Intervention-induced temperament changes in children: Evidence from a randomized controlled trial of the Incredible Years parent program.

Huijzer-Engbrenghof, M.; van Rijn-van Gelderen, L.; van den Akker, A.; Jorgensen, T.D.; Overbeek, G.

DOI
10.1037/dev0001591

Publication date
2023

Document Version
Final published version

Published in
Developmental Psychology

AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

ISSN: 0012-1649

2023, Vol. 59, No. 10, 1839–1851
https://doi.org/10.1037/dev0001591

Intervention-Induced Temperament Changes in Children: Evidence From a Randomized Controlled Trial of the Incredible Years Parent Program

Marijke Huijzer-Engbrenghof, Loes van Rijn-van Gelderen, Alithe van den Akker, Terrence D. Jorgensen, and Geertjan Overbeek

Research Institute of Child Development and Education (RICDE), University of Amsterdam

Child temperament has long been viewed as a potential susceptibility factor in the link between parenting and child disruptive behavior (CDB). Specifically, the idea is that children with higher negative emotionality, surgency, and lower effortful control are more affected by their received parenting, but experimental evidence is scarce. Also, others have argued that child temperament might not be a susceptibility factor but a factor that can change through parents' participation in a parenting intervention. To test both hypotheses, we analyzed pretest, posttest, and 4-month follow-up data from 386 mostly Dutch parents, mainly mothers (92%; $M_{age} = 38.1$, $SD = 4.8$) with children ($M_{age} = 6.31$, $SD = 1.33$; 54.2% boys). The children had above-average disruptive behavior (i.e., ≥75th percentile Eyberg Child Behavior Inventory questionnaire; Eyberg & Pincus, 1999). The families participated in a randomized controlled trial of the Incredible Years (IY) parenting program. Hierarchical regression analyses showed that child temperament did not moderate IY intervention effects on CDB. Furthermore, parallel process analyses showed that the IY intervention led to direct, simultaneous decreases in both negative emotionality and CDB. These findings counter the widely held belief that temperament traits are static, unchangeable modulators of the links between parenting and CDB. Instead, child temperament (negative emotionality) can at least partly be influenced by parents' participation in a parenting program.

Public Significance Statement

In this randomized controlled trial of the Incredible Years (IY) parenting intervention, we found that the intervention is effective, regardless of child temperament (i.e., negative emotionality, effortful control, and surgency). We did, however, find a first indication that child temperament, like disruptive behavior, may itself change as an outcome of a parenting intervention. These findings indicate that the trait of negative emotionality can be influenced by parents' participation in the IY parenting program.

Keywords: temperament, disruptive behavior, Incredible Years, parenting, randomized controlled trial

Supplemental materials: https://doi.org/10.1037/dev0001591.supp

Early onset disruptive behavior in children, characterized by disobedience, verbal and physical aggression, oppositional behavior, and an irritable mood state, is a known risk factor for clinical diagnoses of externalizing disorders in young adulthood when left untreated (Caspi et al., 2020). Furthermore, it increases the likelihood of health problems, substance abuse, financial hardship, and delinquency in adulthood, leaving high emotional and financial costs to individuals and society in its wake (Moffitt et al., 2011). Even though in normative development, disruptive behavior and negative emotionality decrease with age in childhood (Baillargeon et al., 2012; Murphy et al., 1999; Olson et al., 2017), negative parenting, when parents apply harsh and physical disciplining strategies and lack parental warmth and sensitivity, has been identified as an important contributor to the increase of children's disruptive behavior (Pinquart, 2017). In contrast, experiencing more positive parenting, such as warm, supportive, and responsive parenting behavior, is related to less disruptive behavior in children (Boeldt et al., 2012). However, not all children respond the same to their received

Marijke Huijzer-Engbrenghof https://orcid.org/0000-0002-6989-8715
Loes van Rijn-van Gelderen https://orcid.org/0000-0002-2324-4687

This work was partly supported by the Dutch Research Council (NWO): Project 016.Veni.195.457, awarded to Terrence D. Jorgensen, and Project 016.vici.185.063, awarded to Geertjan Overbeek.

Marijke Huijzer-Engbrenghof served as the lead for writing–original draft. Loes van Rijn-van Gelderen served in a supporting role for conceptualization, methodology, and writing–original draft. Alithe van den Akker served in a supporting role for writing–review and editing. Terrence D. Jorgensen served in a supporting role for formal analysis and methodology. Geertjan Overbeek served as lead for funding acquisition, investigation, and supervision and served in a supporting role for writing–original draft.

Correspondence concerning this article should be addressed to Marijke Huijzer-Engbrenghof, Research Institute of Child Development and Education (RICDE), University of Amsterdam, Nieuwe Achtergracht 127, Postbus 15780, 1001 NG Amsterdam, Noord-Holland, The Netherlands. Email: m.huijzer-engbrenghof@uva.nl

parenting, and scholars have argued that responsiveness to parenting might depend on a child's temperament (Belsky et al., 2007). Experimental evidence for this idea is scarce (Slagt et al., 2016). Furthermore, some scholars have argued that child temperament is not only a moderator of the relationship between parenting and child disruptive behavior (CDB) but that temperamental and disruptive behavior develop in tandem (e.g., Overbeek, 2017; Rettew & McKee, 2005). We will test this idea using an experimental design. We pose that children's temperament *and* children's disruptive behavior can change through an intervention.

Temperament as Susceptibility Factor

Child temperament is defined as "constitutionally based individual differences in reactivity and self-regulation, in the domains of affect, activity, and attention" (Rothbart & Bates, 2006, p. 100). A challenge in research including child temperament is the existence of various different, overlapping temperament conceptualizations. In this study, we chose to use the three dimensions of negative emotionality, effortful control, and surgency, as it is possible to link these dimensions to disruptive behavior and psychopathology and because they have a long history in temperament research (Rettew & McKee, 2005). Negative emotionality entails the tendency to be easily distressed, which encompasses emotions of fear, worry, sadness, discomfort, anger, frustration, and irritability (Putnam et al., 2001). Next, effortful control can be defined as the capacity children have to inhibit a certain response, be able to direct attention and regulate emotions and behaviors (Rothbart & Bates, 2006). Finally, surgency includes the tendency to be, or not be, effortlessly and actively involved with their environment. Those high on surgency tend to approach novelty, enjoy intense activities, and be sociable, active, and impulsive (Putnam et al., 2001). Although scholars often assumed that temperament is stable over time (Komsi et al., 2006; Putnam et al., 2008), research has also shown that the expression of the specific behavior linked to temperament dimensions can vary across ages (de Pauw & Mervielde, 2011; Soto & Tackett, 2016).

Various theories inspired researchers to test the effect of parenting on children's behavior, including CDB, and see whether it is moderated by child temperament. Examples of those theories are the differential susceptibility theory (Belsky, 1997), the diathesis-stress model (Monroe & Simons, 1991), and the vantage sensitivity model (Pluess & Belsky, 2013). The differential susceptibility theory states that children with specific temperamental characteristics are more susceptible to their environment than others, for better and for worse (Belsky et al., 2007; Ellis et al., 2011). For instance, children with a difficult temperament, compared to children with an easier temperament, show more disruptive behaviors when parenting quality is low and fewer disruptive behaviors when parenting quality is high (e.g., Bradley & Corwyn, 2008). The diathesis-stress model argues that children have specific characteristics (e.g., difficult temperament) that make them disproportionately vulnerable to stressors in their environment (including negative parenting; Monroe & Simons, 1991). When a diathesis is activated by a stressor (e.g., harsh parenting), it leads to disruptive behavior in the child, specifically for children with a difficult temperament (Roisman et al., 2012). In contrast to the diathesis-stress model, the vantage sensitivity model states that some children, like children high in negative emotionality, are especially sensitive and positively responsive to environmental advantages. For instance, research demonstrated that children high in negative emotionality within a positive parenting environment had significantly more social and academic skills than children with the same characteristics in a negative parenting environment. In contrast, children with less negative emotionality did not benefit, or at least did not benefit to the same extent, from positive parenting (Pluess & Belsky, 2013).

A recent meta-analysis tested the moderating effect of temperament on the associations between positive and negative parenting and (among other) child externalizing problem behavior (Slagt et al., 2016). Concerning negative emotionality, the results showed that children higher on negative emotionality responded more strongly to both negative and positive parenting compared to children lower on negative emotionality, providing evidence in favor of temperament as a (differential) susceptibility factor. Regarding effortful control, the results showed that associations between positive parenting and externalizing problems were stronger for children low on effortful control, providing evidence for individual differences in environmental sensitivity. Finally, the level of children's surgency did not seem to matter in the relationship between parenting and externalizing problems. However, earlier studies rarely assessed surgency, so only a few studies on this trait were included in the meta-analysis. Also, it is important to consider that a vast majority of the 84 studies included in this meta-analysis had correlational study designs, and only a few exceptions were experimental (Slagt et al., 2016). This implies certain limitations. Correlational studies are informative in generating expectations about crucial moderators. However, experimental studies are necessary to see if there are causal relations between variables and no alternative explanations (van IJzendoorn & Bakermans-Kranenburg, 2012). The current study is such an experimental study.

In particular, we investigated the moderating role of temperament by using data from a randomized controlled trial (RCT). In this trial, the effectiveness of the parenting program Incredible Years (IY) was studied (Chhangur et al., 2012). Earlier research, including a meta-analytic review, has shown that this intervention effectively reduced CDB (mean effect size of $d = 0.27$; Menting et al., 2013) and increased positive parenting behavior (Weeland et al., 2018). Weeland et al. (2018) also investigated whether the mechanisms of change (i.e., improved parental positive affect and parenting behavior) were more important for some children than others because of children's temperamental negative emotionality but did not find evidence for this moderating role. However, this is only one dimension of the temperament construct, while the other dimensions also appear to play an important role. For instance, children low in effortful control are more likely to show disruptive behavior, even more so for children within negative parenting environments (Olson et al., 2017). Additionally, a recent study by Brown et al. (2022) found that for children with an "exuberant" temperament profile (high in surgency and low in effortful control), the relation between disruptive behavior and negative parenting intensified. Children with this profile in a negative parenting environment indeed had more disruptive behavior (Brown et al., 2022). We will therefore test the moderating role of child temperament with all three dimensions (i.e., negative emotionality, effortful control, and surgency) on the relation between a parenting intervention and CDB in an experimental design to gain more insight into the role of the three temperament aspects.

Parent Intervention-Induced Changes in Child Temperament and Disruptive Behavior

Temperament is defined as relatively stable across time and situations, and early childhood temperamental traits indeed predict later child temperament (Komsi et al., 2006; Putnam et al., 2008) and even adult personality (Blatny et al., 2007). Child temperament has mainly been investigated as an individual difference factor that may predict important outcomes. For instance, children low in effortful control and high in negative emotionality have been shown to have more often employment and relational difficulties, have more brushes with the law, and suffer more often from substance abuse and health problems in adulthood compared to children scoring on the other end of these dimensions (Caspi, 2000). However, it has also become clear that temperament and/or personality can change over time due to life experiences, such as relationship factors, stressful life events, and work experiences (Shiner et al., 2017; van Aken et al., 2006).

Next to these life experiences, research has shown that temperament can also change through parenting (Klein et al., 2018; van den Akker et al., 2010, 2014) and that these changes in child temperament predict child outcomes (for a review, see Kiff et al., 2011). For instance, more harsh parenting is related to increases in negative emotionality in children. Also, more parental control, specifically psychological control, is related to more negative reactivity across childhood. Concerning effortful control, the review by Kiff et al. (2011) showed that parental responsiveness, consistency, and warmth in early childhood predicted higher effortful control, but only in infancy and preschool years. Finally, impulsivity, as a part of surgency, appears to improve when parenting is clear, consistent, and nonpunitive (Kiff et al., 2011). When we translate this knowledge to parenting interventions, we would expect that an effective parenting intervention would lead to changes in negative emotionality, effortful control, and surgency. Empirical evidence for this idea comes from different studies. Specifically related to surgency, an RCT with children high in temperamental inhibition and meeting diagnostic criteria for an anxiety disorder (according to *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed.) showed that when parents received the parenting intervention Cool Kids Program, their children not only decreased in the frequency and severity of anxiety disorder but also showed a reduction in inhibited temperament (Kennedy et al., 2009). Another RCT, testing an intervention that consisted of common elements for parent programs targeting conduct problems (Somech & Elizur, 2012), also showed that the parenting program led to decreases in disruptive behavior as well as increases in children's effortful control. Concerning negative emotionality, correlational studies have found that parenting is related to changes in negative emotionality; a longitudinal study with preschoolers showed that maternal negativity predicted increases in child frustration (as a part of negative emotionality) over time and that child frustration and maternal negativity both predicted adjustment problems (Klein et al., 2018). Even though these studies partly show that parenting and changes in parenting can predict changes in temperamental outcomes, it must be noted that they targeted different child problem behaviors (e.g., anxiety, conduct problems). To be able to determine whether disruptive behavior can change with temperament after a parenting intervention, more RCT research appears necessary.

We are the first to test both hypotheses (temperament as a moderator of parenting intervention effectiveness or as an outcome of a parenting intervention) within one study. While doing so, we controlled for age, as disruptive behavior and child temperament can change a part of normative development (Baillargeon et al., 2012; Murphy et al., 1999; Olson et al., 2017). Firstly, we tested the moderating role of child temperament (i.e., negative emotionality, effortful control, and surgency) on the relation between a parenting intervention and CDB. Specifically, we hypothesized that a decrease in CDB differs for children with various temperament dimensions: children high in negative emotionality and surgency and/or low in effortful control may reap more benefit from the intervention and thus show a stronger decrease in disruptive behavior compared to children without these characteristics. Secondly, we investigated whether an effective parenting intervention decreases children's disruptive behavior, simultaneously with child temperament. Specifically, we expected a decrease in disruptive behavior and negative emotionality and an increase in effortful control. We tested surgency exploratory, as we do not know whether we can expect a change in this dimension in the context of this specific parenting program. Although the Cool Kids program studied by Kennedy et al. (2009) did lead to lower temperamental inhibition, it must be noted that this was with a very different intervention (i.e., focusing on how parents can manage their child's anxiety) with very different behavioral problems (i.e., clinical levels of child anxiety).

Method and Materials

Participants

Participants of this RCT were 386 parents with their children who took part in the Observational Randomized Controlled Trial on Childhood Differential Susceptibility (ORCHIDS) study. For more detailed information about the study design and sample characteristics, see Chhangur et al. (2012). At the pretest, children (54.2% boys) were between 4 and 8 years old (M_{age} = 6.31, SD = 1.32) and were primarily born in the Netherlands (96.9%). Parents (92% mothers) were between 23 and 51 years old (M_{age} = 38.09, SD = 4.83) and were primarily highly educated (50.8% with a college degree or higher). Most participating parents were married or lived together (87.0%) and worked parttime (90.5%). See Table 1 for more details.

Procedure

The families were recruited in two cohorts via two Dutch regional healthcare organizations. In the first screening phase, parents of 26,084 children ages 4–8 received the Eyberg Child Behavior Inventory questionnaire (ECBI; Eyberg & Pincus, 1999) to assess the parent-reported intensity of their child's disruptive behavior. In total, 5,874 questionnaires were returned in time (response rate: 22.5%). Parents scoring their child at or above the 75th percentile on the ECBI intensity scale (n = 1,524) were eligible for inclusion. If parents reported moderate-to-high levels of externalizing behavior of multiple children within a family, the child with the highest ECBI intensity score was selected. A total of 1,393 mother–child or father–child dyads received an invitation to participate, and of this group, 46% (n = 387) agreed to participate. Participating families did not significantly differ from nonparticipating families regarding demographic characteristics and parenting behavior. However, the ECBI scores of participating and nonparticipating children slightly differed, in that CDB problems were higher in participating families compared to nonparticipating families (Weeland et al., 2017).

Table 1
Sociodemographic Characteristics of Participants at Baseline

Baseline characteristic	Condition				
	Total $N = 386$	IY $n = 196$	Control $n = 190$	t or χ^2	p
Age child (*M*; *SD*)	6.3; 1.3	6.3; 1.4	6.3; 1.3	0.04	.97
Gender child (boy)	54.5%	57.7%	51.1%	1.71	.19
Ethnicity child (Dutch)	96.9%	96.4%	97.4%	0.28	.60
Ethnicity mother (Dutch)	81.3%	77.6%	85.3%	3.78	.05
Ethnicity father (Dutch)	81.9%	80.1%	83.7%	0.83	.36
Age-participating parent (*M*; *SD*)	38.1; 4.8	38.0; 4.9	38.2; 4.8	−0.30	.76
Gender-participating parent (female)	92.0%	92.3%	91.6%	0.08	.78
Highest education mother: College degree or higher	50.8%	48.5%	53.2%	0.85	.34
Highest education father: College degree or higher	47.1%	46.2%	48.1%	0.15	.70
Work status of participating parent				3.48	.18
Employed part-time	90.5%	93.2%	87.6%		
Employed full time	9.5%	6.8%	12.4%		
Stay-at-home parent	28.0%	26.0%	30.0%		
Marital status				3.88	.14
Married/living together	87.0%	85.6%	88.4%		
Single	8.8%	11.3%	6.3%		
Other[a]	4.2%	3.1%	5.3%		

Note. IY = Incredible Years.
[a] New family constellation ($n = 4$), new relationship ($n = 3$), registered partnership ($n = 5$), divorced ($n = 2$), long distance relationship ($n = 1$), and mother living with grandmother ($n = 1$).

Furthermore, one family's response appeared empty on inspection and was deleted from the set. We, therefore, ended up with 386 participants in the study.

The RCT consisted of three measurements: pretest before randomization; posttest immediately after intervention (i.e., 4 months after pretest); and follow-up 4 months after intervention (i.e., 8 months after pretest). During each measurement, parents filled in online questionnaires to measure parent and child behavior. Families were visited by a researcher or trained assistant to videotape a structured play situation (four 5-min play sessions: free play, child-directed play, parent-directed play, and clean up). After the first home visit, families were randomly assigned to either the control group ($n = 190$) or the intervention group ($n = 196$). Parents in the intervention group received the IY parent training. Parents in the control group were approached to ask whether they needed any parenting support or other mental health services. If so, they were directed to relevant contact points (care as usual). The ORCHIDS study was approved by a relevant Institutional Review Board in the Netherlands (METC UMC Utrecht, protocol number 11-320/K). The trial was preregistered (NTR 3594), and information about the study design, background, measures, and a priori hypotheses was published open access (Chhangur et al., 2012) at the outset. Data and study materials (e.g., syntaxes) on this specific study can be requested by contacting the first author of this article.

Intervention

The IY parenting intervention is designed to prevent and intervene in the development of child externalizing problems by reducing harsh and unresponsive parenting and increasing positive and warm parenting (Webster-Stratton et al., 2008). The intervention is a group behavioral parent training and consists of 14 weekly sessions and one booster session after 1 month after finishing the intervention. Following the IY procedure, each group was led by two group leaders, who served as a facilitator rather than experts by focusing on empowerment instead of "training" the parents. At least one of the group leaders was an experienced and certified IY group leader, and all main leaders had a background in clinical child psychology. Issues and topics addressed in the intervention included parents' child-directed play, their coaching of social behaviors and emotions, their use of praise, rewards, and incentives to reinforce appropriate behavior, and their consistency in using appropriate, noncorporal discipline practices. During the sessions, parents watched and discussed videos of parent–child interactions, engaged in role plays, and debated experiences in their families in small subgroups. After each session, parents received exercises to practice at home, read relevant literature, and practice behavior management skills with their child. Protocol adherence, which was monitored by a protocol checklist of session elements, was, on average, 86%. Previous analyses on this ORCHIDS sample (Weeland et al., 2017) showed that parents in the intervention group increased in reported positive parenting (posttest: $d = 0.49$; follow-up: $d = 0.45$) and decreased in reported negative parenting (posttest: $d = 0.29$; follow-up: $d = 0.25$).

Measures

Child Disruptive Behavior

CDB, characterized by disobedience, verbal and physical aggression, oppositional behavior, and an irritable mood state, was measured using the ECBI (Eyberg & Pincus, 1999). The ECBI is a widely used and validated instrument to assess the frequency and severity of children's disruptive behaviors between the ages of two and 18 (Funderburk et al., 2003). The ECBI was conducted at pretest, posttest, and follow-up, based on parents' self-reports. It consists of 36 items and yields an intensity subscale that measures parent-reported frequency of children's disruptive behavior on a

Likert scale from 0 (*never*) to 7 (*always*) and a problem subscale that asks parents whether they experience the specific behavior as problematic (0 = *no*, 1 = *yes*). We only used the intensity subscale for this study, as we are interested in actual behavior. The ECBI asks parents about disruptive behaviors such as temper tantrums, whining, and oppositional behavior (refusing to do specific tasks). Example items of the intensity subscale are: "My child often acts defiant when told to do something," "My child often teases or provokes other children," and "My child constantly seeks attention." Higher scores indicate higher levels of disruptive child behavior. Cronbach's αs were .89, .90, and .91 at the pretest, posttest, and follow-up. Previous research showed that the ECBI has satisfactory construct and convergent validity (Hukkelberg et al., 2018).

Child Temperament

Child temperament, defined as "constitutionally based individual differences in reactivity and self-regulation, in the domains of affect, activity, and attention" (Rothbart & Bates, 2006, p. 100), was assessed with the 36 items long Child Behavior Questionnaire-Very Short Form (CBQ-VSF; Putnam & Rothbart, 2006). The CBQ-VSF assesses child temperament on three dimensions (negative emotionality, effortful control, and surgency). Parents scored their child's temperament at the pretest, posttest, and follow-up. Previous research showed that the CBQ-VSF has satisfactory construct and convergent validity (Allan et al., 2013). Negative emotionality, measured with 12 items, reflects elevated sadness, anger, fear, discomfort, and lower soothability. Example items are: "gets quite frustrated when prevented from doing something s/he wants to do," "is very difficult to soothe when s/he has become upset," and "gets quite upset by a little cut or bruise." Parents score the items on a Likert scale from 1 (*extremely untrue for your child*) to 7 (*extremely true for your child*). Effortful control involves self-regulation, including voluntary regulation of attention and behavior, and is measured with 12 items. Some example items are: "when drawing or coloring in a book, shows strong concentration," "is good at following instructions," and "is quickly aware of some new item in the living room." Surgency, which assesses active behavior and the tendency to act impulsive, is also measured with 12 items. Example items are: "often rushes into new situations," "likes rough and rowdy games," and "seems to be at ease with almost any person." For the analyses, we used mean scores on the subscales. Cronbach's αs at pretest, posttest, and follow-up for the subscales were .73, .70, and .72 for effortful control, .67, .63, and .64 for surgency, and .67, .71, and .78 for negative emotionality.

Randomization Check

Participants (child or parent) in the intervention and control conditions did not significantly differ in age, gender, ethnicity, parental education level, work status, or marital status (see Table 1). Also, children did not differ in disruptive behavior at pretest—$t(>1{,}000^{1}) = 1.89$, $p = .059$—or temperamental characteristics—negative emotionality; $t(>1{,}000) = 1.48$, $p = .139$, effortful control; $t(>1{,}000) = -1.02$, $p = .309$, and surgency; $t(>1{,}000) = -0.97$, $p = .331$. Also, parents did not differ at the pretest on dysfunctional parenting practices—$t(>1{,}000) = 1.68$, $p = .092$—and positive parenting practices—$t(>1{,}000) = 0.37$, $p = .713$. See previous work (Weeland et al., 2017) for more details.

Analyses

Given the ubiquity of missing data due to attrition in longitudinal designs, it was no surprise that there were more missing data at posttest ($n = 20$–55 missing values, 5.18%–14.25%) and follow-up ($n = 24$–57 missing values, 6.22%–14.77%) than there were at pretest ($n = 1$–35 missing values, 0.26%–9.07%) for the disruptive behavior items and the temperament items. Before describing our target analyses, we explain how incomplete data were handled.

Missing Values and Imputation Procedure

Multiple imputation is a state-of-the-art method that preserves all available information observed in our sample, preventing unnecessary loss of power by replacing missing values with a distribution of plausible values (Little et al., 2014). Multiple imputation assumes data are missing at random (MAR), given the observed data used to predict what values would have been observed. A more restrictive assumption that data are missing *completely* (i.e., unconditionally) at random (MCAR) is ideal but unnecessary, as both MCAR and MAR mechanisms are considered "ignorable" when using state-of-the-art methods like full-information maximum likelihood and multiple imputation (Schafer & Graham, 2002). Given the size of our data set (over a dozen constructs measured by multiple-item scales on multiple occasions, along with several demographic and contextual variables), we had a high probability of capturing useful auxiliary variables (Little et al., 2014) that would be related to any systematic missing mechanism, particularly attrition (i.e., what might lead a family to drop out of the study). It, therefore, seems reasonable to assume that at least some of our missing data were MAR, minimizing the negative impact of missing data, although it is also reasonable to assume at least some data were MCAR (also ignorable) or missing not at random (MNAR) (potentially introducing some bias in our results). Given these realistic expectations, multiple imputation is the best-preferred method, yielding the least deleterious effects of any MNAR data (Gomer & Yuan, 2023).

Uncertainty about imputed values is reflected by the variance of the distribution of each missing value's imputations, which is combined with complete-data uncertainty (due to sampling error) using Rubin's (1987) rules to pool standard errors (*SE*s). We imputed the data 100 times (Graham et al., 2007) using fully conditional specification (a.k.a. "chained equations") in the R package mice (van Buuren & Groothuis-Oudshoorn, 2011). The online supplemental materials include R syntax and a detailed account of the imputation process, so we only summarize the relevant information here. Fractions of missing information (FMI; Graham et al., 2007) for scale-mean descriptive statistics are summarized in Tables 4 and 5. Specific details on how the data were imputed for this study can be found in the second online supplemental materials.

Temperament as Susceptibility Factor

The moderator hypothesis implies that the decreased CDB through IY is moderated by child temperament (i.e., negative emotionality, surgency, and/or effortful control). The moderating effect of child temperament on the effect of IY on disruptive behavior was

[1] The *df* for pooled test statistics are a function of *N*, the number of imputations, and the fraction of missing information.

examined using multiple regression analysis in R; the mitml package (Grund et al., 2021) was used to pool results across multiple imputations. The independent variable was "treatment group" (1 = *intervention group*, 2 = *control group*). The dependent variable was disruptive behavior. The moderators were child temperament, specified to effortful control, surgency, or negative emotionality at the pretest. To control whether the intervention predicts disruptive behavior above and beyond age, we added child age as a covariate. We first performed a simple regression model for each of the moderators between the pretest and posttest. This was followed by interactive models to rule out whether (the improved) disruptive behavior is dependent on child temperament. Hereafter, we repeated the analyses for the measures between the posttest and follow-up. In all analyses, disruptive behavior at the pretest was entered as a predictor to reflect the effect of IY on the predictor-adjusted change in disruptive behavior as moderated by child temperament.

Intervention-Induced Changes in Child Temperament and Disruptive Behavior

Concerning our second hypothesis, we investigated whether an effective parenting intervention decreases children's disruptive behavior simultaneously with temperamental characteristics. To examine this hypothesis, we used lavaan (Rosseel, 2012) to fit parallel-process latent difference score (LDS) models in R; the semTools package (Jorgensen et al., 2022) was used to pool results across multiple imputations. We fitted four contemporaneous developmental trajectories: one for CDB, one for negative emotionality, one for effortful control, and one for surgency. We first used the LDS models to examine change trajectories. Next, by correlating the slopes of these variables, we examined whether development trajectories are correlated among each pair of variables. A significant correlation would imply that disruptive behavior and the specific temperamental trait indeed change simultaneously. Again, to control whether the intervention predicts disruptive behavior and change in temperament above and beyond age, we added child age as a covariate.

Results

See Table 2 (means and *SD*s) and Table 3 (correlations) for the results of our descriptive analyses. Before testing our hypotheses, we also ran preliminary hierarchical regression analyses to determine whether the IY parenting program was effective in reducing CDB, and the results showed that the program was indeed effective, $b = 0.14$, $SE = 0.05$, $t(>1{,}000) = 3.01$, $p < .003$. More specifically, children whose parents received the IY intervention showed a significantly stronger decrease in disruptive behavior than children whose parents did not receive the IY intervention.

Temperament as Susceptibility Factor

Next, we tested our first hypothesis. We examined whether children's negative emotionality, effortful control, and surgency as susceptibility factors, acted as moderators within the causal link from participating in the parenting intervention IY to children's disruptive behavior while controlling for children's age. The hierarchical regression analyses showed no support for this susceptibility hypothesis (see Table 4). No significant moderator effects of child temperament factors were identified. This means that for children either lower or higher on negative emotionality, effortful control, and surgency, the IY parenting program effects are not significantly different from one another. Similar results were found between posttest and 4-month follow-up.

Intervention-Induced Changes in Child Temperament and Disruptive Behavior

To test our second hypothesis, we first looked at the mean changes in disruptive behavior and the three temperament dimensions between the pre- and posttest and between the posttest and follow-up (see Table 5). Due to the imputed dataset and the multiple tests, we also controlled for familywise false discovery rate (FDR). For disruptive child behavior, the results showed that disruptive behavior significantly decreased between the pretest and posttest in the intervention group ($M_{change} = -0.34$, $p_{FDR} < .001$) with a large effect size ($d = -1.04$). In the control group, disruptive behavior also decreased ($M_{change} = -0.16$, $p_{FDR} < .015$), but with a more modest effect size ($d = -0.56$). Between the posttest and follow-up, there was no difference in the intervention group ($M_{change} = -0.04$, $p_{FDR} = .616$, $d = -0.16$). In the control group, disruptive behavior seemed to decrease further ($M_{change} = -0.12$, $p = .025$). However, when we controlled for the FDR this effect disappeared ($p_{FDR} = .058$, $d = -0.51$).

For child temperament, the results showed that negative emotionality decreased in the intervention group between the pretest and posttest ($M_{change} = -0.21$, $p = .017$, $d = -0.44$), however when we controlled for the FDR, the effect disappeared to trend-level ($p_{FDR} = .054$). Between posttest and follow-up, negative emotionality did not change in the intervention group ($M_{change} = -0.05$, $p_{FDR} = .638$, $d = -0.13$). In the control group, negative emotionality showed no change in either two time periods ($M_{change} = -0.15$, $p_{FDR} = .113$, $M_{change} = -0.15$, $p_{FDR} = .129$). Effortful control showed an increase in the intervention group between pretest and posttest ($M_{change} = 0.20$, $p \leq .026$, $d = 0.41$) but disappeared when we applied the FDR ($p_{FDR} = .058$). Between posttest and follow-up effortful increased within the intervention group ($M_{change} = 0.38$, $p_{FDR} \leq .001$, $d = 1.11$). In the control group, effortful control also increased, but only between posttest and follow-up ($M_{change} = .29$, $p_{FDR} = .004$, $d = 0.77$). For surgency, there were no changes in either the intervention group or the control group in both time periods. These results suggest that just like disruptive behavior, the child temperament dimension of effortful control changed because of the IY parenting intervention. For negative emotionality, after correction for multiple testing, this effect was not significant, although the effect size showed a small to moderately large change nevertheless.

Second, to see whether CDB simultaneously decreases with children's temperament, we next examined the correlations between the latent change scores of negative emotionality, effortful control, surgency, and disruptive behavior from the pretest to posttest and from posttest to follow-up. Correlations between latent-change scores were significant in the intervention group for disruptive behavior and negative emotionality between the pretest and posttest ($r = .28$, $p_{FDR} < .01$). Between posttest and follow-up correlations for disruptive behavior and negative emotionality were also significant, but only without controlling for FDR ($r = .20$, $p < .05$). The same was true within the control group: disruptive behavior and negative emotionality correlated significantly between posttest and follow-up, but only without controlling for FDR ($r = .20$, $p < .05$).

Table 2
Means and SDs of Child Disruptive Behavior and Child Temperament

Group	Variable	Time	Pooled *M*	FMI (%)[a]	Pooled *SD*	FMI (%)[a]
Control	Disruptive behavior	Pretest	3.65	0.0	0.51	0.0
		Posttest	3.51	3.3	0.55	17.2
		Follow-up	3.42	7.6	0.54	23.6
	Effortful control	Pretest	4.79	0.4	0.78	6.4
		Posttest	4.70	4.5	0.80	23.4
		Follow-up	4.75	6.5	0.95	25.3
	Negative emotionality	Pretest	4.08	0.6	0.85	6.3
		Posttest	3.99	5.1	0.83	20.6
		Follow-up	3.90	7.0	0.87	20.5
	Surgency	Pretest	4.37	0.6	0.78	7.8
		Posttest	4.36	6.4	0.74	25.7
		Follow-up	4.43	12.3	0.69	35.5
Intervention	Disruptive behavior	Pretest	3.76	0.4	0.55	4.5
		Posttest	3.44	8.7	0.52	27.5
		Follow-up	3.42	11.8	0.58	30.0
	Effortful control	Pretest	4.71	0.6	0.79	9.3
		Posttest	4.66	13.3	0.82	31.5
		Follow-up	4.80	12.9	0.83	36.6
	Negative emotionality	Pretest	4.21	0.5	0.91	6.2
		Posttest	4.06	8.4	0.87	24.4
		Follow-up	4.07	11.4	0.85	23.2
	Surgency	Pretest	4.29	0.6	0.82	6.9
		Posttest	4.30	11.8	0.77	34.8
		Follow-up	4.25	13.7	0.82	25.5

Note. FMI = fractions of missing information.
[a] Fraction of missing information.

Although we also expected to see a significant correlation between effortful control and disruptive behavior in the intervention group, we found no support for this on either two time periods ($p > .05$). As expected, none of the correlations were significant in the control group on either two time periods. For surgency, there were no significant correlations in either the intervention or control group on both time periods. This suggests that, due to the intervention effect, both negative emotionality and disruptive behavior decrease simultaneously. All change correlations can be found in Table 6.

Discussion

The notion that there are individual differences in children's susceptibility to changes in their environment is well established (Belsky et al., 2007; Boyce & Ellis, 2005; Ellis et al., 2011; Pluess & Belsky, 2013; van Aar et al., 2017). Even though several studies suggested that the effect of parenting on children's disruptive behavior depends on child temperament (Karreman et al., 2009; Scott & O'Connor, 2012; Slagt et al., 2016), we did not find causal evidence for this. In this RCT of IY, we found no proof that child temperament (i.e., negative emotionality, effortful control, and surgency) moderates the efficacy of the parenting intervention IY. We did, however, find a first indication that child temperament, like disruptive behavior, may itself change as an outcome of a parenting intervention. Specifically, we found that children in the intervention group decreased in disruptive behavior, simultaneously with the temperamental characteristic of negative emotionality.

Even though several correlational studies find a moderating effect of temperament between parenting and disruptive behavior (Pluess & Belsky, 2013; Roisman et al., 2012; van Aken et al., 2007), experimental studies that find the same results are scarce. Those who did find some effects (Mouton et al., 2018; Scott & O'Connor, 2012; Slagt et al., 2017) did not include RCT data of parenting interventions, aimed at decreasing CDB, or did not look at all three temperament dimensions (e.g., negative emotionality, effortful control, and surgency), and thus are limited in providing evidence for the moderating effect of temperament on the link between parenting and CDB. While we measured whether a parenting intervention (of which we know it led to changes in parenting; see Weeland et al., 2017) decreased CDB, we did not find the expected moderating effect. Our results were, nevertheless, in line with the majority of the meta-analytic results of Slagt et al. (2016). They found that effortful control and surgency did not moderate associations between parenting and child behavior. However, the same meta-analysis found that negative emotionality did moderate the associations between parenting and child behavior (Slagt et al., 2016). At the same time, this finding was not robust to the effects of potential publication bias. Possibly more, but unpublished nonresults exist, skewing the moderation effect unseemly in favor of temperament as a susceptibility factor.

This study yields a first indication that a parenting intervention may jointly impact CDB and child temperament. This is in line with earlier research in which an intervention-induced change in personality traits (Roberts et al., 2017). More specifically, we found that when parents participated in an effective parenting program, their child's disruptive behavior decreased, and their child's negative emotionality decreased as well, yet only when we did not control for multiple testing. Furthermore, we found that the decrease in CDB and negative emotionality occurred simultaneously. Although it should be noted that children's negative emotionality decreases with age as a normative development (Murphy et al.,

Table 3
Pearson Correlations

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Disruptive behavior pretest	—	.57***	.52***	−.13	−.12	−.07	.46***	.35***	.38***	.00	.07	.09
2. Disruptive behavior posttest	.62***	—	.68***	−.04	−.08	−.08	.31***	.43***	.42***	.12	.18*	.13
3. Disruptive behavior follow-up	.58**	.68***	—	−.13	−.08	−.15	.21**	.31***	.43***	.07	.15	.07
4. Effortful control pretest	−.21**	−.15**	−.10	—	.58***	.63***	.13	.05	−.05	.06	.00	.03
5. Effortful control posttest	−.24**	−.18*	−.12	.72***	—	.71***	.12	.16	.06	.01	.01	−.01
6. Effortful control follow-up	−.20*	−.15	−.16	.63***	.70***	—	.15	.13	.07	.00	−.03	−.05
7. Negative emotionality pretest	.22**	.12	.16	.12	.11	.07	—	.62***	.61***	−.19*	−.16	−.20*
8. Negative emotionality posttest	.17*	.21**	.23**	.11	.17*	.07	.68***	—	.68***	−.09	−.02	−.11
9. Negative emotionality follow-up	.19*	.24**	.37***	.12	.18*	.16*	.60***	.69***	—	−.18*	−.09	−.14
10. Surgency pretest	.03	.01	−.07	−.10	−.11	−.03	−.34***	−.40***	−.22**	—	.74***	.69***
11. Surgency posttest	−.02	.00	−.11	−.14	−.10	−.05	−.32***	−.35***	−.25**	.72***	—	.72***
12. Surgency follow-up	.07	.03	−.02	−.14	−.16	−.07	−.32***	−.31***	−.22**	.65***	.68***	—

Note. Pooled across imputed datasets. Upper triangle = intervention group; lower triangle = control group. Fractions of missing information ranged from 0% to 17.2% (M = 6.21%, Mdn = 6.11%).
* $p < .05$. ** $p < .01$. *** $p < .001$.

1999), we controlled for the child's age in this study. As the decrease in negative emotionality was only true for the children in the intervention group, this makes it plausible that this decrease could be ascribed to the intervention. This might indicate that parenting programs are not only successful in reducing certain problem behaviors but might also positively impact less optimal temperament characteristics, preventing future maladjustment. Furthermore, our results may be consistent with a spectrum approach (Overbeek, 2017; Rettew & McKee, 2005) in which CDB and negative emotionality might not be two separate entities but reside on the same sliding scale, and both are changeable through a parenting intervention.

We also encountered several unexpected findings within the study. First, we expected an increase in effortful control in the intervention group. Effortful did increase between pretest and posttest, but only when we did not control for multiple testing. Between posttest and follow-up, effortful increased further, also within the adjusted significant levels. However, in the control group effortful control also increased between posttest and follow-up. This would suggest that effortful control increases as a part of normative development (Eisenberg et al., 2010; Posner & Rothbart, 2000). Yet, the effect of the increase in the intervention group was stronger than in the control group. This might imply that the increase in effortful control has the potential to indicate additional strength to the parenting program IY. A possible explanation for the (delayed) increase in effortful control might be due to the preceding decreases observed in negative emotionality. Scholars suggest that effortful control and emotion regulation are related to one another (Eisenberg et al., 2010). For example, in a longitudinal study, toddlers displayed more anger and frustration when they scored lower on a composite measure of effortful control tasks, whereas toddlers who performed better expressed less intense anger during frustrating situations (Kochanska et al., 2000). Having more effortful control may help children regulate emotions, and/or when children are high in negative emotionality have more difficulty developing effortful control (Kochanska & Knaack, 2003). If the intervention causes a decrease in negative emotionality, this might therefore also cause an increase in effortful control. Future research might be able to shed more light on the temperament dimension of effortful control and if a parenting intervention can impact it.

As explained in the introduction, surgency was tested exploratory. The study with the Cool Kids parenting program (Kennedy et al., 2009) was able to find a temperamental change after the intervention. As the family environment appears to play an important role in shaping individuals' surgent traits (Holmboe, 2016), we were curious whether IY could cause an effect or a change. However, this was not the case. Surgency did not influence the effectiveness of IY and nor did surgency change due to IY. Perhaps this finding can be explained by the fact that we did not specifically study reward sensitivity. Holmboe (2016), for instance, points out that individuals who score high on surgency activate a network of emotion- and reward-related brain areas more strongly when presented with positive and rewarding stimuli. This might suggest that children high in surgency are more susceptible to parenting practices based on rewards (e.g., reinforcing positive behavior by compliments). Even though IY has components that include reinforcing positive behavior, it is only a small part of the intervention. Perhaps by focusing specifically on these rewarding stimuli (within IY or a different parenting program), an effect of surgency might be detected.

Table 4
Moderation of IY Parenting Intervention Effect on Disruptive Behavior by Child Temperament at Pretest

	Pretest to posttest								Posttest to follow-up							
	Model 1				Model 2				Model 1				Model 2			
Variable	*B*	*SE*	*B**	*p*	*B*	*SE*	*B**	*p*	*B*	*SE*	*B**	*p*	*B*	*SE*	*B**	*p*
Negative emotionality																
Intercept	1.25	.20	−0.13	.030	1.23	.22	−0.13	.030	0.80	.20	0.05	.358	0.86	.22	0.05	.339
Condition	0.14	.05	0.26	.003	0.20	.22	0.26	.003	−0.06	.05	−0.10	.214	−0.20	.22	−0.10	.214
Negative emotionality	0.01	.03	0.01	.819	0.01	.04	0.02	.735	0.03	.03	0.04	.324	0.01	.04	0.02	.770
Disruptive behavior	0.59	.05	0.59	.000	0.59	.05	0.59	.000	0.70	.05	0.67	.000	0.70	.05	0.67	.000
Age	−0.01	.02	−0.02	.578	−0.01	.02	−0.02	.583	0.02	.02	0.04	.292	0.02	.02	0.04	.298
Negative Emotionality × Condition					−0.01	.05	−0.02	.799					0.03	.05	0.05	.503
Effortful control																
Intercept	1.26	.26	−0.13	.031	1.15	.29	−0.13	.034	1.06	.24	0.05	.347	1.24	.27	0.05	.371
Condition	0.14	.05	0.26	.003	0.38	.28	0.26	.003	−0.06	.05	−0.11	.202	−0.45	.27	−0.11	.198
Effortful control	−0.00	.03	−0.00	.996	0.02	.04	0.04	.552	−0.04	.03	−0.05	.234	−0.07	.04	−0.10	.062
Disruptive behavior	0.59	.04	0.59	.000	0.60	.04	0.59	.000	0.70	.04	0.67	.000	0.71	.04	0.68	.000
Age	−0.01	.02	−0.02	.576	−0.01	.02	−0.02	.609	0.02	.02	0.04	.347	0.02	.02	0.04	.386
Effortful Control × Condition					−0.05	.06	−0.08	.388					0.08	.06	0.11	.151
Surgency																
Intercept	1.10	.23	−0.13	.035	0.93	.25	−0.12	.039	0.99	.22	0.05	.341	0.91	.24	0.05	.327
Condition	0.14	.05	0.25	.003	0.50	.25	0.25	.003	−0.06	.05	−0.11	.196	0.14	.26	−0.11	.198
Surgency	0.04	.03	0.06	.185	0.08	.04	0.12	.050	−0.03	.03	−0.04	.302	−0.01	.04	−0.01	.858
Disruptive behavior	0.59	.04	0.60	.000	0.60	.04	0.60	.000	0.71	.04	0.68	.000	0.71	.04	0.68	.000
Age	−0.01	.02	−0.02	.579	−0.01	.02	−0.03	.557	0.02	.02	0.04	.292	0.02	.02	0.04	.299
Surgency × Condition					−0.08	.06	−0.13	.149					−0.05	.06	−0.07	.431

Note. For time interval pretest to posttest, disruptive behavior T1 serves as a predictor. For the time interval posttest to follow-up, disruptive behavior T2 serves as a predictor. *B** = standardized betas; IY = Incredible Years; *SE* = standard error; T1 = pretest; T2 = posttest.

Our findings of the effectivity of IY (decrease in disruptive behavior) showed stronger intervention effects compared to the outcomes of an earlier study done by our research group (Weeland et al., 2017). Due to missing data within the sample, we performed multiple imputations, which resulted in 100 completed datasets. Interestingly, the effect sizes were larger with the imputed data than with listwise

Table 5
Mean Latent Change Scores for Child Temperament and Disruptive Behavior

Group	Variable	Change	Mean change	*SE*	*p*	*p* (FDR[a])	*d*
Intervention	Disruptive behavior	Post–pre	−.34	.06	<.001*	<.001*	−1.04
		FUP–post	−.04	.06	.494	.616	−0.16
	Effortful control	Post–pre	.20	.09	.026*	.058	0.41
		FUP–post	.38	.09	.000*	.000*	1.11
	Negative emotionality	Post–pre	−.21	.00	.017*	.054	−0.44
		FUP–post	−.05	.09	.56	.638	−0.13
	Surgency	Post–pre	−.02	.07	.770	.770	−0.07
		FUP–post	−.08	.08	.299	.435	−0.25
Control	Disruptive behavior	Post–pre	−.16	.06	.004*	.015*	−0.56
		FUP–post	−.12	.06	.025*	.058	−0.51
	Effortful control	Post–pre	.15	.08	.059	.113	0.49
		FUP–post	.29	.09	.001*	.004*	0.77
	Negative emotionality	Post–pre	−.15	.08	.064	.113	−0.40
		FUP–post	−.15	.09	.081	.129	−0.15
	Surgency	Post–pre	−.05	.07	.501	.616	−0.04
		FUP–post	.04	.07	.611	.625	0.12
Group differences	Disruptive behavior	Post–pre	−.18	.05	<.001*	.003*	
		FUP–post	.09	.05	.087	.258	
	Effortful control	Post–pre	.05	.07	.523	.598	
		FUP–post	.09	.08	.257	.411	
	Negative emotionality	Post–pre	−.06	.08	.452	.598	
		FUP–post	.10	.08	.213	.411	
	Surgency	Post–pre	.03	.06	.674	.674	
		FUP–post	−.12	.07	.097	.258	

Note. *SE* = standard error; FUP = follow-up; FDR = false discovery rate.
[a] FDR indicates *p* values adjusted to maintain familywise false discovery rate.
* $p < .05$.

Table 6
Latent Change Score Correlations Between Child Temperament Dimensions and Disruptive Behavior

Variable	1	2	3	4	5	6	7	8
1. d1 disruptive behavior	—	−.32**,*	−.05	−.00	.28**	−.03	−.03	−.10
2. d2 disruptive behavior	−.42***	—	.13	−.12	−.01	.20*	.06	−.08
3. d1 effortful control	.00	.02	—	−.49***	.14	−.01	.09	−.07
4. d2 effortful control	−.01	−.10	−.30*,*	—	−.07	.05	−.02	−.02
5. d1 negative emotionality	.20*	−.03	.12	−.07	—	−.45***	.04	−.06
6. d2 negative emotionality	.03	.18	.02	.14	−.35**,*	—	.03	.06
7. d1 surgency	.06	−.03	.08	−.05	.06	−.16	—	−.41**,*
8. d2 surgency	−.07	.09	−.10	.03	.05	−.02	−.42**,*	—

Note. d1 = from pretest to posttest; d2 = from posttest to follow-up. Upper triangle = intervention group, lower triangle = control group. Asterisks indicate significance of p values adjusted for multiple testing (false discovery rate), and asterisks in bold indicate significance of unadjusted p values.
* $p < .05$. ** $p < .01$. *** $p < .001$.

deletion (i.e., complete-case analysis). In addition to preserving power by using all available information for analysis, listwise deletion would have biased our results due to analyzing a nonrandom subsample of data (Asendorpf et al., 2014; Graham, 2009).

Strengths and Limitations

This study had multiple strengths. Most notably, it involved an experimental design that allowed us to examine the effect of participating in a parenting intervention on the development of CDB. Also, this study was the first to longitudinally test whether child temperament can change through an effective parenting intervention and whether the prospective development of child temperament and disruptive behavior was correlated. Finally, all analyses were conducted using sophisticated multiple imputation and latent change analysis strategies, increasing the precision and generalizability of our conclusions.

Several limitations warrant mentioning as well. The most important one—applying to most RCTs—is that the generalization of results might be limited to families willing to be randomized, with an equal chance of being assigned to an experimental and control group. This might have caused an underestimation of effect sizes because especially high-risk families might have decided not to participate. Although we included at-risk families by screening them on disruptive behavior, the sample consisted mainly of well-educated and relatively high-socioeconomic status (SES) families. Compared to high-SES parents, parents of low-SES are more likely to have additional problems and are, because of that, more at risk for negative parenting practices (Conger et al., 2010). Another limitation concerns the modest sample size. Even though this study features one of the largest RCTs on IY worldwide, it is certainly possible that analyses on a larger sample might have revealed significant temperament-by-intervention effects. Also, the modest sample size hindered us in conducting mediation-moderation analyses with parenting changes as a mediator with enough power to detect small to medium effects. Yet another limitation concerns the fact that we could not examine the temperament dimensions (negative emotionality, effortful control, and surgency) in more detail because of the instrument (CBQ-VSF) we used. In future research, it would be interesting to see whether subscales from these dimensions are all similarly closely linked to changes in disruptive behavior. Additionally, our data were all parent-reported. We acknowledge that this limits the strength of our findings. Concerning the parent-reported data, it is also possible that the significant association between slopes is because parents view their children's behavior differently over a year since they were aware that they received a parenting program. Ideally, multi-informant measures, such as observations or teacher reports, are incorporated within a study to test whether the results hold when data from nonparents are used.

Clinical Implications

One important implication may be that in professional communications with parents about their child's development, temperament should be approached as a dynamic rather than a static factor. This can help parents feel in control of challenging situations (e.g., temper tantrums) and gain perspective on how they can positively stimulate the development of their child's behavior. Also, this may help to counter self-fulfilling prophecies based on parental expectations of disruptive behavior, for instance, in children with higher levels of negative emotionality. The fact that temperament is a changeable entity also emphasizes the importance of early screening efforts based on child temperament. Which manifestations and intensities of child negative emotionality can we identify as early warning signals for full-blown disruptive behavior later in childhood? Future research could help to delineate a developmental profile of temper tantrums and other instances of pronounced negative emotionality, for instance, that could be considered "risky" for the development of disruptive behavior (van den Akker et al., 2022; Wakschlag et al., 2012). Longitudinal studies might shed light on whether temperamental change through interventions indeed improves children's outcomes, preventing health problems, substance abuse, financial hardship, and delinquency in adulthood. Finally, based on our findings, it seems unnecessary to differentiate parenting program approaches based on (parent-reported) child temperament traits. At least the IY program achieves equally satisfactory results for children with "difficult" versus "easier" temperaments. However, the small effect sizes of parenting programs (Menting et al., 2013) do indicate that we need to continue our search for the factors—at the individual, family, and contextual level—that explain the large heterogeneity in parenting program effects (see, e.g., van Aar et al., 2017).

Conclusion

In conclusion, our study supports the notion that at least negative emotionality and disruptive behavior can change simultaneously

after implementing an effective parenting program. These findings counter the widely held belief that temperament traits are static, unchangeable modulators of the links between parenting and CDB. Instead, child temperament (negative emotionality) can partly be influenced by parents' participation in a parenting program.

References

Allan, N. P., Lonigan, C. J., & Wilson, S. B. (2013). Psychometric evaluation of the Children's Behavior Questionnaire–Very Short Form in preschool children using parent and teacher report. *Early Childhood Research Quarterly*, *28*(2), 302–313. https://doi.org/10.1016/j.ecresq.2012.07.009

Asendorpf, J. B., van de Schoot, R., Denissen, J. J. A., & Hutteman, R. (2014). Reducing bias due to systematic attrition in longitudinal studies: The benefits of multiple imputation. *International Journal of Behavioral Development*, *38*(5), 453–460. https://doi.org/10.1177/0165025414542713

Baillargeon, R. H., Morisset, A., Keenan, K., Normand, C. L., Séguin, J. R., Japel, C., & Cao, G. (2012). Development of disruptive behaviors in young children: A prospective population-based cohort study. *Infant Mental Health Journal*, *33*(6), 633–650. https://doi.org/10.1002/imhj.21353

Belsky, J. (1997). Variation in susceptibility to environmental influence: An evolutionary argument. *Psychological Inquiry*, *8*(3), 182–186. https://doi.org/10.1207/s15327965pli0803_3

Belsky, J., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). For better *and* for worse. *Current Directions in Psychological Science*, *16*(6), 300–304. https://doi.org/10.1111/j.1467-8721.2007.00525.x

Blatny, M., Jelinek, M., & Osecka, T. (2007). Assertive toddler, self-efficacious adult: Child temperament predicts personality over forty years. *Personality and Individual Differences*, *43*(8), 2127–2136. https://doi.org/10.1016/j.paid.2007.06.024

Boeldt, D. L., Rhee, S. H., Dilalla, L. F., Mullineaux, P. Y., Jay Schulz-Heik, R., Corley, R. P., Young, S. E., & Hewitt, J. K. (2012). The association between positive parenting and externalizing behavior. *Infant and Child Development*, *21*(1), 85–106. https://doi.org/10.1002/icd.764

Boyce, W. T., & Ellis, B. J. (2005). Biological sensitivity to context: I. An evolutionary–developmental theory of the origins and functions of stress reactivity. *Development and Psychopathology*, *17*(02), 271–301. https://doi.org/10.1017/S0954579405050145

Bradley, R. H., & Corwyn, R. F. (2008). Infant temperament, parenting, and externalizing behavior in first grade: A test of the differential susceptibility hypothesis. *Journal of Child Psychology and Psychiatry*, *49*(2), 124–131. https://doi.org/10.1111/J.1469-7610.2007.01829.X

Brown, K. M., Pérez-Edgar, K., & Lunkenheimer, E. (2022). Understanding how child temperament, negative parenting, and dyadic parent–child behavioral variability interact to influence externalizing problems. *Social Development*, *31*(4), 1020–1041. https://doi.org/10.1111/sode.12601

Caspi, A. (2000). The child is father of the man: Personality continuities from childhood to adulthood. *Journal of Personality and Social Psychology*, *78*(1), 158–172. https://doi.org/10.1037/0022-3514.78.1.158

Caspi, A., Houts, R. M., Ambler, A., Danese, A., Elliott, M. L., Hariri, A., Harrington, H. L., Hogan, S., Poulton, R., Ramrakha, S., Rasmussen, L. J. H., Reuben, A., Richmond-Rakerd, L., Sugden, K., Wertz, J., Williams, B. S., & Moffitt, T. E. (2020). Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the Dunedin birth cohort study. *JAMA Network Open*, *3*(4), Article e203221. https://doi.org/10.1001/jamanetworkopen.2020.3221

Chhangur, R. R., Weeland, J., Overbeek, G., Matthys, W., & Orobio De Castro, B. (2012). ORCHIDS: An observational randomized controlled trial on childhood differential susceptibility. *BMC Public Health*, *12*(1), Article 917. https://doi.org/10.1186/1471-2458-12-917

Conger, R. D., Conger, K. J., & Martin, M. J. (2010). Socioeconomic status, family processes, and individual development. *Journal of Marriage and Family*, *72*(3), 685–704. https://doi.org/10.1111/j.1741-3737.2010.00725.x

de Pauw, S. S. W., & Mervielde, I. (2011). The role of temperament and personality in problem behaviors of children with ADHD. *Journal of Abnormal Child Psychology*, *39*(2), 277–291. https://doi.org/10.1007/s10802-010-9459-1

Eisenberg, N., Spinrad, T. L., & Eggum, N. D. (2010). Emotion-related self-regulation and its relation to children's maladjustment. *Annual Review of Clinical Psychology*, *6*(1), 495–525. https://doi.org/10.1146/annurev.clinpsy.121208.131208

Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M. J., & van Ijzendoorn, M. H. (2011). Differential susceptibility to the environment: An evolutionary–neurodevelopmental theory. *Development and Psychopathology*, *23*(1), 7–28. https://doi.org/10.1017/S0954579410000611

Eyberg, S. M., & Pincus, D. (1999). *Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory–Revised: Professional manual*. Psychological Assessment Resources.

Funderburk, B. W., Eyberg, S. M., Rich, B. A., & Behar, L. (2003). Further psychometric evaluation of the Eyberg and Behar rating scales for parents and teachers of preschoolers. *Early Education and Development*, *14*(1), 67–82. https://doi.org/10.1207/s15566935eed1401_5

Gomer, B., & Yuan, K. H. (2023). A realistic evaluation of methods for handling missing data when there is a mixture of MCAR, MAR, and MNAR mechanisms in the same dataset. *Multivariate Behavioral Research*. Advance online publication. https://doi.org/10.1080/00273171.2022.2158776

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*(1), 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, *8*(3), 206–213. https://doi.org/10.1007/s11121-007-0070-9

Grund, S., Lüdtke, O., & Robitzsch, A. (2021). Multiple imputation of missing data in multilevel models with the R package mdmb: A flexible sequential modeling approach. *Behavior Research Methods*, *53*(6), 2631–2649. https://doi.org/10.3758/s13428-020-01530-0

Holmboe, K. (2016). Encyclopedia of personality and individual differences. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp. 1–6). Springer. https://doi.org/10.1007/978-3-319-28099-8_2123-1

Hukkelberg, S. S., Reedtz, C., & Kjøbli, J. (2018). Construct validity of the Eyberg child behavior inventory (ECBI): An exploratory structural equation modeling approach. *European Journal of Psychological Assessment*, *34*(1), 14–21. https://doi.org/10.1027/1015-5759/a000316

Jorgensen, T. D., Mansolf, M., & Enders, C. K. (2022, March). *Pooled score tests for multiply imputed data using semTools*. https://files.osf.io/v1/resources/m2gd8/providers/osfstorage/622f127453a4e80bef51685c?action=download&direct&version=1

Karreman, A., van Tuijl, C., van Aken, M. A. G., & Deković, M. (2009). Predicting young children's externalizing problems: Interactions among effortful control, parenting, and child gender. *Merrill-Palmer Quarterly*, *55*(2), 111–134. https://doi.org/10.1353/mpq.0.0020

Kennedy, S. J., Rapee, R. M., & Edwards, S. L. (2009). A selective intervention program for inhibited preschool-aged children of parents with an anxiety-disorder: Effects on current anxiety disorders and temperament. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(6), 602–609. https://doi.org/10.1097/CHI.0b013e31819f6fa9

Kiff, C. J., Lengua, L. J., & Zalewski, M. (2011). Nature and nurturing: Parenting in the context of child temperament. *Clinical Child and Family Psychology Review*, *14*(3), 251–301. https://doi.org/10.1007/s10567-011-0093-4

Klein, M. R., Lengua, L. J., Thompson, S. F., Moran, L., Ruberry, E. J., Kiff, C., & Zalewski, M. (2018). Bidirectional relations between temperament and parenting predicting preschool-age children's adjustment. *Journal of Clinical Child & Adolescent Psychology*, *47*(Suppl. 1), S113–S126. https://doi.org/10.1080/15374416.2016.1169537

Kochanska, G., & Knaack, A. (2003). Effortful control as a personality characteristic of young children: Antecedents, correlates, and consequences. *Journal of Personality*, *71*(6), 1087–1112. https://doi.org/10.1111/1467-6494.7106008

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology*, *36*(2), 220–232. https://doi.org/10.1037/0012-1649.36.2.220

Komsi, N., Räikkönen, K., Pesonen, A. K., Heinonen, K., Keskivaara, P., Järvenpää, A. L., & Strandberg, T. E. (2006). Continuity of temperament from infancy to middle childhood. *Infant Behavior and Development*, *29*(4), 494–508. https://doi.org/10.1016/j.infbeh.2006.05.002

Little, T. D., Jorgensen, T. D., Lang, K. M., & Moore, E. W. G. (2014). On the joys of missing data. *Journal of Pediatric Psychology*, *39*(2), 151–162. https://doi.org/10.1093/jpepsy/jst048

Menting, A. T. A., Orobio de Castro, B., & & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: A meta-analytic review. *Clinical Psychology Review*, *33*(8), 901–913. https://doi.org/10.1016/j.cpr.2013.07.006

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Roberts, B. W., Ross, S., Sears, M. R., Murray Thomson, W., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*(7), 2693–2698. https://doi.org/10.1073/pnas.1010076108

Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: Implications for the depressive disorders. *Psychological Bulletin*, *110*(3), 406–425. https://doi.org/10.1037/0033-2909.110.3.406

Mouton, B., Loop, L., Stievenart, M., & Roskam, I. (2018). Child differential sensitivity to parental self-efficacy improvement. *International Journal of Behavioral Development*, *42*(2), 203–213. https://doi.org/10.1177/0165025416687416

Murphy, B. C., Eisenberg, N., Fabes, R. A., Shepard, S., & Guthrie, I. K. (1999). Consistency and change in children's emotionality and regulation: A longitudinal study. *Merrill-Palmer Quarterly*, *45*(3), 413–444. https://www.jstor.org/stable/23092580

Olson, S. L., Choe, D. E., & Sameroff, A. J. (2017). Trajectories of child externalizing problems between ages 3 and 10 years: Contributions of children's early effortful control, theory of mind, and parenting experiences. *Development and Psychopathology*, *29*(4), 1333–1351. https://doi.org/10.1017/S095457941700030X

Overbeek, G. (2017). Parenting intervention effects on children's externalizing behavior: The moderating role of genotype and temperament. *Current Opinion in Psychology*, *15*, 143–148. https://doi.org/10.1016/j.copsyc.2017.02.025

Pinquart, M. (2017). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: An updated meta-analysis. *Developmental Psychology*, *53*(5), 873–932. https://doi.org/10.1037/dev0000295

Pluess, M., & Belsky, J. (2013). Vantage sensitivity: Individual differences in response to positive experiences. *Psychological Bulletin*, *139*(4), 901–916. https://doi.org/10.1037/a0030196

Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, *12*(3), 427–441. https://doi.org/10.1017/S0954579400003096

Putnam, S. P., Ellis, L. K., & Rothbart, M. K. (2001). The structure of temperament from infancy through adolescence. In A. Eliasz & A. Angleitner (Eds.), *Advances in research on temperament* (Vol. 165, pp. 165–182). Pabst Scientist Publisher.

Putnam, S. P., & Rothbart, M. K. (2006). Development of short and very short forms of the children's behavior questionnaire. *Journal of Personality Assessment*, *87*(1), 102–112. https://doi.org/10.1207/s15327752jpa8701_09

Putnam, S. P., Rothbart, M. K., & Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. *Infant and Child Development*, *17*(4), 387–405. https://doi.org/10.1002/icd.582

Rettew, D. C., & McKee, L. (2005). Temperament and its role in developmental psychopathology. *Harvard Review of Psychiatry*, *13*(1), 14–27. https://doi.org/10.1080/10673220590923146

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin*, *143*(2), 117–141. https://doi.org/10.1037/bul0000088

Roisman, G. I., Newman, D. A., Fraley, R. C., Haltigan, J. D., Groh, A. M., & Haydon, K. C. (2012). Distinguishing differential susceptibility from diathesis-stress: Recommendations for evaluating interaction effects. *Development and Psychopathology*, *24*(2), 389–409. https://doi.org/10.1017/S0954579412000065

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rothbart, M. K., & Bates, J. E. (2006). Temperament. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Vol 3. Social, emotional, and personality development* (6th ed., pp. 99–166). Wiley.

Rubin, D. B. (Ed.). (1987). *Multiple imputation for nonresponse in surveys*. John Wiley. https://doi.org/10.1002/9780470316696

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Scott, S., & O'Connor, T. G. (2012). An experimental test of differential susceptibility to parenting among emotionally-dysregulated children in a randomized controlled trial for oppositional behavior. *Journal of Child Psychology and Psychiatry*, *53*(11), 1184–1193. https://doi.org/10.1111/j.1469-7610.2012.02586.x

Shiner, R. L., Allen, T. A., & Masten, A. S. (2017). Adversity in adolescence predicts personality trait change from childhood to adulthood. *Journal of Research in Personality*, *67*, 171–182. https://doi.org/10.1016/j.jrp.2016.10.002

Slagt, M., Dubas, J. S., Deković, M., & van Aken, M. A. G. (2016). Differences in sensitivity to parenting depending on child temperament: A meta-analysis. *Psychological Bulletin*, *142*(10), 1068–1110. https://doi.org/10.1037/bul0000061

Slagt, M., Dubas, J. S., van Aken, M. A. G., Ellis, B. J., & Deković, M. (2017). Children's differential susceptibility to parenting: An experimental test of "for better and for worse." *Journal of Experimental Child Psychology*, *154*, 78–97. https://doi.org/10.1016/j.jecp.2016.10.004

Somech, L. Y., & Elizur, Y. (2012). Cooperation in pre-kindergarten children with conduct problems: A randomized. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(4), 412–422. https://doi.org/10.1016/j.jaac.2012.01.019

Soto, C. J., & Tackett, J. L. (2016). Personality traits in childhood and adolescence: Structure, development, and outcomes. *Cancer Research*, *76*(4), 358–362. https://doi.org/10.1177/0963721415589345/FORMAT/EPUB

van Aar, J., Leijten, P., Orobio de Castro, B., & Overbeek, G. (2017). Sustained, fade-out or sleeper effects? A systematic review and meta-analysis of parenting interventions for disruptive child behavior. *Clinical Psychology Review*, *51*, 153–163. https://doi.org/10.1016/j.cpr.2016.11.006

van Aken, C., Junger, M., Verhoeven, M., van Aken, M. A. G., & Deković, M. (2007). The interactive effects of temperament and maternal parenting on toddlers' externalizing behaviours. *Infant and Child Development*, *16*(5), 553–572. https://doi.org/10.1002/icd.529

van Aken, M. A. G., Denissen, J. J. A., Branje, S. J. T., Dubas, J. S., & Goossens, L. (2006). Midlife concerns and short-term personality change in middle adulthood. *European Journal of Personality*, *20*(6), 497–513. https://doi.org/10.1002/per.603

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–67. https://doi.org/10.18637/JSS.V045.I03

van den Akker, A. L., Deković, M., Asscher, J., & Prinzie, P. (2014). Mean-level personality development across childhood and adolescence: A temporary defiance of the maturity principle and bidirectional associations with parenting. *Journal of Personality and Social Psychology*, *107*(4), 736–750. https://doi.org/10.1037/a0037248

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*(4), 485–495. https://doi.org/10.1007/s10802-009-9379-0

van den Akker, A. L., Hoffenaar, P., & Overbeek, G. (2022). Temper tantrums in toddlers and preschoolers: Longitudinal associations with adjustment problems. *Journal of Developmental & Behavioral Pediatrics*, *43*(7), 409–417. https://doi.org/10.1097/dbp.0000000000001071

van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2012). Differential susceptibility experiments: Going beyond correlational evidence: Comment on beyond mental health, differential susceptibility articles. *Developmental Psychology*, *48*(3), 769–774. https://doi.org/10.1037/a0027536

Wakschlag, L. S., Choi, S. W., Carter, A. S., Hullsiek, H., Burns, J., McCarthy, K., Leibenluft, E., & Briggs-Gowan, M. J. (2012). Defining the developmental parameters of temper loss in early childhood: Implications for developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *53*(11), 1099–1108. https://doi.org/10.1111/j.1469-7610.2012.02595.x

Webster-Stratton, C., Jamila Reid, M., & Stoolmiller, M. (2008). Preventing conduct problems and improving school readiness: Evaluation of the Incredible Years teacher and child training programs in high-risk schools. *Journal of Child Psychology and Psychiatry*, *49*(5), 471–488. https://doi.org/10.1111/j.1469-7610.2007.01861.x

Weeland, J., Chhangur, R. R., Jaffee, S. R., van der Giessen, D., Matthys, W., Orobio De Castro, B., & Overbeek, G. (2018). Does the Incredible Years reduce child externalizing problems through improved parenting? The role of child negative affectivity and serotonin transporter linked polymorphic region (5-HTTLPR) genotype. *Development and Psychopathology*, *30*(1), 93–112. https://doi.org/10.1017/S0954579417000499

Weeland, J., Chhangur, R. R., van der Giessen, D., Matthys, W., de Castro, B. O., & Overbeek, G. (2017). Intervention effectiveness of the Incredible Years: New insights into sociodemographic and intervention-based moderators. *Behavior Therapy*, *48*(1), 1–18. https://doi.org/10.1016/j.beth.2016.08.002

Received November 11, 2022
Revision received June 1, 2023
Accepted June 5, 2023 ■