The association between parent-reported and observed parenting

A multi-level meta-analysis

Hendriks, A.M.; Van der Giessen, D.; Stams, G.J.J.M.; Overbeek, G.

DOI
10.1037/pas0000500

Publication date
2018

Document Version
Final published version

Published in
Psychological Assessment

Psychological Assessment

1040-3590/17/$12.00 http://dx.doi.org/10.1037/pas0000500

The Association Between Parent-Reported and Observed Parenting: A Multi-Level Meta-Analysis

A. M. Hendriks
Vrije Universiteit Amsterdam

D. Van der Giessen, G. J. J. M. Stams, and G. Overbeek
University of Amsterdam

The purpose of the present meta-analysis was to investigate the strength of the association between parent-reported and observed parenting, and to investigate which specific characteristics of participants, questionnaires, or observational procedures moderate this association. A systematic search of relevant peer-reviewed articles published between January 2000 and December 2014 yielded 36 articles (N = 8,510) and 89 effect sizes. Results from a 3-level random-effects meta-analysis demonstrated a weak, yet significant, overall association of r = .17 between parent-reported and observed parenting. The magnitude of the effect size depended on questionnaire length (larger effect for more items) and the type of parenting investigated (largest effects for negative parent behaviors, next largest effects for positive parent behaviors, and smallest effect for controlling parent behaviors). In conclusion, this study shows that the strength of the association between parent-reported and observed parenting is small but significant.

Public Significance Statement
Parenting is often measured using self-reports or observations. This study has shown that the association between parenting measured through self-report and through observation is small, but significant. The association is larger for negative and positive parenting compared with controlling parenting and the association is larger for studies using questionnaires with more items.

Keywords: parenting, parent-reports, observation, meta-analysis

Parenting has long been theorized to be important for the development of children (e.g., Campbell, 1997; Caron, Weiss, Harris, & Catron, 2006; Kuppens, Grietens, Onghena, & Michiels, 2009). Parenting is usually assessed through self-report or through observation (e.g., Gardner, 2000; Locke & Prinz, 2002). However, the rationale for choosing one of these methods of measurement is not always clear. Furthermore, whether a higher score on a certain self-reported parenting measure is indicative of a higher score on the same parenting variable when observed, or vice versa, is unclear. Whereas several substantial meta-analyses have found small to medium association between multiple informants for child behavior or child psychopathology (e.g., Achenbach, Krukowski, Dumenci, & Ivanova, 2005; Achenbach, McConaughy, & Howell, 1987; Duhig, Renk, Epstein, & Phares, 2000), to the best of our knowledge the meta-analytic association between different measures of parenting behaviors is still unknown. Therefore, the purpose of the present study is to perform a meta-analysis to investigate the strength of the association between parent-reported and observed parenting.

In assessing parenting, questionnaires and observation instruments are used to measure similar constructs, however, there are differences between questionnaires and observation instruments. For instance, they are used for different purposes. Questionnaires are more often used for the assessment of thoughts, feelings, attitudes, and perceptions of behavior, whereas observational measures are mostly used to assess behavior (Gardner, 2000). In addition, parents and observers differ in their roles toward the child and on the existing knowledge they have on the range of behaviors of the child. For example, parents are more acquainted with the behavior that the child displays in day-to-day interactions compared with observers (Achenbach et al., 2005), which may cause parents and observers to perceive situations differently. Furthermore, assessment types differ in the extent to which social desirability and perceptions influence the parenting scores. Within the construct of parenting, there are elements that increase the likelihood of distortions associated with self-report (Morsbach & Prinz, 2006). For instance, when parents report their own parenting behaviors, they have a tendency to underreport their hostility (Waylen, Stallard, & Stewart-Brown, 2008) and to overreport the extent to which they show warmth and exert control (Bögels & van Melick, 2004; Schwarz, Barton-Henri, & Pruzinsky, 1985). Although the different properties of questionnaires and observation

A. M. Hendriks, Department of Biological Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam; D. Van der Giessen, G. J. J. M. Stams, and G. Overbeek, Department of Child Development and Education, Research Priority Area Yield, University of Amsterdam.

Correspondence concerning this article should be addressed to A. M. Hendriks, Department of Biological Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, van den Boechorststraat 1, 1081 BT, Amsterdam, the Netherlands. E-mail: a.m.hendriks@vu.nl

instruments are known, there does not appear to be a common practice for measuring parenting.

How parents report their own parenting behaviors and how they actually behave in interactions is thought to be associated. Yet, it is still unknown how strong the association between parent-reported and observed parenting actually is, and which factors might increase or decrease the strength of the association. Knowing the strength of the association between parent-reported and observed parenting could inform us on the extent to which self-reports or observations provide similar information on parenting. In addition, knowing which factors are associated with stronger associations between parent-reported and observed parenting is helpful in designing more precise assessments. Overall, this study could provide a first step toward establishing common or "best" practices for measuring parenting.

In this study we examine the strength of the association between parent-reported and observed parenting. With this study, we want to inform research and clinical practice on the extent to which these different measures of parenting are related and thereby provide information on whether they potentially provide different information. In addition, we examine whether the association between parent-reported and observed parenting may be moderated by characteristics of participants, questionnaires, observations, and parenting.

Moderators

Participant characteristics refer to specific parent and child characteristics that may influence the strength of the association between parent-reported and observed parenting, including gender of the parent, sample type for the parents (i.e., general or clinical sample), age of the children, and sample type for the children (i.e., children from the general population, children with a history of maltreatment, ADHD, externalizing behavior problems, or an autism spectrum disorder).

Research on parenting is mostly focused on mothers (e.g., Bögels & Phares, 2008; Phares, 1992; Smith, 2011), but there has been an increase of interest in the influence of fathers (e.g., Calzada, Eyberg, Rich, & Querido, 2004; McBride, Schoppe, & Rane, 2002; Phares, 1992). If the association between parent-reported and observed parenting is moderated by whether there are more mothers in a sample, this might have implications for including mothers and fathers in future research. If the association between parent-reported and observed parenting differs for parents from general samples or from clinical samples, this should be considered when measuring parenting in different types of samples. Because as children age, disciplinary techniques and nurturing behaviors of parents change to adapt to the changing needs of their children (Locke & Prinz, 2002), child age might affect the strength of the association between parent-reported and observed parenting, and thus have implications for assessing parenting for different age groups. If whether children are from clinical or general populations samples influences the association between parent-reported and observed parenting, this should be considered when choosing an assessment method for parenting in a sample with children from the general population or a clinical sample.

Questionnaire characteristics consist of specific features that might influence the strength between parent-reported and observed parenting. For this study the length of the questionnaire and which questionnaire is administered are of interest. Psychometric theory indicates that a greater number of items in a measure often leads to increased reliability (e.g., Cronbach, 1951; Smith, McCarthy, & Anderson, 2000), whereas administrating fewer questions has economic advantages. Many questionnaires exist to assess parenting, however, it could be possible that parenting scores assessed by different questionnaires are related differently to parenting scores from observations.

Observation characteristics refer to specific aspects of an observation procedure that might influence the strength of the association between parent-reported and observed parenting. Observation characteristics of interest are the duration of the observation, the location of the observation, whether the observation was performed directly or from a video, which observation instrument was used, and which type of task was employed.

There are many decisions following the choice to include an observational measure of parenting. Whether the association between parent-reported and observed parenting is moderated by duration of an observation, the location of the observation, whether the observation was performed directly or from a video, or which task to give to the parents, would indicate that scores on parent-reported and observed parenting are related differently across these variables. Similarly, as with questionnaires, the association between parent-reported and observed parenting could be different for different observational instruments.

Parenting refers to which type of parenting was measured (e.g., positive parenting, negative parenting, and controlling parenting). The literature distinguishes between several parenting dimensions: positive parenting (e.g., warmth, responsiveness, positive affect), negative parenting (e.g., hostility, criticism, intrusiveness), and control (e.g., commands, monitoring, discipline). This categorization follows the dominant differentiation between parenting behaviors in the literature (Baumrind, 1971; Locke & Prinz, 2002; McKee, Colletti, Rakow, Jones, & Forehand, 2008; Steinberg, Mounts, Lamborn, & Dornbusch, 1991). Measuring positive, negative, or controlling parenting could yield different associations between parent reports and observations.

The Present Study

The main aim of this study was to examine to what extent parent-reported and observed parenting were associated. In addition, we explore which participant characteristics, questionnaire characteristics, observation characteristics, and parenting moderated the association between parent-reported and observed parenting.

Method

Articles were collected through a computerized database search of PsycARTICLES, PsycINFO, ERIC, and MEDLINE searching for articles published from January 2000 through December 2014 to cover a period of the last 15 years following the reviews on measurements of parenting by Gardner (2000) and by Locke and Prinz (2002). The search terms were: report, observ*, parent*, child, adolescent, convergence, correspondence, convergent validity, and correlation put together in different combinations. With these search terms, 14,070 candidate articles were identified. Because an investigation of the association between self-reported and

observed parenting in most studies was not the primary goal, during the first collection of articles abstracts were scanned on whether both questionnaire and observational measures of parenting had been administered. After screening the abstracts, 290 articles were included for further analyses of eligibility. These 290 articles were then scanned to determine whether there were correlations reported either between parent-reported positive parenting and observed positive parenting, parent-reported negative parenting and observed negative parenting, or parent-reported control and observed control, and to determine whether other inclusion criteria were met. Of 33 articles that did not report correlations, authors were contacted (yielding an additional 17 articles to include). To further ensure that all relevant articles were included, lists of references of the useful articles were scanned. These processes together led to the inclusion of 36 articles (N = 8,510) with a total of 89 effect sizes, as summarized in Table 1. Because we conducted secondary analyses on already gathered data, there was no involvement of a research ethics committee.

Criteria for the Selection of Studies

Articles included in the analyses had to meet the following inclusion criteria:

1. The study was published between January 2000 and December 2014.[1]
2. Children in the sample were between 1 and 18 years.[2]
3. The parent who reported parenting behavior did not have any impairment that would restrict behavioral or cognitive functioning, such as autism or intellectual disabilities.
4. The study provided a measure of association between observed positive parenting, negative parenting, or control combined with assessment using a questionnaire that measured parenting behavior on a similar dimension.
5. The report on parenting behavior was filled in by the parent under scrutiny.
6. The reports were not filled in retrospectively.
7. Observations were either directly performed or videotaped.

Moderators

Participant characteristics. The percentage of mothers was coded as a continuous variable, summarizing the percentage of mothers from the sample that reported on the parenting behaviors. Sample type for the parents was coded either as a general population sample or a clinical sample. Age of the children was coded as a continuous variable. Sample type for the children was coded categorically as one of the following categories: general sample, children with maltreatment history, children with ADHD, children with externalizing behavior problems, and children with an autism spectrum disorder. There were no studies including children with internalizing problems.

Questionnaire characteristics. The number of questions to measure the relevant parenting construct was coded as a continuous variable, with the score reflecting the actual number of items in the questionnaires. Which questionnaire was administered was coded categorically with the following categories: Parenting Dimensions Inventory, Parenting Practices Interview, Alabama Parenting Questionnaire, Parenting Scale, scores of multiple questionnaires combined, or other. Each of these categories represented the questionnaires that were used in at least three studies.

Observation characteristics. Duration of the observation was coded as a continuous variable representing the duration in minutes of the concerning observation task. The location of the observation referred to whether an observation took place either at home or in a laboratory setting. Direct observation versus videotaped referred to whether the observation was performed while the interaction took place or later on from a videotape. Which coding instrument was used for the observation was coded categorically with the following categories: Coders Impression Inventory, Dyadic Parent-Child Interaction Coding System, Family Observation System, a combination of multiple observation systems, or other. These categories represented the observation instruments that were used in at least three studies. Which task was coded for the observation was one of the following categories: a play task, a combination of playing and cleaning up, a homework task, a discussion task, observations made during the interview, a compliance task, a problem solving task, or other.

Parenting. Which type of parenting was measured for the association between parent reports and observations of parenting was coded categorically. Categories consisted of positive (e.g., warm, prosocial, supportive parenting), negative (e.g., hostile, critical, harsh parenting), or controlling (e.g., commands, monitoring, discipline).

Coding the Studies

To capture the information regarding moderators, we developed a coding scheme. The moderators were coded according to this coding scheme by the first author. If studies reported associations between parent-reported and observed parenting before and after an intervention, we only included the results from before an intervention. In total, 22% of the articles were randomly selected to be double coded by the second author to check for interrater reliability. For continuous variables, we calculated intraclass correlations. The average intraclass correlation measuring correspondence between the coders was .90 on the continuous variables. The lowest intraclass correlations between the two raters were for the

[1] To make sure that no relevant body of literature before 2000 was missed, we conducted an additional electronic literature search for articles published before 2000. This search yielded 5916 results. Based on the titles and abstracts, we fully scanned 31 articles; only three articles appeared to be eligible. This number of studies was minimal compared with the far greater number of studies found in the period of 2000–2015, and thus indicated that no meaningful proportion of literature was missed due to our time frame.

[2] The review by Locke and Prinz (2002) focused on measures of parenting for children aged 1-18, because parenting during infancy has not yet developed into stable behavioral patterns. Specifically, Lecuyer-Maus (2000) explains that only after a year, mothers and children have practiced their interactions to such an extent that meaningful patterns can be investigated. Thus, because there are significant developmental differences between children younger and older than 1 year, we decided to not include studies that focused on parenting for children below 1 year of age.

Table 1
The Included Articles

Authors	*N*	% mothers	Type sample parents	Age children	Type sample children	No. items	Quest.	Dur. obs.	Loc.	Direct	Obs.	Task	Type parenting	*r*
Asscher Hermanns, and Deković (2008)	105	100	G	2,5	G	8	Other	12	H	V	Other	P	Pos	−.24
Asscher et al. (2008)	105	100	G	2,5	G	6	PDI	12	H	V	Other	P	Neg	.17
Asscher et al. (2008)	105	100	G	2,5	G	8	Other	12	H	D	CII	Other	Pos	−.31
Asscher et al. (2008)	105	100	G	2,5	G	6	PDI	12	H	D	CII	Other	Neg	.21
Bennett Sullivan, & Lewis (2006)	139	100	G	5	M	5	Other	20	L	V	Other	PS	Neg	.08
Bhandari & Barnett (2007)	78	96	G	4,3	G		Other	7	L	V	Other	P	Cont	.30
Blair et al. (2014)	1,100	99	G	1,3	G	18	Other		L		Other	PS	Pos	.44
Blair et al. (2014)	1,050	99	G	2	G	18	Other		H		Other	PS	Pos	.34
Breitenstein et al. (2012)	532	90	G	2,8	G	22	Other	15		V	DPICS		Pos	.83
Breitenstein et al. (2012)	532	90	G	2,8	G	6	Other	15		V	DPICS		Cont	−.08
Brotman et al. (2005)	92	90	G	4,1	EXT	1	PPI	10	H	V	Multiple	P	Pos	.28
Brotman et al. (2005)	92	90	G	15,6	EXT	13	PPI	10	H	V	Multiple	P	Neg	.17
Brotman et al. (2011)	171	90	G	15,6	G	12	PPI	15	H	V	Other	PC	Cont	−.02
Browne, Meunier, O'Connor, and Jenkins (2012)	381	83	G	1,6	G	5	Other	15	H	V	Other	P	Pos	.02
Browne et al. (2012)	381	83	G	1,6	G	5	Other	15	H	V	Other	P	Neg	−.02
Chronis-Tuscano et al. (2008)	70	89	C	8,1	ADHD	10	APQ	5	H	V	DPICS	P	Pos	.37
Chronis-Tuscano et al. (2008)	70	100	C	8,1	ADHD	10	APQ	10	H	V	DPICS	HW	Pos	.16
Feinberg Neiderhiser, Howe, and Hetherington (2001)	702	100	G	14,5	G	11	Other	10	H	V	Other	D	Pos	.06
Feinberg et al. (2001)	702	100	G	14,5	G	4	Other	10	H	V	Other	D	Neg	.01
Feinberg et al. (2001)	699	100	G	12,9	G	11	Other	10	H	V	Other	D	Pos	.07
Feinberg et al. (2001)	699	100	G	12,9	G	4	Other	10	H	V	Other	D	Neg	.05
Feinberg et al. (2001)	694	0	G	14,5	G	11	Other	10	H	V	Other	D	Pos	.03
Feinberg et al. (2001)	694	0	G	14,5	G	4	Other	10	H	V	Other	D	Neg	.06
Feinberg et al. (2001)	701	0	G	12,9	G	11	Other	10	H	V	Other	D	Pos	−.02
Feinberg et al. (2001)	701	0	G	12,9	G	4	Other	10	H	V	Other	D	Neg	.03
Føssum, Mørch, Handegård, and Drugli (2007)	532	100	G	6,6	EXT	15	PPI	15	H	V	DPICS	PC	Pos	.07
Føssum et al. (2007)	532	100	G	6,6	EXT	14	PPI	15	H	V	DPICS	PC	Neg	.02
Friesen, Woodward, Horwood, and Fergusson (2013)	146	77	G		G	13	Multiple		H	D	Other	DI	Pos	.42
Friesen et al. (2013)	146	77	G		G	6	Multiple		H	D	Other	DI	Pos	.47
Hahlweg, Heinrichs, Kuschel, Bertram, and Naumann (2014)	276	100	G	4,5	G		PS	20	H	V	FOS	PC	Neg	.06
Hanisch Hautmann, Plück, Eichelberger, and Döpfner (2014)	155	100	G	4,2	G	13	Multiple	20	H	V	CII	PC	Pos	.13
Hawes & Dadds (2006)	56		G	6,3	EXT	14	APQ	60	H	V	FOS	Other	Pos	.32
Hawes & Dadds (2006)	56		G	6,3	EXT	14	APQ	20	H	V	FOS	PC	Pos	.31
Hill et al. (2008)	335	100	G	4,5	EXT		PDI	20	L	V	Other	C	Pos	.18
Hill et al. (2008)	335	100	G	4,5	EXT		PDI	20	L	V	Other	C	Cont	.01
Hill et al. (2008)	335	100	G	4,5	EXT		Other	20	L	V	Other	C	Pos	.16
Hill et al. (2008)	335	100	G	4,5	EXT		Other	20	L	V	Other	C	Neg	.19
Hutchings et al. (2007)	151		G	3,9	EXT	30	PS	30	H	V	DPICS	PC	Neg	.57
Karreman, van Tuijl, van Aken, and Deković (2008)	89	100	G	3	G		PDI		H	V	Other	P	Pos	−.03
Karreman et al. (2008)	89	100	G	3	G		PDI		H	V	Other	P	Cont	.06
Karreman et al. (2008)	89	0	G	3	G		PDI		H	V	Other	P	Pos	−.07
Karreman et al. (2008)	89	0	G	3	G		PDI		H	V	Other	P	Cont	.16
Kochanska, Kim, and Koenig Nordling (2012)	186	100	G	2,5	G	13	Other	55	L	D	Other	C	Cont	.04

(*table continues*)

Table 1 (*continued*)

Authors	*N*	% mothers	Type sample parents	Age children	Type sample children	No. items	Quest.	Dur. obs.	Loc.	Direct	Obs.	Task	Type parenting	*r*
Linares, Montalto, Li, and Oza (2006)	128	87	G	6,2	M	15	PDI	20	H	D	Other	DI	Pos	.31
Lui, Johnston, Lee, and Lee-Flynn (2013)	94	100	G	9,8	ADHD		APQ	20	L	V	Other	P	Pos	.20
Metsäpelto & Pulkkinen (2005)	54	100	G	10,5	G	10	Other	20	L	V	Other	PS	Pos	.11
Metsäpelto & Pulkkinen (2005)	52	0	G	10,5	G	10	Other	20	L	V	Other	PS	Pos	.19
Morawska & Sanders (2007)	126	100	G	2,2	G	10	PS	30	H	V		PC	Neg	.34
Parent et al. (2014)	180	89	C	11,5	G		APQ	15	H	V	Other	DI	Pos	.22
Parent et al. (2014)	180	89	C	11,5	G		APQ	15	H	V	Other	DI	Neg	.27
Perry et al. (2013)	263	100	G	3	G	12	Other		L	V		P	Pos	.30
Perry et al. (2013)	244	100	G	4	G	12	Other		L	V		P	Pos	.32
Perry et al. (2013)	228	100	G	5	G	12	Other		L	V		P	Pos	.32
Scott, O'Connor, et al. (2010)	174		G	5,7	G	2	Other	23	H	V	Other	PC	Pos	−.02
Scott, O'Connor, et al. (2010)	174		G	5,7	G	2	Other	23	H	V	Other	PC	Pos	.09
Scott, O'Connor, et al. (2010)	174		G	5,7	G	2	Other	23	H	V	Other	PC	Cont	.13
Scott, Sylva, et al. (2010)	112		G	5,2	EXT		Other	15	H	V	Other	PC	Neg	.30
Scott, Sylva, et al. (2010)	112		G	5,2	EXT	5	Other	15	H	V	Other	PC	Neg	.22
Sessa, Avenevoli, Steinberg, and Morris (2001)	84	100	G	5,2	G	8	Other	25	H	V		PC	Pos	.03
Sessa et al. (2001)	84	100	G	5,2	G	5	Other	25	H	V		PC	Neg	−.06
Sessa et al. (2001)	84	100	G	5,2	G	5	Other	25	H	V		PC	Cont	−.02
Sheeber et al. (2012)	70	100	C	4,6	G		Other	20	H	V	Multiple	PC	Neg	−.02
Sheeber et al. (2012)	67	100	C	4,6	G		Other	20	H	V	Multiple	PC	Neg	.14
Tellegen & Sanders (2014)	64	94	G	5,7	ASD		PS	30	H	V	FOS		Neg	−.02
Van den Akker, Deković, M., Prinzie, and Asscher (2010)	96	100	G	2,5	G	8	Other	12	H	D	CII	PC	Pos	.10
Van den Akker et al. (2010)	96	100	G	2,5	G	6	PDI	12	H	D	CII	PC	Neg	.25
Van den Akker et al. (2010)	96	100	G	2,5	G	8	Other	12	H	V	Other	PC	Pos	.03
Van den Akker et al. (2010)	96	100	G	2,5	G	6	PDI	12	H	V	Other	PC	Neg	.28
Van den Akker et al. (2010)	96	100	G	3	G	8	Other	12	H	D	CII	PC	Pos	.06
Van den Akker et al. (2010)	96	100	G	3	G	6	PDI	12	H	D	CII	PC	Neg	.22
Van den Akker et al. (2010)	96	100	G	3	G	8	Other	12	H	V	Other	PC	Pos	.11
Van den Akker et al. (2010)	96	100	G	3	G	6	PDI	12	H	V	Other	PC	Neg	.31
Van den Akker et al. (2010)	96	100	G	3,3	G	8	Other	12	H	D	CII	PC	Pos	−.05
Van den Akker et al. (2010)	96	100	G	3,3	G	6	PDI	12	H	D	CII	PC	Neg	.16
Van den Akker et al. (2010)	96	100	G	3,3	G	8	Other	12	H	V	Other	PC	Pos	.16
Van den Akker et al. (2010)	96	100	G	3,3	G	6	PDI	12	H	V	Other	PC	Neg	.02
Van den Akker et al. (2010)	96	100	G	3,5	G	8	Other	12	H	D	CII	PC	Pos	−.04
Van den Akker et al. (2010)	96	100	G	3,5	G	6	PDI	12	H	D	CII	PC	Neg	.13
Van den Akker et al. (2010)	96	100	G	3,5	G	8	Other	12	H	V	Other	PC	Pos	−.05
Van den Akker et al. (2010)	96	100	G	3,5	G	6	PDI	12	H	V	Other	PC	Neg	.06

(*table continues*)

Table 1 (*continued*)

Authors	*N*	% mothers	Type sample parents	Age children	Type sample children	No. items	Quest.	Dur. obs.	Loc.	Direct	Obs.	Task	Type parenting	*r*
Waller et al. (2012)	731	100	G	2,5	G	10	PS	67	H	V	Multiple	PC	Neg	.07
Waller et al. (2012)	731	100	G	3,5	G	10	PS	67	H	V	Multiple	PC	Neg	.20
Weis & Lovejoy (2002)	99	100	G	3,3	G	10	Other	15	L	V	Other	P	Pos	.30
Weis & Lovejoy (2002)	99	100	G	3,3	G	10	Other	15	L	V	Other	P	Neg	.39
Weis & Lovejoy (2002)	99	100	G	3,3	G	10	Other	15	L	V	Other	PS	Pos	.47
Weis & Lovejoy (2002)	99	100	G	3,3	G	10	Other	15	L	V	Other	PS	Neg	.59
Winter, Morawska, and Sanders (2012)	62	97	G	2,5	G	9	Other	15	H	V	FOS	PC	Pos	.07
Wymbs (2011)	99	50	G	11	ADHD	10	Other	25	L	V	FOS	PC	Pos	.07
Wymbs (2011)	90	50	G	11	ADHD	10	Other	25	L	V	FOS	PC	Neg	.19
Zaslow et al. (2006)	278	100	G	4,1	G	4	Multiple	−	H	V		PS	Pos	.05

Note. *N*= sample size in the studies; % mothers = percentage of mothers in the sample; Type sample parents = which type of parents the sample consisted of (G = general sample, C = clinical); Age children = mean age of the children in the sample in years; Type sample children = which type of children the sample consisted of (G = general, M = with a maltreatment history, ADHD = children with attention-deficit hyperactivity disorder, EXT = children with externalizing behavior problems, ASD = children with an autism spectrum disorder); No. items = questionnaire length for the parenting construct of interest; Quest. = which questionnaire was administered (PDI = Parenting Dimensions Inventory, PPI = Parenting Practices Interview, APQ = Alabama Parenting Questionnaire, PS = Parenting Scale, Multiple = scores of multiple questionnaires combined, Other = another parenting questionnaire); Dur. obs. = duration of the observation in minutes; Loc. = where the observation took place (H = home, L = laboratory); Direct = whether the observation was performed directly of from a videotape; Obs. = which observation instrument was used (CII = Coders Impression Inventory, DPICS = Dyadic Parent-Child Interaction Coding System, FOS = Family Observation System, Multiple = a combination of multiple observation systems, Other = another observation instrument); Task = which type of task parents and children had to participate in (C = compliance, D = discussion, HW = homework, *p* = play, PC = play and clean up, PS = problem solving, Other = another type of task); Type of parenting = which type of parenting was measured (Pos = positive, Neg = negative, Cont = controlling); *r* = the strength of the observed correlation between parent-reported parenting and observed parenting; − indicates negative observed correlations between parent-reported and observed parenting. Negative correlations did not appear to be caused by differences in meaning of high and low scores on the parenting construct.

variables observation duration, number of positive items, and sample size; their values were, respectively, .75, .82, and .85. For categorical variables, we calculated Cohen's kappas. The average of the Cohen's kappas measuring correspondence between the coders was .92 on the categorical variables. The lowest agreement between the two raters was for the variables observation location and sample type parents; their values were, respectively, .74 and.77. For the other categorical values, the raters fully agreed.

Effect Sizes

In order to make generalizations beyond the articles, we chose to use random effect methods for the analyses, yielding a more conservative significance test for combined effects (Field, 2001). Furthermore, it enabled us to estimate the distribution of effect sizes across studies (Thompson & Higgins, 2002). To investigate whether particular covariates or moderators accounted for the heterogeneity of effect sizes between studies, metaregression was used. To summarize the strength of the associations, Pearson's correlation coefficient *r*, and Fisher's *z* were either obtained or calculated for all the studies. Fisher's *z* is used to correct for skewness in the sampling distribution of *r* (Field, 2001). Analyses were performed with the Fisher's *z* corrected effect sizes notated as ES_z. After analyses the outcomes for the significant models were transformed into correlations (ρ) for interpretation purposes.

Publication Bias

Because for most studies the associations between parent-reported and observed parenting were not the primary research interest, we assumed that there was a relatively low risk for publication bias. To check this assumption, we inspected a funnel plot using Fisher's *z* transformations, as displayed in Figure 1. The distribution of the effect sizes was symmetrical with the effect sizes of the studies with a larger sample size spread out more narrowly around the estimated overall effect size compared with the studies with a smaller sample size. This demonstrated that there was no bias in the correlations (Lipsey & Wilson, 2001). In addition, we tested asymmetry using Egger's regression test (Egger, Davey Smith, Schneider, & Minder, 1997), which was not

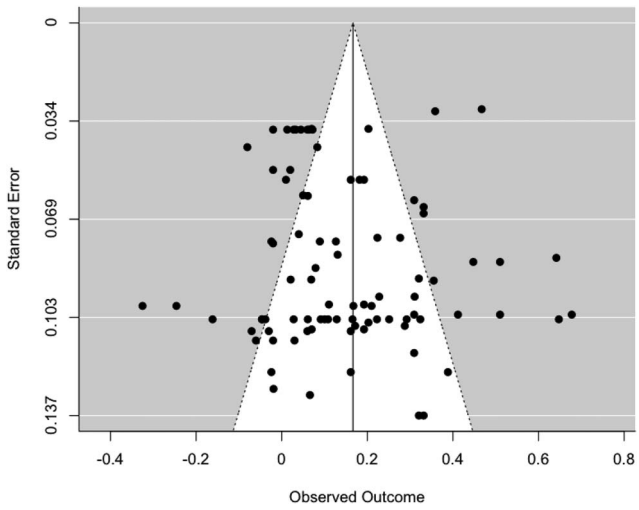

Figure 1. Funnel plot with Fisher's *z* transformed correlations. On the *y*-axis are the standard errors of the studies, with smaller standard errors representing larger sample sizes. On the *x*-axis are the associations between parent-reported and observed parenting.

significant ($z = 0.88$, $p = .38$) and therefore provided additional evidence that there was no publication bias present in this meta-analysis.

Analyses

Many of the included studies contained multiple effect sizes for different groups or conditions and therefore there was dependence between effect sizes. To account for this dependence, we applied a multilevel approach (Hedges, Tipton, & Johnson, 2010) with a level of sampling variance around the estimated population effect size (Level 1), a level of variance between effect sizes within studies (Level 2), and a level of variance between effect sizes across studies (Level 3; Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2013; Wibbelink & Assink, 2015). The estimation of the parameters was performed with a restricted maximum likelihood estimation method (Wibbelink & Assink, 2015). The analyses were performed in R using the metafor package (Viechtbauer, 2010). First, an overall effect size was estimated. Second, to decide whether the variance between effect sizes within studies and the variance between effect sizes were significant, two log-likelihood ratio tests were performed to see whether constraining the variances of level 2 and level 3 to zero significantly deteriorated model fit. Third, we investigated how the variance was distributed across the three levels using the Higgins and Thompson method (Higgins & Thompson, 2002) as proposed by Cheung (2014), showing how much of the variance was a result of sampling variability (Level 1), within-study variance (Level 2), and between-study variance (Level 3).

Next, we performed separate moderator analyses. For the categorical moderators we created dummy variables. We investigated which moderators significantly explained heterogeneity of the effect sizes with univariate analyses. These moderators were included in the final model. To establish each moderator's unique contribution to the model, we performed a multivariate analysis with a model including all moderators that were significant in the univariate analyses. Finally, the proportion of explained variance obtained by adding the moderators was calculated using the method as described in Cheung (2014).

Results

Overall Effect Size

The estimate of the true effect size was .17, which significantly deviated from zero ($SE_z = .03$, $t(88) = 5.88$, $p < .01$, $\rho = .17$, 95% CI [.11, .22]). This indicated that overall, there was a significant but weak association between parent-reported and observed parenting. The Q statistic was significant, $Q(88) = 590.57$, $p < .01$, indicating that there was significant heterogeneity in the effect sizes not accounted for by this main effects model. The Level 2 variance, or the variance of effect sizes within studies, was .01. Constraining this variance to zero significantly deteriorated model fit, $\chi^2(1) = 18.09$, $p < .01$, indicating that the Level 2 variance did significantly deviate from zero. The Level 3 variance, or the variance between the effect sizes, was .02. Constraining this variance to zero also significantly degraded model fit, $\chi^2(1) = 22.49$, $p < .01$, indicating that the Level 3 variance did significantly deviate from zero. Next, we calculated how the total variance was divided across the three levels. The percentage of variance at Level 1, or sampling variance of the effect sizes around the mean estimate, was 14.06%. The percentage of variance at Level 2 was 26.25%, indicating that around 26% of the variance in effect sizes could be attributed to differences between effect sizes within studies. The percentage of variance at Level 3 was 65.15%, indicating that 65% of the total variance in associations in this study could be explained by differences between studies. Because the heterogeneity coefficient Q—containing the heterogeneity at Levels 1, 2, and 3—in our model was significant, the next step was to perform univariate moderator analyses to investigate which moderator variables significantly explained heterogeneity of effect sizes across studies.

Participant Characteristics

The Q_M statistics for whether the moderators were significant, and the Q_E statistics for whether there was significant residual heterogeneity are displayed in Table 2. None of the examined participant characteristics (i.e., percentage of mothers, type of sample parents, age of the children, and type of sample children) significantly moderated the association between parent-reported parenting and observed parenting.

Questionnaire Characteristics

In Table 2, the Q_M statistics for whether the moderators were significant, and the Q_E statistics for whether there was significant residual heterogeneity are displayed. Which questionnaire was used was not a significant moderator. Questionnaire length was a significant moderator, indicating that the amount of items had a significant overall effect on the strength of the association between parent-reported and observed parenting. The amount of residual heterogeneity was significant. The coefficients for this model are displayed in Table 3. The intercept for zero items did not significantly deviate from zero. The slope for more items was positive and significant, indicating that the association between parent-reported increased significantly per increase in questionnaire length.

Table 2
All the Q_M Statistics for Whether the Moderators Have a Significant Effect, and the Q_E Statistics for Whether There Is Residual Heterogeneity

Moderator	Number of studies	Q_M (*df*)	*p*	Q_E (*df*)	*p*
% mothers	32	.05 (1)	.83	503.95 (79)	<.01
Type sample parents	36	.47 (1)	.83	586.65 (87)	<.01
Age children	35	.00 (1)	.95	485.43 (85)	<.01
Type sample children	36	.77 (4)	.55	580.61 (84)	<.01
Questionnaire length	27	4.44 (1)	.04	403.83 (72)	<.01
Which questionnaire	36	1.57 (5)	.18	567.28 (83)	<.01
Duration observation	31	.05 (1)	.83	313.33 (75)	<.01
Location	35	.68 (1)	.41	543.57 (85)	<.01
Direct vs. video	35	2.05 (1)	.16	390.81 (85)	<.01
Which observation instrument	32	.46 (4)	.77	541.64 (76)	<.01
Type of task	34	.58 (7)	.77	363.01 (78)	<.01
Type of parenting	36	3.34 (2)	.04	548.03 (86)	<.01

Table 3
Coefficients for the Univariate Moderator Analysis With Questionnaire Length

Parameter estimates	ES_z	*SE*	*t*	95% CI	*p*	ρ
Intercept (items = 0)	.09	.05	1.71	[−.02, .19]	.09	.09
Slope	.01	.00	2.11	[.00, .02]	.04	.01

Note. Number of studies = 28. ES = effect size; *SE* = standard error; CI = confidence interval.

Observation Characteristics

In Table 2, the Q_M statistics for whether the moderators were significant, and the Q_E statistics for whether there was significant residual heterogeneity are displayed. Type of observation instrument used, duration and location of the observation, whether the observation was scored directly or from a videotape, and which task was used were not found to significantly moderate the association between parent reported and observed parenting.

Parenting

Which type of parenting was reported and observed *was* a significant moderator, indicating that there was an overall significant effect of the type of parenting. The amount of residual heterogeneity of the model was significant. The coefficients for the moderator are displayed in Table 4. The mean effect sizes for positive and negative parenting were positive and significant, indicating that for studies reporting either positive or negative parenting, there was a small to medium association between parent-reported and observed parenting. On the contrary, the mean effect size for controlling parenting did not significantly deviate from zero, indicating that in studies that assessed controlling parenting there was no significant association between parent-reported and observed parenting. In short, for studies that assessed positive or negative parenting behavior, there was a small to medium significant association between parent-reported and observed parenting.

Model With Significant Moderators

The moderators that were significant in the univariate analyses were questionnaire length and which type of parenting was reported and observed. Next, we analyzed a model including the moderators that emerged as significant in the univariate analyses (i.e., type of parenting and type of questionnaire used). The effect of the moderators in this final model was significant, Q_M (3) = 4.21, $p = .01$, indicating that there was an effect of which questionnaire and which type of parenting were included on the strength of the association between parent-reported parenting and observed parenting. The amount of residual heterogeneity was significant, Q_E (70) = 379.66, $p < .01$, indicating that there was heterogeneity of the effect sizes not accounted for by the model. Figure 2 displays a forest plot of the final model in which the residual heterogeneity is visualized. The mean effect sizes for this model are presented in Table 5. In combination with the intercept for questionnaire length, the mean effect size for negative was small, but significant. For positive, negative, and controlling parenting, the slope for items was significant, indicating that for all types of parenting, the association between parent-reported and observed parenting increased significantly with increase in questionnaire length.

The Level 2 variance of the model with the significant moderators was .01. Constraining this variance to zero significantly deteriorated model fit, $\chi^2(1) = 13.78$, $p < .01$, indicating that the Level 2 variance significantly deviated from zero. The Level 3 variance was .02. Constraining this variance to zero significantly degraded model fit, $\chi^2(1) = 15.66$, $p < .01$, indicating that the Level 3 variance significantly deviated from zero. Next, we calculated how the total variance was divided across the three levels. The percentage of variance at Level 1, or the sampling variance around the mean estimate, was 15.20%. The percentage of variance at Level 2 was 30.47%, indicating that around 30% of the variance could be assigned to differences between effect sizes within studies. The percentage of variance at Level 3 was 61.44%, indicating that 61% of the total variance could be explained by differences between studies. The increase in the proportion of Level 2 variance of the model with the moderators compared with the model without moderators yielded a negative proportion. According to Cheung (2014), negative values are truncated to zero, indicating that adding the moderators did not explain more of the variance between studies. However, calculating the proportion of Level 3 variance explained by the moderators was .13%, indicating that adding the moderators explained about 38% of the variance within studies.

Discussion

The purpose of the present meta-analysis was to investigate the strength of the association between parent-reported and observed parenting, and to investigate which variables would moderate this association. Overall, we found an estimate of the true effect size of .17, indicating that parent-reported parenting was weakly associated with observed parenting. Moderator analyses demonstrated that this effect size was dependent on questionnaire length, and on the type of parenting measured. Specifically, effect sizes were stronger (i.e., more positive) for studies that used questionnaires with more items, and effect sizes were stronger for studies that assessed negative parenting, compared with studies that assessed controlling parenting.

The small positive significant association indicates that if a parent scores higher on a parent-reported assessment of parenting, this parent will probably also score higher on an observational

Table 4
Coefficients for the Univariate Moderator Analysis With Which Type of Parenting Was Measured

Parameter estimates	ES_z	*SE*	*t*	95% CI	*p*	ρ
Positive	.15	.03	5.03	[.09, .21]	<.01	.15
Negative	.21	.03	6.76	[.16, .29]	<.01	.22
Controlling	.09	.05	1.69	[−.02, .19]	.09	.09
Slope pos-neg	.06	.03	1.92	[−.00, .11]	.06	.05
Slope pos-cont	−.08	.05	−1.42	[−.18, .03]	.16	−.08
Slope neg-cont	−.13	.06	−2.31	[−.24, −.02]	.02	−.13

Note. Number of studies = 37. ES = effect size; *SE* = standard error; CI = confidence interval.

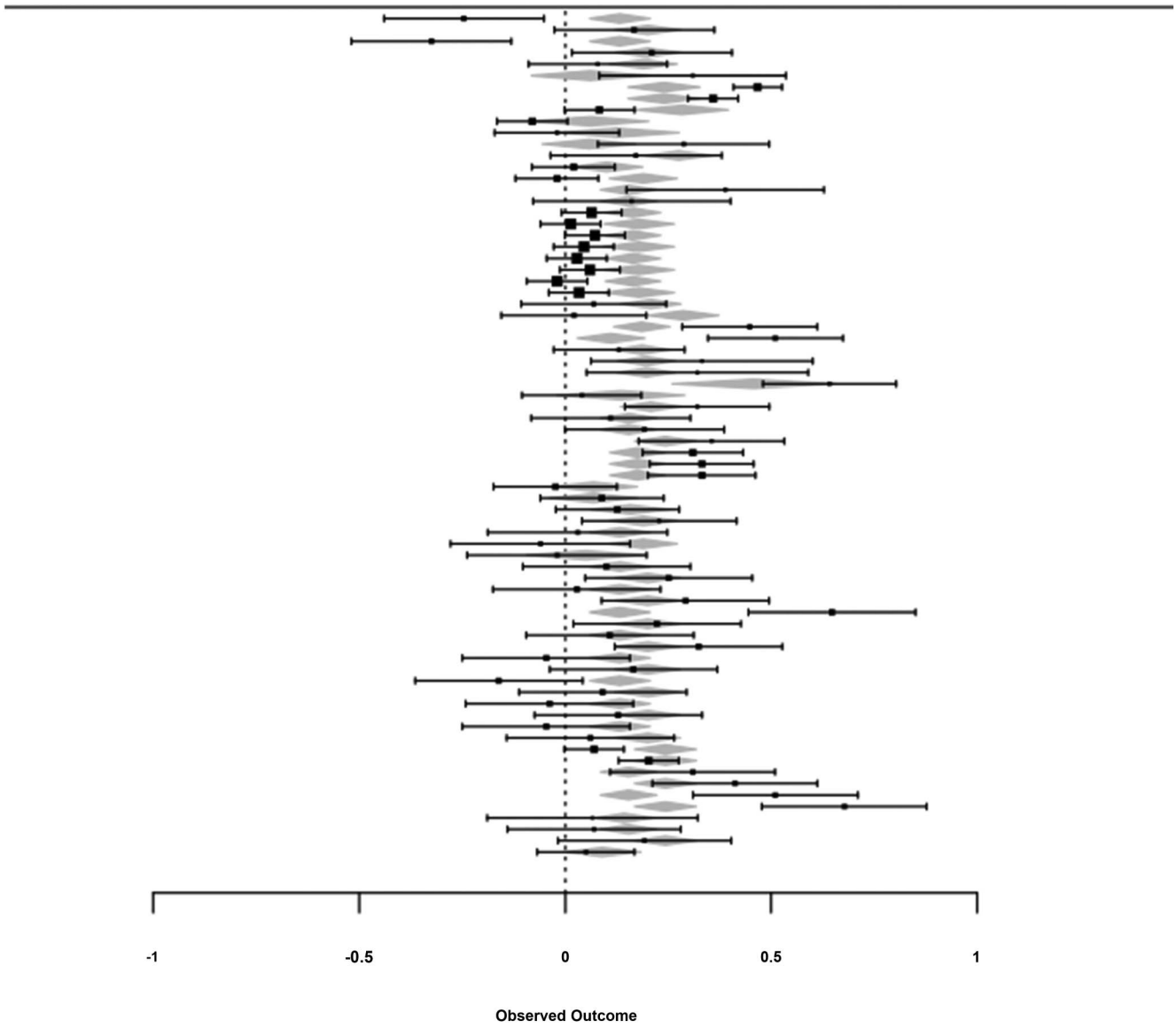

Figure 2. Forest plot of the final model. The gray areas represent the values of the association based on the values on the moderators, the black stripes represent the observed associations.

assessment, and vice versa. Furthermore, the association provides support for the assumption that parent reports and observations of parenting measure similar underlying constructs (Gardner, 2000).

Table 5
Coefficients for the Final Model

Parameter estimates	ES_z	*SE*	*t*	95% CI	*p*	ρ
Mean positive (items = 0)	.05	.06	.78	[−.07, .17]	.44	.05
Slope items positive	.01	.00	2.45	[.00, .02]	.02	.01
Mean negative (items = 0)	.14	.05	2.52	[.03, .24]	.01	.14
Slope items negative	.01	.00	2.45	[.00, .02]	.02	.01
Mean controlling (items = 0)	−.00	.08	−.05	[−.16, .15]	.96	−.00
Slope items controlling	.01	.00	2.45	[.00, .02]	.02	.01

Note. ES = effect size; *SE* = standard error; CI = confidence interval.

However, the association is weak, indicating that there are more variables explaining variance in parent-reported or observed parenting. For instance, the weak association may indicate that parent reports and observations have different functions (Gardner, 2000), and that they can still assess the same underlying construct but different aspects of this construct (i.e., perceptions and attitudes vs. visible behaviors). Moreover, as in similar research on children (Achenbach et al., 1987, 2005), parents and observers are exposed to different situations when reporting on the subject and construct of interest, which may cause parents and observers to perceive behaviors differently. In addition, the weak association could indicate that scores from parent reports and observations are influenced differently by social desirability. Within self-reports, positive biases can alter how parents report on their own behavior (e.g., Bögels & van Melick, 2004; Morsbach & Prinz, 2006;

Schwarz et al., 1985; Waylen et al., 2008). Similarly, during an observation parents may behave more socially desirable due to the presence of an observer (e.g., Aspland & Gardner, 2003; Jacob, Tennenbaum, Seilhamer, Bargiel, & Sharon, 1994). In other words, this study cannot speak to what the association between parent-reported and observed parenting means. Based on the arguments presented above and on the strength of the association found in this study, we conclude that parent reports and observations of parenting are weakly associated.

In addition to establishing the overall strength of the association between parent-reported and observed parenting, this study also aimed to investigate which study characteristics—participant, questionnaire, observation, or parenting—moderated the strength of the association between parent-reported and observed parenting. Overall, few moderators were significant, resulting in a robust mean effect size. More specifically, moderators indicative of methodological rigor or study quality (e.g., observation duration, direct vs. video observation, observation location) were not significant, resulting in a robust mean effects size across study quality. Of the tested moderators, two moderators did significantly influence the association between parent-reported and observed parenting, namely, questionnaire length and type of parenting.

For questionnaire length the association between parent-reported and observed parenting was stronger for studies using questionnaires with more items. As mentioned in the introduction, according to psychometric theory, a greater number of items often leads to increased reliability (e.g., Cronbach, 1951; Smith, McCarthy, & Anderson, 2000). Furthermore, previous studies provided evidence that aggregating measures of parenting improved stability and representativeness of scores (e.g., Rushton, Brainerd, & Pressley, 1983; Wachs, 1987). If the questionnaires provided a more reliable and valid representation of parenting, it is also likely that the rating of parenting is more consistent across situations and informants. Finally, if the questionnaires provide a better rating of parenting, it is likely that their scores are less susceptible to the influence of social desirability and therefore have less nuances in the assessment of parenting, allowing for detection of a stronger association between parent-reported and observed parenting.

In addition to questionnaire length, the type of parenting assessed also moderated the strength of the association between parent-reported and observed parenting. Specifically, the association was stronger for negative parenting compared with controlling behavior. This finding indicates that mechanisms for measuring parenting are different for positive, negative, or controlling parenting. An explanation could be that negative behaviors are more salient for parents and observers (e.g., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Lovejoy, Weis, O'Hare, & Rubin, 1999), and positive and negative parenting are more stable across situations whereas the expression of controlling parenting is somewhat dependent on interactional sequences with the child (Hill, Maskowitz, Danis, & Wakschlag, 2008). Therefore, for negative parenting, the self-reported perceptions and attitudes are possibly more in line with the visible, observable behaviors. Similarly, it could be that for more salient parenting constructs there is less situational specificity for the different informants, causing parents and observers to perceive the situations less differently compared with the less salient parenting behaviors. Furthermore, the influence of social desirability could be similar for parent reports and observations of these parenting behaviors. For example, parents may report less negative and more positive behaviors, and during observation display less negative and more positive behaviors because of a positive bias and altered behavior due to the presence of an observer.

Strengths and Limitations

The main strength of this study is that it, as a first, provides a mean estimate of the effect size for the association between parent-reported and observed parenting—reviewing 15 years of research on parenting. Although previously scholars emphasized that what parents report is not the same as how they actually parent (e.g., Bornstein, Cote, & Venuti, 2001; Locke & Prinz, 2002), until now the mean strength of the association remained unknown. Moreover, the knowledge that the association is stronger for certain questionnaires and parenting behaviors may provide some insight into the mechanisms that underlie the association between parent-reported and observed parenting. As we have previously described, the association between parent-reported and observed parenting was stronger for more salient types of parenting behaviors. Differences between purposes of measurement instruments, informants, and influences of social desirability may become less important when assessing certain types of parenting, which could indicate that these behaviors are detected more accurately and stably. A second strength of the present study was the use of a three-level analysis model that corrected for the dependency between correlations reported within and across studies. By taking the dependency of effect sizes into account, the liability for a Type I error (i.e., a false positive finding) improves (Van den Noortgate et al., 2013).

Several limitations of the present study also warrant mentioning. First, a weakness of using meta-analytical techniques is that it uses summary statistics, and therefore nuances might get lost that would have been detected using raw data. Second, the outcomes do not explain what the association between parent-reported and observed parenting actually entails. The strength of the association does indicate that agreement between parents and observers on parenting is weak; this needs to be taken into account when researchers or clinicians want to assess parenting. Nevertheless, this study has provided a starting point for future research by exploring the strength of the association between parent-reported and observed parenting based on 15 years of research.

Future Recommendations and Implications

Although this study provides insight into the strength of the association between parent-reported and observed parenting, it does not—nor did it intend to—provide insight in "which measure is best to use." A suggestion for future research to gain more insight in which measure of parenting is best to use is to investigate which methods are most strongly associated with outcome variables. For instance, in research on child anxiety, several meta-analyses have shown that certain parenting dimensions are more strongly associated with child anxiety disorders (McLeod, Wood, & Weisz, 2007; Möller, Nikolić, Majdandžić, & Bögels, 2016; Van der Brugge, Stams, & Bögels, 2008), and that observed parenting is more strongly associated with child anxiety disorders than parent reports of parenting (McLeod et al., 2007). Although that would not provide insight into which measure provides the

most accurate representation of parenting, it would provide insight in which measure is best to use for research on certain outcome variables.

This study takes a first step to investigate the strength of the association between parent-reported and observed parenting, and which moderators influence this strength. One main conclusion of our study is that except for which type of parenting was assessed and questionnaire length, none of the other participant, questionnaire, or observation characteristics that we examined moderated the main effect size. This indicates that the significant-but-small association between parent-reported and observed parenting is a robust finding. The significant portion of unexplained heterogeneity between study findings emphasizes the need for future research in this domain. Awaiting these future endeavors, what has become clear on the basis of our current findings is that the association between parent-reported and observed parenting is relatively weak and varies across types of parenting and parenting questionnaires. Researchers and clinicians should be aware of this weak association when choosing a method to assess parenting.

References

References marked with an asterisk were included in the analyses.

Achenbach, T. M., Krukowski, R. A., Dumenci, L., & Ivanova, M. Y. (2005). Assessment of adult psychopathology: Meta-analyses and implications of cross-informant correlations. *Psychological Bulletin, 131,* 361–382. http://dx.doi.org/10.1037/0033-2909.131.3.361

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213–232. http://dx.doi.org/10.1037/0033-2909.101.2.213

Aspland, H., & Gardner, F. (2003). Observational measures of parent-child interaction: An introductory review. *Child and Adolescent Mental Health, 8,* 136–143. http://dx.doi.org/10.1111/1475-3588.00061

*Asscher, J. J., Hermanns, J. M. A., & Deković, M. (2008). Effectiveness of the Home-Start parenting support program: Behavioral outcomes for parents and children. *Infant Mental Health Journal, 29,* 95–113. http://dx.doi.org/10.1002/imhj.20171

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Is bad stronger than good. *Review of General Psychology, 5,* 323–370. http://dx.doi.org/10.1037/1089-2680.5.4.323

Baumrind, D. (1971). Current patterns of parental authority. *Developmental Psychology, 4,* 1–103. http://dx.doi.org/10.1037/h0030372

*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. http://dx.doi.org/10.1177/1077559505283589

*Bhandari, R. P., & Barnett, D. (2007). Restrictive parenting buffers Head Start students from stress. *Infants & Young Children, 20,* 55–63. http://dx.doi.org/10.1097/00001163-200701000-00006

*Blair, C., Raver, C. C., & Berry, D. J., & the Family Life Project Investigators. (2014). Two approaches to estimating the effect of parenting on the development of executive function in early childhood. *Developmental Psychology, 50,* 554–565. http://dx.doi.org/10.1037/a0033647

Bögels, S., & Phares, V. (2008). Fathers' role in the etiology, prevention and treatment of child anxiety: A review and new model. *Clinical Psychology Review, 28,* 539–558. http://dx.doi.org/10.1016/j.cpr.2007.07.011

Bögels, S. M., & van Melick, M. (2004). The relationship between child-report, parent self-report, and partner report of perceived parental rearing behaviors and anxiety in children and parents. *Personality and Individual Differences, 37,* 1583–1596. http://dx.doi.org/10.1016/j.paid.2004.02.014

Bornstein, M. H., Cote, L. R., & Venuti, P. (2001). Parenting beliefs and behaviors in northern and southern groups of Italian mothers of young infants. *Journal of Family Psychology, 15,* 663–675. http://dx.doi.org/10.1037/0893-3200.15.4.663

*Breitenstein, S. M., Gross, D., Fogg, L., Ridge, A., Garvey, C., Julion, W., & Tucker, S. (2012). The Chicago Parent Program: Comparing 1-year outcomes for African American and Latino parents of young children. *Research in Nursing & Health, 35,* 475–489. http://dx.doi.org/10.1002/nur.21489

*Brotman, L. M., Calzada, E., Huang, K. Y., Kingston, S., Dawson-McClure, S., Kamboukos, D., . . . Petkova, E. (2011). Promoting effective parenting practices and preventing child behavior problems in school among ethnically diverse families from underserved, urban communities. *Child Development, 82,* 258–276. http://dx.doi.org/10.1111/j.1467-8624.2010.01554.x

*Brotman, L. M., Gouley, K. K., Chesir-Teran, D., Dennis, T., Klein, R. G., & Shrout, P. (2005). Prevention for preschoolers at high risk for conduct problems: Immediate outcomes on parenting practices and child social competence. *Journal of Clinical Child and Adolescent Psychology, 34,* 724–734. http://dx.doi.org/10.1207/s15374424jccp3404_14

*Browne, D. T., Meunier, J. C., O'Connor, T. G., & Jenkins, J. M. (2012). The role of parental personality traits in differential parenting. *Journal of Family Psychology, 26,* 542–553. http://dx.doi.org/10.1037/a0028515

Calzada, E. J., Eyberg, S. M., Rich, B., & Querido, J. G. (2004). Parenting disruptive preschoolers: Experiences of mothers and fathers. *Journal of Abnormal Child Psychology, 32,* 203–213. http://dx.doi.org/10.1023/B:JACP.0000019771.43161.1c

Campbell, S. B. (1997). Behavior problems in preschool children: Developmental and family issues. *Advances in Clinical Child Psychology, 19,* 1–26. http://dx.doi.org/10.1007/978-1-4757-9035-1_1

Caron, A., Weiss, B., Harris, V., & Catron, T. (2006). Parenting behavior dimensions and child psychopathology: Specificity, task dependency, and interactive relations. *Journal of Clinical Child and Adolescent Psychology, 35,* 34–45. http://dx.doi.org/10.1207/s15374424jccp3501_4

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods, 19,* 211–229. http://dx.doi.org/10.1037/a0032968

*Chronis-Tuscano, A., Raggi, V. L., Clarke, T. L., Rooney, M. E., Diaz, Y., & Pian, J. (2008). Associations between maternal attention-deficit/hyperactivity disorder symptoms and parenting. *Journal of Abnormal Child Psychology, 36,* 1237–1250. http://dx.doi.org/10.1007/s10802-008-9246-4

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334. http://dx.doi.org/10.1007/BF02310555

Duhig, A. M., Renk, K., Epstein, M. K., & Phares, V. (2000). Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science and Practice, 7,* 435–453. http://dx.doi.org/10.1093/clipsy.7.4.435

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315,* 629–634. http://dx.doi.org/10.1136/bmj.315.7109.629

*Feinberg, M., Neiderhiser, J., Howe, G., & Hetherington, E. M. (2001). Adolescent, parent, and observer perceptions of parenting: Genetic and environmental influences on shared and distinct perceptions. *Child Development, 72,* 1266–1284. http://dx.doi.org/10.1111/1467-8624.00346

Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects methods. *Psychological Methods, 6,* 161–180. http://dx.doi.org/10.1037/1082-989X.6.2.161

*Fossum, S., Mørch, W. T., Handegård, B. H., & Drugli, M. B. (2007). Childhood disruptive behaviors and family functioning in clinically referred children: Are girls different from boys? *Scandinavian Journal*

of Psychology, 48, 375–382. http://dx.doi.org/10.1111/j.1467-9450.2007.00617.x

*Friesen, M. D., Woodward, L. J., Horwood, L. J., & Fergusson, D. M. (2013). Quality of parent-child relations in adolescence and later adult parenting outcomes. *Social Development, 22,* 539–554. http://dx.doi.org/10.1111/j.1467-9507.2012.00657.x

Gardner, F. (2000). Methodological issues in the direct observation of parent–child interaction: Do observational findings reflect the natural behavior of participants? *Clinical Child and Family Psychology Review, 3,* 185–198.

*Hahlweg, K., Heinrichs, N., Kuschel, A., Bertram, H., & Naumann, S. (2010). Long-term outcome of a randomized controlled universal prevention trial through a positive parenting program: Is it worth the effort? *Child and Adolescent Psychiatry and Mental Health, 4,* 14. http://dx.doi.org/10.1186/1753-2000-4-14

*Hanisch, C., Hautmann, C., Plück, J., Eichelberger, I., & Döpfner, M. (2014). The prevention program for externalizing problem behavior (PEP) improves child behavior by reducing negative parenting: Analysis of mediating processes in a randomized controlled trial. *Journal of Child Psychology and Psychiatry, 55,* 473–484. http://dx.doi.org/10.1111/jcpp.12177

*Hawes, D. J., & Dadds, M. R. (2006). Assessing parenting practices through parent-report and direct observation during parent-training. *Journal of Child and Family Studies, 15,* 554–567. http://dx.doi.org/10.1007/s10826-006-9029-x

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1,* 39–65. http://dx.doi.org/10.1002/jrsm.5

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21,* 1539–1558. http://dx.doi.org/10.1002/sim.1186

*Hill, C., Maskowitz, K., Danis, B., & Wakschlag, L. (2008). Validation of a clinically sensitive, observational coding system for parenting behaviors: The Parenting Clinical Observation Schedule. *Parenting: Science and Practice, 8,* 153–185. http://dx.doi.org/10.1080/15295190802045469

*Hutchings, J., Gardner, F., Bywater, T., Daley, D., Whitaker, C., Jones, K., . . . Edwards, R. T. (2007). Parenting intervention in Sure Start services for children at risk of developing conduct disorder: Pragmatic randomised controlled trial. *British Medical Journal, 334,* 678. http://dx.doi.org/10.1136/bmj.39126.620799.55

Jacob, T., Tennenbaum, D., Seilhamer, R. A., Bargiel, K., & Sharon, T. (1994). Reactivity effects during naturalistic observation of distressed and nondistressed families. *Journal of Family Psychology, 8,* 354–363. http://dx.doi.org/10.1037/0893-3200.8.3.354

*Karreman, A., van Tuijl, C., van Aken, M. A. G., & Deković, M. (2008). Parenting, coparenting, and effortful control in preschoolers. *Journal of Family Psychology, 22,* 30–40. http://dx.doi.org/10.1037/0893-3200.22.1.30

*Kochanska, G., Kim, S., & Koenig Nordling, J. (2012). Challenging circumstances moderate the links between mothers' personality traits and their parenting in low-income families with young children. *Journal of Personality and Social Psychology, 103,* 1040–1049. http://dx.doi.org/10.1037/a0030386

Kuppens, S., Grietens, H., Onghena, P., & Michiels, D. (2009). Measuring parenting dimensions in middle childhood. *European Journal of Psychological Assessment, 25,* 133–140. http://dx.doi.org/10.1027/1015-5759.25.3.133

Lecuyer-Maus, E. A. (2000). Maternal sensitivity and responsiveness, limit-setting style, and relationship history in the transition to toddlerhood. *Issues in Comprehensive Pediatric Nursing, 23,* 117–139. http://dx.doi.org/10.1080/01460860050121439

*Linares, L. O., Montalto, D., Li, M., & Oza, V. S. (2006). A promising parenting intervention in foster care. *Journal of Consulting and Clinical Psychology, 74,* 32–41. http://dx.doi.org/10.1037/0022-006X.74.1.32

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Locke, L. M., & Prinz, R. J. (2002). Measurement of parental discipline and nurturance. *Clinical Psychology Review, 22,* 895–929. http://dx.doi.org/10.1016/S0272-7358(02)00133-2

Lovejoy, M. C., Weis, R., O'Hare, E., & Rubin, E. C. (1999). Development and initial validation of the Parent Behavior Inventory. *Psychological Assessment, 11,* 534–545.

*Lui, J. H. L., Johnston, C., Lee, C. M., & Lee-Flynn, S. C. (2013). Parental ADHD symptoms and self-reports of positive parenting. *Journal of Consulting and Clinical Psychology, 81,* 988–998. http://dx.doi.org/10.1037/a0033490

McBride, B. A., Schoppe, S. J., & Rane, T. R. (2002). Child characteristics, parenting stress, and parental involvement: Fathers versus mothers. *Journal of Marriage and Family, 64,* 998–1011. http://dx.doi.org/10.1111/j.1741-3737.2002.00998.x

McKee, L., Colletti, C., Rakow, A., Jones, D. J., & Forehand, R. (2008). Parenting and child externalizing behaviors: Are the associations specific or diffuse? *Aggression and Violent Behavior, 13,* 201–215. http://dx.doi.org/10.1016/j.avb.2008.03.005

McLeod, B. D., Wood, J. J., & Weisz, J. R. (2007). Examining the association between parenting and childhood anxiety: A meta-analysis. *Clinical Psychology Review, 27,* 155–172. http://dx.doi.org/10.1016/j.cpr.2006.09.002

*Metsäpelto, R. L., & Pulkkinen, L. (2005). The moderating effect of extraversion on the relation between self-reported and observed parenting. *Journal of Applied Developmental Psychology, 26,* 371–384. http://dx.doi.org/10.1016/j.appdev.2005.04.005

Möller, E. L., Nikolić, M., Majdandžić, M., & Bögels, S. M. (2016). Associations between maternal and paternal parenting behaviors, anxiety and its precursors in early childhood: A meta-analysis. *Clinical Psychology Review, 45,* 17–33. http://dx.doi.org/10.1016/j.cpr.2016.03.002

*Morawska, A., & Sanders, M. R. (2007). Are parent-reported outcomes for self-directed or telephone-assisted behavioral family intervention enhanced if parents are observed? *Behavior Modification, 31,* 279–297. http://dx.doi.org/10.1177/0145445506293784

Morsbach, S. K., & Prinz, R. J. (2006). Understanding and improving the validity of self-report of parenting. *Clinical Child and Family Psychology Review, 9,* 1–21. http://dx.doi.org/10.1007/s10567-006-0001-5

*Parent, J., Forehand, R., Dunbar, J. P., Watson, K. H., Reising, M. M., Seehuus, M., & Compas, B. E. (2014). Parent and adolescent reports of parenting when a parent has a history of depression: Associations with observations of parenting. *Journal of Abnormal Child Psychology, 42,* 173–183. http://dx.doi.org/10.1007/s10802-013-9777-1

*Perry, N. B., Nelson, J. A., Swingler, M. M., Leerkes, E. M., Calkins, S. D., Marcovitch, S., & O'Brien, M. (2013). The relation between maternal emotional support and child physiological regulation across the preschool years. *Developmental Psychobiology, 55,* 382–394. http://dx.doi.org/10.1002/dev.21042

Phares, V. (1992). Where's poppa? The relative lack of attention to the role of fathers in child and adolescent psychopathology. *American Psychologist, 47,* 656–664. http://dx.doi.org/10.1037/0003-066X.47.5.656

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94,* 18–38. http://dx.doi.org/10.1037/0033-2909.94.1.18

Schwarz, J. C., Barton-Henry, M. L., & Pruzinsky, T. (1985). Assessing child-rearing behaviors: A comparison of ratings made by mother, father, child, and sibling on the CRPBI. *Child Development, 56,* 462–479.

*Scott, S., O'Connor, T. G., Futh, A., Matias, C., Price, J., & Doolan, M. (2010). Impact of a parenting program in a high-risk, multi-ethnic

community: The PALS trial. *Journal of Child Psychology and Psychiatry, 51,* 1331–1341. http://dx.doi.org/10.1111/j.1469-7610.2010.02302.x

*Scott, S., Sylva, K., Doolan, M., Price, J., Jacobs, B., Crook, C., & Landau, S. (2010). Randomised controlled trial of parent groups for child antisocial behaviour targeting multiple risk factors: The SPOKES project. *Journal of Child Psychology and Psychiatry, 51,* 48–57. http://dx.doi.org/10.1111/j.1469-7610.2009.02127.x

*Sessa, F. M., Avenevoli, S., Steinberg, L., & Morris, A. S. (2001). Correspondence among informants on parenting: Preschool children, mothers, and observers. *Journal of Family Psychology, 15,* 53–68. http://dx.doi.org/10.1037/0893-3200.15.1.53

*Sheeber, L. B., Seeley, J. R., Feil, E. G., Davis, B., Sorensen, E., Kosty, D. B., & Lewinsohn, P. M. (2012). Development and pilot evaluation of an Internet-facilitated cognitive-behavioral intervention for maternal depression. *Journal of Consulting and Clinical Psychology, 80,* 739–749. http://dx.doi.org/10.1037/a0028820

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment, 12,* 102–111. http://dx.doi.org/10.1037/1040-3590.12.1.102

Smith, M. (2011). Measures for assessing parenting in research and practice. *Child and Adolescent Mental Health, 16,* 158–166. http://dx.doi.org/10.1111/j.1475-3588.2010.00585.x

Steinberg, L., Mounts, N. S., Lamborn, S. D., & Dornbusch, S. M. (1991). Authoritative parenting and adolescent adjustment across varied ecological niches. *Journal of Research on Adolescence, 1,* 19–36.

*Tellegen, C. L., & Sanders, M. R. (2014). A randomized controlled trial evaluating a brief parenting program with children with autism spectrum disorders. *Journal of Consulting and Clinical Psychology, 82,* 1193–1200. http://dx.doi.org/10.1037/a0037246

Thompson, S. G., & Higgins, J. P. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine, 21,* 1559–1573. http://dx.doi.org/10.1002/sim.1187

*van den Akker, A. L., Deković, M., Prinzie, P., & Asscher, J. J. (2010). Toddlers' temperament profiles: Stability and relations to negative and positive parenting. *Journal of Abnormal Child Psychology, 38,* 485–495. http://dx.doi.org/10.1007/s10802-009-9379-0

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods, 45,* 576–594. http://dx.doi.org/10.3758/s13428-012-0261-6

Van der Bruggen, C. O., Stams, G. J. J. M., & Bögels, S. M. (2008). Research review: The relation between child and parent anxiety and parental control: A meta-analytic review. *Journal of Child Psychology and Psychiatry, 49,* 1257–1269. http://dx.doi.org/10.1111/j.1469-7610.2008.01898.x

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36,* 1–48. http://dx.doi.org/10.18637/jss.v036.i03

Wachs, T. D. (1987). Short-term stability of aggregated and nonaggregated measures of parental behavior. *Child Development, 58,* 796–797. http://dx.doi.org/10.2307/1130216

*Waller, R., Gardner, F., Hyde, L. W., Shaw, D. S., Dishion, T. J., & Wilson, M. N. (2012). Do harsh and positive parenting predict parent reports of deceitful-callous behavior in early childhood? *Journal of Child Psychology and Psychiatry, 53,* 946–953. http://dx.doi.org/10.1111/j.1469-7610.2012.02550.x

Waylen, A., Stallard, N., & Stewart-Brown, S. (2008). Parenting and health in mid-childhood: A longitudinal study. *European Journal of Public Health, 18,* 300–305. http://dx.doi.org/10.1093/eurpub/ckm131

*Weis, R., & Lovejoy, M. C. (2002). Information processing in everyday life: Emotion-congruent bias in mothers' reports of parent-child interactions. *Journal of Personality and Social Psychology, 83,* 216–230. http://dx.doi.org/10.1037/0022-3514.83.1.216

Wibbelink, C. J. M., & Assink, M. (2015). *Handleiding voor het uitvoeren van een drie-level meta-analyse in R (Versie 3.0)* [Manual for performing a three-level meta-analysis in R (Version 3.0)]. Unpublished manual.

*Winter, L., Morawska, A., & Sanders, M. (2012). The Knowledge of Effective Parenting Scale (KEPS): A tool for public health approaches to universal parenting programs. *The Journal of Primary Prevention, 33,* 85–97. http://dx.doi.org/10.1007/s10935-012-0268-x

*Wymbs, B. T. (2011). Mechanisms underlying the influence of disruptive child behavior on interparental communication. *Journal of Family Psychology, 25,* 873–884. http://dx.doi.org/10.1037/a0025372

*Zaslow, M. J., Weinfield, N. S., Gallagher, M., Hair, E. C., Ogawa, J. R., Egeland, B., . . . De Temple, J. M. (2006). Longitudinal prediction of child outcomes from differing measures of parenting in a low-income sample. *Developmental Psychology, 42,* 27–37. http://dx.doi.org/10.1037/0012-1649.42.1.27

Received April 15, 2016
Revision received March 7, 2017
Accepted April 12, 2017 ■