., & Conger, R. D. (1978). Family interaction in abusive, neglectful, and normal families. *Child Development*, 40, 1163–1173.
- Burgoon, J. K., & Newton, D. A. (1991). Applying a social meaning model to perceptions of conversational involvement: Comparing observer and participant perspectives. *Southern Journal of Communication*, 56, 96–113.
- Camras, L. A., Ribordy, S., Hill, J., Martino, S., Sachs, V., Spaccarelli, S., & Stefani, R. (1990). Maternal facial behavior and the recognition and production of emotional expression by maltreated and nonmaltreated children. *Developmental Psychology*, 26, 304–312.
- Cegala, D. J., Savage, G. T., Brunner, C. T., & Conrad, A. B. (1982). An elaboration of the meaning of interaction involvement: Toward the development of a theoretical concept. *Communication Monographs*, 51, 320–338.
- Cerezo, M. A. (1997). Abusive family interaction: A review. *Aggression and Violent Behavior*, 2, 215–240.
- *Cerezo, M. A., & D’Ocon, A. (1995). Maternal inconsistent socialization: An interactional pattern with maltreated children. *Child Abuse Review*, 4, 14–31.
- *Cerezo, M. A., & D’Ocon, A. (1999). Sequential analyses in coercive mother–child interaction: The predictability hypothesis in abusive versus nonabusive dyads. *Child Abuse & Neglect*, 23, 99–113.
- Cerezo, M. A., D’Ocon, A., & Dolz, L. (1996). Mother–child interactive patterns in abusive families versus in nonabusive families: An observational study. *Child Abuse & Neglect*, 20, 573–587.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Coster, W. J., Gersten, M. S., Beeghly, M., & Cicchetti, D. (1989). Communicative functioning in maltreated toddlers. *Developmental Psychology*, 25, 1020–1029.
- Crittenden, P. M. (1982). Abusing, neglecting, problematic, and adequate dyads: Differentiating by patterns of interaction. *Merrill-Palmer Quarterly*, 27, 201–218.
- Crittenden, P. M. (1992). Children’s strategies for coping with adverse home environments: An interpretation using attachment theory. *Child Abuse & Neglect*, 16, 329–343.
- Crittenden, P. M., & Ainsworth, M. D. S. (1989). Child maltreatment and attachment theory. In D. Cicchetti & V. Carlson (Eds.), *Child maltreatment: Theory and research on the causes and consequences of child abuse and neglect* (pp. 432–463). Cambridge, UK: Cambridge University Press.
- DiLalla, D. L., & Crittenden, P. M. (1990). Dimensions of maltreated children’s home behavior: A factor analytic approach. *Infant Behavior & Development*, 13, 439–460.
- Dolz, L., Cerezo, M. A., & Milner, J. S. (1997). Mother–child interactional patterns in high and low risk families. *Child Abuse & Neglect*, 21, 1149–1158.

¹References marked with an asterisk indicate studies included in the meta-analysis.

- *Edwards, A., Shipman, K., & Brown, A. (2005). The socialization of emotional understanding: A comparison of neglectful and nonneglectful mothers and their children. *Child Maltreatment, 10*, 293–304.
- Eyeberg, S. M., Nelson, M. M., Duke, M., & Boggs, S. R. (2005). *Manual for the dyadic parent–child interaction coding system* (3rd ed.). Retrieved March 12, 2007 from www.pcit.org
- *Fagan, J., & Dore, M. M. (1993). Mother–child play interaction in neglecting and non-neglecting mothers. *Early Child Development and Care, 87*, 59–68.
- *Givens, D. B. (1978). Contrasting nonverbal styles in mother–child interaction: Examples from a study of child abuse. *Semiotica, 24*, 33–47.
- *Haskett, M. E., Ahern, L. S., Ward, C. S., & Allaire, J. C. (2006). Factor structure and validity of the parenting stress index-short form. *Journal of Clinical Child and Adolescent Psychology, 35*, 302–312.
- Herrenkohl, E. C., Herrenkohl, R. C., Toedter, L., & Yanushefski, M. (1984). Parent–child interactions in abusive and nonabusive families. *Journal of the American Academy of Child Psychiatry, 23*, 641–648.
- Howes, P. W., Cicchetti, D., Toth, S. L., & Rogosch, F. A. (2000). Affective, organizational, and relational characteristics of maltreating families: A systems perspective. *Journal of Family Psychology, 14*, 95–110.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.
- *Kavanagh, K. A., Youngblade, L., Reid, J. B., & Fagot, B. I. (1988). Interactions between children and abusive versus control parents. *Journal of Clinical Child Psychology, 17*, 137–142.
- *Koenig, A. L., Cicchetti, D., & Rogosch, F. A. (2000). Child compliance/noncompliance and maternal contributors to internalization in maltreating and non-maltreating dyads. *Child Development, 71*, 1018–1032.
- *Lahey, B. B., Conger, R. D., Atekson, B. M., & Treiber, F. A. (1984). Parenting behavior and emotional status of physically abusive mothers. *Journal of Consulting and Clinical Psychology, 52*, 1062–1071.
- Lau, A. S., Leeb, R. T., English, D., Graham, J. C., Briggs, E. C., Brody, K. E., & Marshall, J. M. (2005). What's in a name? A comparison of methods for classifying predominant type of maltreatment. *Child Abuse & Neglect, 29*, 533–551.
- *Lau, A. S., Valerie, S. M., McCarty, C. A., & Weitz, J. R. (2006). Abusive parents' reports of child behavior problems: Relationship to observed parent–child interactions. *Child Abuse & Neglect, 30*, 639–655.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- *Lorber, R., Felton, D. K., & Reid, J. B. (1984). A social learning approach to the reduction of coercive processes in child abusive families: A molecular analysis. *Advances in Behavioral Research and Therapy, 6*, 29–45.
- *Lyons-Ruth, K., Connell, D. B., Zoll, D., & Stahl, J. (1987). Infants at social risk: Relations among maltreatment, maternal behavior, and infant attachment. *Developmental Psychology, 23*, 223–232.
- *Mash, E. J., Johnston, C., & Kovitz, K. (1983). A comparison of the mother–child interactions of physically abused and non-abused children during play and task situations. *Journal of Clinical Child Psychology, 12*, 337–346.
- Milner, J. S. (2003). Social information processing in high-risk and physically abusive parents. *Child Abuse & Neglect, 27*, 7–20.
- Morton, N., & Browne, K. D. (1998). Theory and observation of attachment and its relation to child maltreatment: A review. *Child Abuse & Neglect, 22*, 1093–1104.
- Nastasi, B. K., & Hill, S. D. (1982). Interactions between abusing mothers and their children in two situations. *Bulletin of the Psychonomic Society, 20*, 79–81.
- *Oldershaw, L., Walters, G. C., & Hall, D. K. (1986). Control strategies and noncompliance in abusive mother–child dyads: An observational study. *Child Development, 57*, 722–732.
- *Oldershaw, L., Walters, G. C., & Hall, D. K. (1989). A behavioral approach to the classification of different types of physically abusive mothers. *Merrill-Palmer Quarterly, 35*, 255–279.
- Olds, D. L. (2006). The Nurse–Family Partnership: An evidence-based preventive intervention. *Infant Mental Health Journal, 27*, 5–25.
- Reid, J. B., Kavanagh, K., & Baldwin, D. V. (1987). Abusive parents' perceptions of child problem behaviors: An example of parental bias. *Journal of Abnormal Child Psychology, 15*, 457–466.
- *Reid, J. B., Taplin, P. S., & Lorber, R. (1981). A social interactional approach to the treatment of abusive families. In R. B. Stuart (Ed.), *Violent behavior: Social learning approaches to prediction, management, and treatment* (pp. 83–101). New York: Brunner/Mazel.
- Reid, J. S. (1986). Social interactional patterns in families of abused and non-abused children. In C. Zahn-Waxler, E. M. Cummings, & R. Iannotti (Eds.), *Altruism and aggression: Biological and social origins* (pp. 238–257). Cambridge, UK: Cambridge University Press.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* ((Rev. ed.)). Newbury Park, CA: Sage.
- *Schaeffer, S. (1983). Peer behavior, caregiver-directed behavior, and mother–child interaction in maltreated children. Unpublished doctoral dissertation, City University of New York.
- *Schindler, F., & Arkowitz, H. (1986). The assessment of mother–child interactions in physically abusive and nonabusive families. *Journal of Family Violence, 1*, 247–257.
- *Shipman, K. L., & Zeman, J. (1999). Emotional understanding: A comparison of physically maltreating and nonmaltreating mother–child dyads. *Journal of Clinical Child Psychology, 28*, 407–417.
- *Silber, S., Bermann, E., Henderson, M., & Lehman, A. (1993). Patterns of influence and response in abusing and nonabusing families. *Journal of Family Violence, 8*, 27–38.
- Straus, M. A., Hamby, S. L., Finkelhor, D., Moore, D. W., & Runyan, D. (1998). Identification of child maltreatment with the Parent–Child Conflict Tactics Scales: Development and psychometric data from a national sample of American parents. *Child Abuse & Neglect, 22*, 249–270.
- *Timmer, S. G., Borrego, J., Jr., & Urquiza, A. J. (2002). Antecedents of coercive interactions in physically abusive mother–child dyads. *Journal of Interpersonal Violence, 17*, 836–853.
- Trickett, P. K., & Kuczynski, L. (1986). Children's misbehaviors and parental discipline strategies in abusive and non-abusive families. *Developmental Psychology, 22*, 115–123.
- Urquiza, A. J., & Timmer, S. G. (2002). Patterns of interaction within violent families: Use of social interaction research methodology. *Journal of Interpersonal Violence, 17*, 824–835.
- *Valentino, K., Cicchetti, D., Toth, S. L., & Rogosch, F. A. (2006). Mother–child play and emerging social behaviors among infants from maltreating families. *Developmental Psychology, 42*, 474–485.
- *Wasserman, G. A., Green, A., & Allen, R. (1983). Going beyond abuse: Maladaptive patterns of interaction in abusing mother–infant pairs. *Journal of the American Academy of Child Psychiatry, 22*, 245–252.
- *Webster-Stratton, C. (1985). Comparison of abusive and nonabusive families with conduct-disordered children. *American Journal of Orthopsychiatry, 55*, 59–60.
- *Whipple, E. E., & Webster-Stratton, C. (1991). The role of parental stress in physically abusive families. *Child Abuse & Neglect, 15*, 279–291.
- Whipple, E. E., & Wilson, S. R. (1996). Evaluation of a parent education and support program for families at risk for physical child abuse. *Families in Society, 77*, 227–239.
- Wilson, S. R., Morgan, W. M., Hayes, J., Bylund, C. L., & Herman, A. (2004). Mothers' child abuse potential as a predictor of maternal and child behaviors during play-time interactions. *Communication Monographs, 71*, 395–421.
- Wilson, S. R., Shi, X., Tirmenstein, L., Norris, A., & Rack, J. (2006). Parental physical negative touch and child noncompliance in abusive, neglectful, and comparison families: A meta-analysis of observational studies. In L. Turner & R. West (Eds.), *Family communication: A reference for theory and research* (pp. 237–258). Newbury Park, CA: Sage.
- Wolfe, D. A., & Sandler, J. (1981). Training abusive parents in effective child management. *Behavior Modification, 5*, 320–335.