Letter

Treating Parental Burnout: Impact of Two Treatment Modalities on Burnout Symptoms, Emotions, Hair Cortisol, and Parental Neglect and Violence

Forthcoming in *Psychotherapy and Psychosomatics*

Maria Elena Brianda[1]*, Isabelle Roskam[1], James J. Gross[2], Aline Franssen[3], Florence Kapala[3], France Gérard[3], Moïra Mikolajczak[1]

[1] Psychological Sciences Research Institute, Department of Psychology, UCLouvain, Louvain-la-Neuve, Belgium

[2] Department of Psychology, Stanford University, Stanford, CA, USA

[3] Mutualité Chrétienne-Christelijke Mutualiteit, Health Promotion Service, Brussels, Belgium

Short Title: Treating Parental Burnout

*Corresponding Author

Maria Elena Brianda

Psychological Sciences Research Institute

Department of Psychology

UCLouvain

Place Cardinal Mercier, 10

Louvain-la-Neuve, 1348, Belgium

Tel: +32(0)10474687

E-mail: maria.brianda@uclouvain.be

Keywords: burn-out; parent; treatment; efficiency

Being a parent can be rewarding, but it can also be stressful [1]. Pressures on parents have intensified since the 1980s, to the point that parenting stress has become the rule rather than the exception [2]. When parents lack the resources needed to handle parenting stress, they may develop *parental burnout* (PB) [3, 4], a state of intense exhaustion related to one's parental role, in which one becomes emotionally detached from one's children and doubtful of one's capacity to be a good parent [4]. PB is a prevalent condition (5-6% of parents are affected in Western countries), with severe consequences for parents (e.g., poor health and suicidal ideation [4, 5]) and children (e.g., parental neglect and violence [4, 5]).

These consequences highlight the urgent need for effective treatments for PB. Unfortunately, no such treatments are currently available. One possible approach might be to guide parents in decreasing their parenting stressors and/or increasing their parenting resources. This approach would directly address the defining features of PB [3]. A second approach might be to offer a setting in which people can be heard and understood without judgment. This is an approach that has shown tremendous therapeutic power in other contexts [6], and it might be particularly helpful in the context of PB, where guilt and shame often prevent parents from sharing their difficulties with friends and family members [7].

The current study aimed to compare these two approaches by examining the efficiency of two group-based interventions. The group setting seems indeed particularly appropriate for parents suffering from PB who often tend to isolate themselves, and feel extremely lonely and unsupported [7]. Both interventions are briefly described below and fully detailed in the Supplementary Material. Interventions Manuals are also available. The first intervention (Directive intervention) aimed to restore the balance between parental stressors and resources. This intervention was led by psychologists who were expert in PB and trained to guide parents to actively work on the main parental stressors (e.g., societal pressure, parental

perfectionism, lack of coparental support…) and resources (e.g., emotional competences, stress managing abilities, effective childrearing practices, help-seeking…) [3], by providing them with psychoeducation and targeted exercises. The second intervention (Non-Directive intervention) aimed to offer parents a space where they could share their difficulties. This intervention was led by psychologists trained in active listening [6], who offered a setting of active, empathetic, caring, and authentic listening in which participants could share, perceive themselves as fundamentally worthy of consideration regardless of their difficulties, and believe in their capacity to find their own path and resources [6, 8].

To compare these interventions, 142 parents were randomly assigned to either directive (N = 77) or non-directive interventions (N = 65). Interventions consisted of 8 weekly two-hour group-format sessions. Measures were taken three times: before the intervention, immediately after the intervention, and 3 months after the intervention. In a subgroup of 40 parents, the intervention took place two months or more after they enrolled in the study, enabling us to obtain measures two times prior to the intervention (with an 8-week intervening period that matched the intervention duration), as well as the assessments at the end of the intervention and then three months later. This assessment strategy allowed us to examine how parental PB spontaneously evolves without intervention.

Measures were of three types: self-reported, informant-reported, and biological. Self-reported measures assessed PB (*Parental Burnout Assessment – PBA* [9]), parental neglect (*Parental Neglect Scale [5]*), parental violence (*Parental Violence Scale* [5]), positive and negative emotions (*Positive and Negative Affect Schedule* [10]), and the balance between risks and resources (*BR²* [3]). Informant-reported measures (completed by the partner or by an adult close to them who saw them at least three times a week) assessed the participant's symptoms of PB, positive and negative emotions, and irritability (see Supplementary

Material). The biological measure consisted of hair cortisol, a measure of cortisol accumulated over the last three months (taken before and three months after the intervention). The flowchart, procedure, and measures are detailed in Supplementary Material.

Results are shown in Table 1 (see Supplementary Material for detailed statistical procedures, preliminary checks, and complementary analyses). Whereas PB remains stable in the absence of intervention – neither participants nor their partners reported a spontaneous reduction of PB symptoms over the two month-period preceding the interventions – it decreased significantly after both interventions. Findings show a significant main effect of Time on all dependent variables, but no Time x Condition effect, indicating that PB evolved in the same way in the two conditions. This was true both immediately after the intervention and at the 3-month follow-up. Both interventions were accompanied by a decrease in self-reported PB symptoms (-37%), negative emotions (-29%), parental neglect (-35%), and parental violence (-32%). Self-reported positive emotions also increased after both interventions (+18%). These effects were consistent with the informant-reported evaluation, including reduced PB (-28%) and negative emotions (-24%), and increased positive emotions (+28%). Cortisol decreased by 52% after the interventions.

Beyond suggesting that PB can potentially be treated, these results suggest that it may not be necessary to be expert in PB to deliver effective support for parents suffering from PB, at least in a group setting. Psychologists trained in active listening and with expertise in group interventions have the necessary skills to help parents suffering from PB, which is very good news given the high prevalence of this condition and its damaging effect on children. A framework of active listening, empathy, and comprehension, along with an invitation to consider selected topics relevant for parents with PB, seems sufficient to achieve positive and lasting effects on parents' well-being. Given the comparably positive effect of the directive intervention, a fruitful avenue for future research would consist in examining the effect of a

combined intervention. Such research would benefit from overcoming the main limitations of the current study, namely the self-selection of the participants, the size and nature of the control group, the extent of missing data at follow-up, and the small proportion of fathers.

The database is publicly available on Open Science Framework (https://osf.io/kn7bw/?view_only=None).

Disclosure Statement

M.E.B., J.J.G., A.F., F.K. and F.G. have no conflicts of interest to declare. M.M. and I.R. have founded the Training Institute for Parental Burnout (TIPB) which delivers training on PB to professionals. TIPB was founded after the completion of the study (including the analysis of the results). Thus, the institute did not participate in the funding of this study nor did it influence the process or the results in any manner.

Funding Sources

This study was co-funded by a public center for the promotion of health, Infor Santé (Mutualité Chrétienne) and an FSR-2016 Grant of UCLouvain to M.M. and I.R

References

1 Nomaguchi KM, Milkie MA: Costs and Rewards of Children: The Effects of Becoming a Parent on Adults' Lives. J Marriage Fam 2003;65:356-374.

2 Faircloth C: Intensive parenting and the expansion of parenting. In: Lee E, editor. Parenting culture studies. London: Palgrave Macmillan, 2014, pp 25-50.

3 Mikolajczak M, Roskam I: A theoretical and clinical framework for parental burnout: The Balance between Risks and Resources (BR²). Front Psychol 2018;9:886.

4 Mikolajczak M, Gross JJ, Roskam I: Parental Burnout: What Is It, and Why Does It Matter? Clin Psychol Sci 2019;7:1319-1329.

5 Mikolajczak M, Brianda ME, Avalosse H, Roskam I: Consequences of parental burnout: Its specific effect on child neglect and violence. Child Abuse Negl 2018;80:134-145.

6 Lambert MJ, Barley DE: Research summary on the therapeutic relationship and psychotherapy outcome. Psychother-Theor Res 2001;38:357.

7 Hubert S, Aujoulat I: Parental Burnout: When Exhausted Mothers Open Up. Front Psychol 2018;9:1021-1021.

8 Rogers CR, Wood JK: Client-centered theory. In: Burton A, editor. Operational theories of personality. Oxford: Brunner/Mazel, 1974

9 Roskam I, Brianda M-E, Mikolajczak M: A step forward in the conceptualization and measurement of parental burnout: The Parental Burnout Assessment (PBA). Front Psychol 2018;9

10 Watson D, Clark LA, Tellegen A: Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54:1063.

Table 1. Means, standard deviations, and mean differences of repeated measure ANOVA tests for self-reported, informant-reported and biological measures

		Directive				Non-Directive					
		T1	T2	T3	Mean difference	T1	T2	T3	Mean difference	Repeated measure ANOVA	
		M (SD)	*M (SD)*	*M (SD)*	%	*M (SD)*	*M (SD)*	*M (SD)*	%	F	Cohen's d
Self-reported outcomes (N = 62 to 96)	PB	86.49 (29.77)	61.40 (33.66)	54.12 (33.43)	-37.43%	73.12 (30.43)	64.25 (35.25)	50.70 (31.32)	-30.66%	64.24***	0.86
	Parental Neglect	20.12 (14.40)	14.90 (11.08)	13.07 (9.48)	-35.04%	16.35 (11.34)	15.77 (10.03)	14.35 (12.07)	-12.23%	16.56***	0.43
	Parental Violence	19.14 (13.62)	13.71 (9.54)	12.98 (8.85)	-32.18%	17.51 (10.78)	15.23 (10.45)	13.28 (7.64)	-24.16%	28.45***	0.58
	Positive Emotions	10.28 (5.09)	12.08 (5.21)	13.16 (5.15)	+28.02%	10.67 (4.66)	10.19 (4.47)	12.29 (4.32)	+15.18%	21.76***	-0.50
	Negative Emotions	18.40 (5.29)	14.20 (6.04)	13.58 (6.12)	-26.20%	19.33 (5.89)	15.19 (6.43)	13.74 (5.29)	-28.92%	70.31***	0.91
	Balance R-R	9.74 (51.54)	11.10 (60.62)	25.16 (55.54)	+158.32%	1.07 (41.27)	-9.45 (49.23)	16.49 (38.89)	+1441.12	8.17**	0.31
	Job Burnout	37.15 (12.41)	37.45 (19.12)	32.44 (15.60)	-12.68	32.43 (16.56)	33.71 (18.84)	33.46 (17.72)	+3.18	1.59	0.13
Informant-reported outcomes (N = 36 to 37)	PB_i	70.70 (30.11)	50.61 (35.37)	-	-28.42%	75.71 (24.85)	62.57 (27.72)	-	-17.36%	10.31**	0.58
	Positive $Emotions_i$	6.39 (2.84)	8.17 (3.69)	-	+27.86%	6.79 (2.39)	8.71 (3.07)	-	+28.28%	13.16**	0.62
	Negative $Emotions_i$	15.09 (4.71)	11.43 (4.96)	-	-24.25%	14.71 (3.38)	12.43 (2.65)	-	-15.50%	24.63***	0.88
	$Irritability_i$	33.73 (17.73)	30.23 (19.68)	-	-10.38%	39.79 (13.01)	38.50 (13.11)	-	-3.24%	0.84	0.17
Biological Outcome (N = 88)	Hair Cortisol	89.19 (102.71)	-	56.89 (60.76)	-36.22%	128.50 (126.35)	-	43.72 (20.34)	-65.98%	28.03***	0.53
	Hair Cortisol_log	4.04 (0.96)	-	3.75 (0.74)	-7.18%	4.48 (0.85)	-	3.67 (0.49)	-18.08%	37.09***	0.60

Notes. *N*'s ranges are due to missing data. For self-reported measures, since no significant differences were observed between T2 and T3 and in order to overcome missing data issue, only overall mean differences and effect sizes between T1 and T3 are reported. Hair cortisol is measured in pg/mg of hair. Balance R-R = score of the balance of risks and resources. Hair Cortisol_log = log transformation of hair cortisol scores. *** = $p < 0.001$; ** = $p < 0.01$.

Supplementary Material

Treating Parental Burnout: Impact of Two Treatment Modalities on Burnout Symptoms, Emotions, Hair Cortisol, and Parental Neglect and Violence

Maria Elena Brianda, Isabelle Roskam, James J. Gross, Aline Franssen, Florence Kapala, France Gérard, Moïra Mikolajczak

Supplementary Background

Several preliminary studies provide support for the idea that PB can be treated. Two pilot studies tested the effects of non-psychological interventions (a nurse-administered education program focused on child disability [1] and a lifestyle-nutritional counselling program [2]) on burnout in a sample of mothers. Despite the low statistical power, the results showed an overall reduction of self-reported burnout symptoms after the intervention. Two additional studies looked at group psychological interventions for parents with PB, with promising results. Lindström and colleagues [3] were the first to develop a group intervention for PB, specifically conceived for parents of chronically ill children, and this intervention focused on coping strategies. Their pilot study with a single-group design (N = 16) showed a significant reduction (9%) in self-reported burnout symptoms after the intervention. Anclair and colleagues [4] carried out a waiting list controlled pilot study (N = 19 participants in the experimental group, and N = 28 in the control condition), comparing the effects of cognitive behavioral therapy (N = 10) and mindfulness (N = 9) in a sample of parents suffering from PB who had children with chronic conditions. Their results showed a significant reduction in self-reported burnout symptoms in both experimental groups (25% and 23% respectively). Although preliminary, and evaluated via self-reports only, these studies show a significant reduction of burnout symptoms immediately after the interventions.

The group setting used in the foregoing studies seems particularly interesting and appropriate for parents suffering from PB in light of the extreme sense of loneliness and isolation that they report [5]. A sense of guilt and shame at not being a good enough parent and at having reprehensible feelings, thoughts, and behaviors towards children [5] prevents parents from sharing their difficulties with their close relatives and asking for help. The conviction that they are terrible, unworthy parents leads them to isolate themselves [5], exacerbating the lack of support. Offering them a setting in which they can meet other parents suffering in their

parental role may help to normalize their feelings. Moreover, the group can provide the support and comprehension necessary to say the unspeakable.

For these reasons, we decided to create, test, and compare two types of group interventions for parents suffering from PB: a relatively directive intervention (henceforth "Directive") and a relatively non-directive intervention (henceforth "Non-Directive"). The Directive intervention was based on the etiological process underlying PB. Although each parent lives in a different situation and "*each parental burnout has its own history*", research has highlighted a common etiological process [6]: the presence of a chronic imbalance between parental stressors and parental resources. In other words, in parents with PB, the balance between stress-enhancing factors (such as parental perfectionism, ineffective childrearing practices, or children with special needs) and stress-relieving factors (such as support from the coparent or the family or high emotional competences) is negative: their risks chronically outweigh their resources. In light of this, one way to treat PB is to try to help parents to "restore their balance". In doing so, it must be kept in mind that some factors play a greater role than others in causing PB. Whereas sociodemographic factors (i.e., the number and age of children) are less predictive of PB than might be expected, parenting factors (e.g., parenting role restriction, childrearing practices), family functioning factors (e.g., support from the coparent or the family) and parents' characteristics (e.g., parental perfectionism, level of emotional competences) play a stronger role [7] (see Sánchez-Rodríguez et al. [8] and Mikolajczak & Roskam [6] for reviews). An intervention aiming to restore parents' balance by actively working on improving these factors should therefore be effective in reducing PB symptoms.

The feasibility of this first intervention approach was suggested by the fact that the literature had already shown that parental practices, parental cognitions, coparenting, and emotional competences could be efficiently improved through targeted interventions. Extensive research

[9-11] supports the effectiveness of parenting programs aimed at improving parents' childrearing practices and parent-child relationships (e.g., the Triple P-Positive Parenting Program [12], the Parent-Child Interaction Therapy [13]), or adjusting parental cognitions, such as distorted beliefs or thoughts that might be influencing their behavior (e.g., the Incredible Years Parent Training [14], the Chicago Parent Program [15]). Some programs also provide tools to work on the coparental relationship and foster joint parenting (e.g., mutual support, childrearing agreements, division of duties) [16-19]. Others strengthen personal traits involved in stress management, such as emotional competences [20]. Together, these programs cover the main set of parenting, family, and personal factors which play an important role in PB. However, none of them proposes the full combination of all these factors within the same intervention. The few programs that offer a combination of some of these factors target parents of children with very specific characteristics (behavioral and emotional disorders, disabilities, physical or mental illnesses) [10, 21, 22]. One is the Parent Stress Management Training [23], a good example of multi-dimensional intervention specifically addressing parents with ADHD children. Given the importance of targeting all the above-mentioned factors and of addressing all kinds of parents, none of the existing parenting programs seemed an ideal fit for the treatment of parents suffering from PB. However, we took inspiration from some parts of these programs in developing a specific and comprehensive intervention, the evaluation of which was one of the goals of this study (details about the content of this intervention are provided in the Method section).

The second group program tested in this study, the Non-Directive intervention, was based on the assumption that directly targeting the etiological process of a condition is not the only way to help parents get better [24]. A non-directive setting that offers active, empathetic, caring, and authentic listening might provide participants with a sense of belonging and unconditional acceptance. This would allow participants to believe in themselves, to find their own path and

resources, and to perceive themselves as fundamentally worthy of consideration, regardless of their past actions [25-27]. Such a setting in which anything can be said (or not said) and any emotion, desire or need can be expressed, heard and respected, might be particularly therapeutic for parents with PB who often live in shame at what they feel – and, sometimes, at what they have thought, done, or not done with respect to their children – and in guilt arising from their own judgment and/or the disapproval of the outside world. Sharing their exhaustion, their actions, their guilt, their shame, their doubts, in a context where they could be heard without judgment might have a powerful therapeutic effect. According to users' reported perceptions of parenting group programs, acceptance and support from other parents play an important role in reducing participants' sense of inadequacy and improving their well-being [28]. In addition, a less directive approach, aiming to enhance parents' own resources to deal with parental challenges, would, it was thought, improve self-esteem and self-confidence [29, 30]. The evaluation of the effectiveness of this intervention constituted the second goal of this study (details about the intervention are provided in the Method section below).

Overall, with this study we wanted to extend previous research on the treatment of PB in five ways: (1) by comparing two theoretically contrasting interventions; (2) by investigating the effectiveness of these interventions in a sufficiently large sample of parents; (3) by including parents who did not necessarily have an ill child (since previous studies had shown that PB can also occur in the absence of such problems on the part of the child [5]), (4) by adopting a multi-method evaluation of intervention efficiency (by adding informant-reported and biological measures to the self-report evaluation), and (5) by assessing whether treatment effects were maintained over three months. Our focus was on testing the effectiveness of two 8-week group interventions (one aimed at "restoring the balance" of parents and one focused on active listening and group dynamics). Except for the assessment strategy used to examine

the spontaneous evolution of PB (see below), the study was conducted following the CONSORT guidelines for psychological randomized control trials (RCT) [31, 32].

Supplementary Materials and Methods

The study used a two-arm parallel-design for the evaluation of the effectiveness of two different group interventions for PB. The interventions under investigation have been implemented through 14 group treatments (eight in the Directive and six in the Non-Directive condition), which took place in different cities across Belgium. Randomization to the two experimental conditions was based on the self-enrollment of participants to a specific location (see below for further details). Participants were assessed at three testing times: pre-test (just before the intervention), post-test (after the intervention) and follow-up (three months after the end of the intervention). As explained below, a subset of participants provided measures four times: two months before the intervention, just before the intervention, after the end of the intervention and three months after the end of it. This subset of participants allowed to examine how parental PB spontaneously evolves in the absence of intervention. Participants voluntary chose to be assigned to this "waiting" condition. The ethical rationale for this and further details are provided in the procedure section.

CONSORT guidelines [31, 32] on RCTs were used to report the methods and results of this study. The database is publicly available on Open Science Framework (https://osf.io/kn7bw/?view_only=None).

Participants

The sample consisted of 142 parents at baseline (88.7% mothers), which is sufficient to detect moderate Time x Condition interaction effects with 95% power (G*Power). The vast majority of the participants (76.7%) were aged between 25 and 44 years (21.1% from 45 to 54 years and only 2.1% over 55). Participants had up to 6 children: 12.0% had one child, 73.2% had 2 or 3 children, and the rest more than 3 children. Most of the parents were in a relationship

(only 14.8% were single parents). 71.8% of participants had a bachelor's or a master's degree and 74.0% were employed at the time of the study (among the unemployed, 11.3% had a work disability). Net monthly household income was between €1,000 and €2,500 ($1,120 - 2,800) for 18.2% of participants, between €2,500 and €4,000 ($2,800 - 4,480) for 39.8%, between €4,000 and €5,500 ($4,480 - 6,160) for 28.4% and over €5,500 ($6,160) for the rest of the sample. Participants were recruited between January and February 2017, mostly through advertisements on the internet, but also through flyers and newsletters promoting the study across the networks of the largest mutual health insurance fund in the country. Health professionals were also informed about the study so that they could refer parents. The treatments under study were promoted as an 8-week (2h/week) group intervention for exhausted parents. Parents were not informed that there were two different interventions. The announcement also provided a brief description of PB symptoms, together with an explanation of formal aspects of the study (dates, locations, evaluation protocol). Parents wishing to participate could voluntarily apply for registration via an online application form. A brief screening questionnaire was administered in order to assess the following eligibility criteria: (i) having at least one child still living at home, (ii) experiencing PB symptoms, and (iii) being willing to engage in an intervention specifically aimed at reducing PB symptoms. PB was screened via a subscale of the *Parental Burnout Assessment – PBA* [33] consisting of nine items assessing emotional exhaustion symptoms experienced in the parental role (e.g., *I feel completely run down by my role as a parent*). Within the screening questionnaire, we asked parents to provide a brief explanation of the reasons why they wanted to join the group. We excluded parents whose main expectations were unrelated to reducing their PB symptoms (e.g., improving their child's development in cases of disabilities, resolving the couple's problems, etc.).

Procedure

14 groups were held in total in ten different cities across Belgium (see Table S1 for details). Cities were paired according to their average socio-economic level beforehand, in order to ensure that socio-economic background would not represent a confound in the comparison between the two experimental conditions. In four cases, two groups (one for each type of intervention) were proposed at the same location but at different dates.

After the online screening, eligible parents signed an informed consent form. Randomization was assured through participants' self-enrollment to a specific training location and date. Parents were blind to the existence of different treatment conditions. We invited participants to complete the assessment protocol three times: just before the beginning of the treatment (T1), immediately after the end of the eight-week intervention (T2), and three months after the end of the intervention (T3). Intervention groups were scheduled in two phases: seven groups received the intervention between March and April 2018, and seven groups between May and June 2018. Eligible parents who self-enrolled to the second wave groups were invited to complete the additional evaluation 8 weeks *before* the beginning of their group (which corresponds to the treatment length).

The reason why participants were not *randomly* assigned to the waiting list is primarily ethical. Since PB is associated with a high suicide risk [34, 35], we could not take the risk of allocating eligible parents (i.e. exhausted parents) to waiting list and forcing them to wait several months for the intervention. Instead, participants were free to register in whichever group they wanted, based on date and location, and voluntarily self-enrolled in the first or second wave. Thus the subset of participants who provided two pre-test measures consisted only of parents who were voluntarily waiting for the intervention.

All participants were informed about their right to withdraw from the research at any time, without compromising their participation in the group. To ensure the confidentiality of data,

participants were identified by anonymous codes. The design of the study was approved by the Institutional Review Board.

Interventions

The current study aimed to test the efficacy of two different group-based interventions for parents suffering from PB: a relatively directive intervention (henceforth "Directive") and a relatively non-directive intervention (henceforth "Non-Directive"). Each group of parents was led by two trained psychologists. In order to prevent any experimenter effect, 23 psychologists were recruited to lead the 14 groups. They were all clinicians either working at or cooperating with the faculty's consultation center for parenting. Most psychologists were assigned to only one condition (either Directive or Non-Directive, according to their own preference); five psychologists led one group in each condition (they had no preference and were willing to lead both, see Table S1). Except for PhD students and interns, all psychologists were paid for their participation in the study, in order to ensure maximum commitment. Psychologists were trained to lead the groups as follows. Note that, except for the psychology interns (N = 2) who were paired with an experienced psychologist, all psychologists were experienced clinicians who had previous experience in group therapy/interventions in other domains (parents of difficult children; psychiatric inpatients; stress-management groups; etc.). Thus, we did not need to train psychologists to run groups. Training therefore concerned: PB on the one hand, and the specific modality of their intervention on the other hand. All psychologists were invited to a general information session providing general information on PB (i.e., nature, prevalence, risk factors, consequences, testimonies). Psychologists in charge of the directive intervention group had to further read two books on PB: the first book provided detailed explanation on risk-factors for PB and the second book consisted of the Treatment Manual of this condition, providing detailed information about how to concretely work with parents to "restore the balance". Psychologists

in charge of the non-directive intervention received a 4-hour training in active listening and had to read the Treatment Manual of their condition, which provided further guidance about how to implement active listening and Rogerian attitudes in a group setting. The two detailed treatment manuals were created in order to ensure the standardization of treatment provided. The manual of the Directive intervention included specific instructions for each session: the strict timetable, a detailed explanation of each topic, descriptions of activities, exercises and materials to be employed, together with recommendations about the optimal attitude that would involve participants and respect their needs and expectations. The manual of the Non-Directive intervention was less structured and focused mostly on the attitude to adopt in order to ensure active listening and unconditional positive regard. It also identified and explained the theme of each session and gave an example of an activity or open-ended question to launch each session. Both manuals are available upon request from the first author for replication studies in other contexts. Except for their content (see below), the two treatment conditions shared the same format: the length of the intervention (8 sessions), the frequency of sessions (one session per week, between 1.5 and 2 hours) and the groups' size (between 6 and 15 participants).

Directive intervention

This consisted of a highly standardized 8-session group intervention founded on the Balance between Risks and Resources framework [6]. According to this framework, PB occurs when parental stress-enhancing factors outweigh parental stress-relieving factors. The main goal of this standardized and directive intervention was to "restore the balance" of participants. One session was thus entirely devoted to explaining the Balance between Risks and Resources and drawing up participants' personal balances. Since previous studies showed that personal, parental, and family/couple factors account for most of the variance in PB [7] (see also Mikolajczak & Roskam [6], Appendix A for a meta-analysis), the following topics were

addressed in turn: societal pressures on parents, parental perfectionism, emotional competences, parent-child relationship, coparental support, and ways to ask for help. The closing session focused on relapse prevention. Each session involved both psychoeducation and pre-established structured activities. At the end of every session, participants received exercises to do at home. The main goal of the intervention was to lead participants to select the best ways to minimize stress-enhancing factors and maximize stress-relieving factors in their specific situation. In parallel, psychologists helped parents identify uncontrollable stressors, in order to work on acceptance and avoid wasting what little energy they had trying to change the unchangeable.

Non-Directive intervention

This consisted of a semi-standardized, highly flexible 8-session experiential support group intervention. The intervention was standardized in the sense each session had a fixed duration and predetermined topic. Apart from that, the focus was primarily on common therapeutic factors, such as active listening, unconditional positive regard, and group dynamics, adopting a person-centered approach. Previous research on psychotherapy outcomes has revealed that such non-specific therapeutic factors account for 30% of patients' improvement [24]. In this study, the Non-Directive approach was chosen to examine the effect of an intervention the same length as the Directive one but in which psychologists did not provide any psychoeducation or advice to parents. The rationale for this was that although a directive intervention, rich in information and advice, might be helpful and reassuring for exhausted parents who felt that they were failing, the possibility could not be excluded that giving suggestions about better ways to handle things could be counterproductive for parents who had burned out precisely because of their tendency to fulfil too many external expectations. Moreover, the practical aspect of the directive intervention, consisting of exercises and activities to be done both in the session and at home, might be useful for parents who needed

to feel that they were immediately benefiting from the intervention, but could also be too hard for exhausted parents who were completely devoid of energy. In the Non-Directive intervention, much more time was devoted to free expression of participants' experiences and emotions, and to free exchanges, compared to the Directive intervention. Participants were regarded as self-examining agents with a metacognitive capacity to reflect on both their thoughts and actions as parents. The role of the psychologist was to start each session with an activity or specific open-ended questions (the pre-established standardized topics), aimed at activating parents' thinking skills and awareness about parenting challenges. The following topics were addressed in turn: demands and expectations regarding the intervention, the experience of PB, the onset and course of PB, being parents in the 21st century, PB consequences, and how participants see themselves as parents and how others see them. In addition, an entire session was devoted to a topic voluntarily chosen by participants. The closing session focused on relapse prevention.

Measures

In order to obtain a multi-dimensional, multi-method, and multi-informant evaluation of the effects of these two treatments, the assessment protocol consisted of self-reported outcomes (completed by parents who received the treatment), informant-reported outcomes (completed by each participant's partner or by an adult who saw the participant at least three times a week), and a biological measure of stress (the analysis of hair cortisol).

Sociodemographics

Participants were asked for the following sociodemographic information: gender, age range, marital status (in a couple or single parent), number of children, educational level (namely, the highest degree obtained), employment status (full-time, part-time or unemployed) and range of net monthly household income. For age and household income, we proposed several ranges and asked parents to note which one they belonged to (e.g., for age:

25-34 years old, 35-44 years old, etc.; for monthly income: between €1,000 and €2,500, between €2,500 and €4,000, etc.). We decided to gather participants' age in ranges rather than asking them for their precise age because precise values would have made it easy to identify participants in each group (for example, there was only one very young mother of four children). This would have compromised the anonymity of the study (which had not only been promised to participants but was also necessary since we were measuring variables such as neglect and violence and would otherwise have been legally required to report parents who were guilty of neglect or violence against their children).

Self-reported outcomes

Parental Burnout was assessed using the *Parental Burnout Assessment – PBA* [33], a 23-item self-report questionnaire designed on the basis of testimonies of parents suffering from PB, which is currently the reference measure used in the International Investigation of Parental Burnout - IIPB (a large study involving 40 countries and aiming to examine the prevalence and cultural variations of PB across the world). The items represent the most common PB symptoms (e.g. *I'm so tired out by my role as a parent that sleeping doesn't seem like enough*, or *I'm no longer able to show my children how much I love them*), and participants are asked to rate them using a 7-point frequency scale (never, a few times a year, once a month or less, a few times a month, once a week, a few times a week, every day). Test scores can therefore theoretically vary from 0 to 138. The PB score is computed by summing the item scores: higher scores reflect higher PB levels. In the current sample, Cronbach's alpha was 0.96 at T1, and 0.97 at T2 and T3.

Parental neglect was assessed with the Parental Neglect Scale [34], a 17-item questionnaire encompassing physical neglect (e.g., *I don't care about my children when I know I should (meals, hygiene, etc.)*), educational neglect (e.g., *I don't help my children when they really need it (for their homework, to make a decision, to resolve a conflict, etc.)*), and

emotional neglect (e.g., *I don't comfort my children when they are sad, frightened, or distraught*). Items are rated on an 8-point scale (never, less than once a month, about once a month, a few times a month, once a week, several times a week, every day, several times a day). Test scores can therefore theoretically vary from 0 to 119. The parental neglect score is obtained by summing the item scores: higher scores reflect higher frequency of neglectful behaviors. In the current sample, Cronbach's alpha coefficients were 0.84 at T1, and 0.82 at T2 and T3.

Parental violence was assessed with the Parental Violence Scale [34], a 15-item questionnaire encompassing verbal violence (e.g., *I say things to my children that I then regret: threats, insults, ridiculous nicknames, etc.*), physical violence (e.g., *I spank or slap my children*) and psychological violence (e.g., *I tell my children that I will abandon them if they are not good*). Items are rated on an 8-point scale (never, less than once a month, about once a month, a few times a month, once a week, several times a week, every day, several times a day). Test scores can therefore theoretically vary from 0 to 105. The parental violence score is obtained by summing the item scores: higher scores reflect higher frequency of violent behaviors. In the current sample, Cronbach's alpha coefficients were 0.85 at T1, 0.84 at T2 and 0.80 at T3.

Positive and negative emotions were assessed using an adaptation of the *Positive and Negative Affect Schedule – PANAS* [36] including the emotions most commonly affected by PB. The resulting items are grouped into two mood subscales: positive emotions (seven items: happy, enthusiastic, interested, serene, relaxed, peaceful, proud) and negative emotions (eight items: stressed, anxious, impatient, angry, guilty, upset, ashamed, and desperate). Participants are invited to appraise the frequency of each emotion during the previous four weeks using a 5-point scale (never, rarely, regularly, often, very often). Test scores can therefore theoretically vary from 0 to 28 for positive emotions, and from 0 to 32 for negative emotions.

Higher scores reflect higher frequency on positive and negative emotions. Cronbach's alpha coefficients in the current sample were 0.87 and 0.83 for the positive and negative subscales respectively at T1, 0.89 and 0.87 at T2, and 0.89 and 0.84 at T3.

Balance between Risks and Resources vis-à-vis PB was assessed via *the Balance between Risk and Resources – BR²* [6], which measures known protective and risk factors for PB through 39 bipolar items encompassing 11 levels, from -5 to +5. Test scores can therefore theoretically vary from -195 to +195. The negative pole represents the risk while the positive pole represents the corresponding resource. For example, -5: *My partner denigrates me as a mother/father*; +5: *My partner says that I am a good mother/father*. The BR² score is computed by summing the 39 items so that positive scores indicate that the parent has more (or more significant) resources than risks, and negative scores indicate that the parent has more (or more significant) risks than resources. Reliabilities have not been computed, as risk/protective factors are not necessarily expected to covary (i.e., for some parents, a poor coparental relationship may be associated to poor childrearing practices, but for others it may not).

Job burnout was assessed among employed parents only by the *Maslach Burnout Inventory-General Survey – MBI-GS* [37], a widely used measure of job burnout. The MBI-GS encompasses 16 items, and participants are invited to indicate how often they experience each symptom, on a 7-point frequency scale (never, a few times a year or less, once a month or less, a few times a month, once a week, a few times a week, every day). Test scores can therefore theoretically vary from 0 to 96. The MBI score is computed by summing the item scores: higher scores reflect higher job burnout levels. Cronbach's alpha coefficients in the current sample were 0.89 at T1, 0.90 at T2, and 0.89 at T3.

Medication. Participants were first asked if they were taking any of the following drugs in a yes or no format: Corticoids (e.g., cortisone); pituitary or steroid hormone;

antidepressant; tranquilizer; mood regulator; antipsychotic/neuroleptic; sleeping pill (see Supplementary Table S2). If the participants replied "yes" to any of the above, he was redirected for each "yes" to the following open-ended question: *Could you please indicate below the name of the medication, the dosage and how often you are taking this treatment?*

Informant-reported outcomes

We invited participants to ask their partner to complete the following questionnaires online. Single parents or parents with unavailable partners (e.g. divorced or separated couples) were invited to ask an adult close to them (and who saw them at least three times a week) to complete the questionnaire. In order to match self-reports and informants' reports, we asked informants to complete the protocol using the same identification code as their partner. We made it clear to participants that asking their partner to complete the questionnaires was not compulsory and that even if they agreed to forward the invitation, then partners (or close adults) must feel free to accept or refuse. We assured respondents that participants would never have access to their responses. We received 76 questionnaires completed at T1, 38 completed for a second time at T2, and only 21 at T3. The majority of completers (89.2%) were partners.

Partner's perception of the participant's symptoms of PB was assessed via the *Parental Burnout Assessment – informant form (PBA-i)*, an adaptation of the Parental Burnout Assessment [33] designed to capture the partner's perception of the parent's level of burnout. To this end, items of the PBA were converted into a hetero-evaluation format: e.g., the item *I'm so tired out by my role as a parent that sleeping doesn't seem like enough* became *I have the impression that my partner is so tired out by his/her role as a parent that sleeping doesn't seem like enough*. Test scores can theoretically vary from 0 to 138. The partner's perception of PB score was obtained by summing the item scores. Cronbach's alpha coefficients for the current sample were 0.96 at T1 and T2, and 0.97 at T3.

Partner's perception of participant's positive and negative emotions was assessed via the questionnaire measuring positive and negative emotions of participants (see above) converted into a hetero-evaluation format. Respondents were asked to indicate how often their partner had displayed five positive emotions (happiness, enthusiasm, interest, serenity, relax), and six negative emotions (stress, anxiety, impatience, irritation, anger, nervousness) over the previous four weeks. The rating system was the same as in the corresponding self-report questionnaire. Test scores can theoretically vary from 0 to 20 for positive emotions, and from 0 to 24 for negative emotions. Cronbach's alphas in the current sample were 0.84 and 0.79 for the positive and negative subscales respectively at T1, 0.86 for both subscales at T2, and 0.89 and 0.82 for positive and negative subscales at T3.

Partner's perception of participant's level of irritability was evaluated via an adaptation of the *Carer's Irritability Questionnaire – CIRQ* [38], a scale designed to measure relatives' perception of irritable behaviors. The original instructions were adjusted to our purpose so that respondents were asked to rate the frequency of their partner's irritable behaviors using an 8-point scale (never, less than once a month, about once a month, a few times a month, once a week, a few times a week, every day, several times a day). Two CIRQ items (*He/she loses his temper and snaps or shouts at others* and *He/she has threatened violence against him/herself or others*) were split into two in order to assess the two parts of each question separately. Three extra items were added in order to fully capture the variety of expressions of irritability among parents with PB (*He/she is quite sensitive to others' remarks*, *Lately, he/she has felt bitter about things* and *He/she's been feeling like a bomb, ready to explode*). The final version consisted of 15 items: higher scores reflected higher levels of irritability. Test scores can therefore theoretically vary from 0 to 105. Cronbach's alpha coefficients in the current simple were 0.86 at T1, 0.91 at T2 and 0.89 at T3.

Biological outcomes

Hair cortisol levels, which provide an indication of chronic stress over the previous three months [39-41], were measured through hair samples of approximately 150 strands of hair collected from the posterior vertex of the head [42]. Hair samples were cut proximal to the scalp, wrapped in a piece of aluminum foil (to protect them from light and humidity) and stored in an envelope at room temperature. They were sent for analysis to a hair cortisol specialist at the Faculty of Pharmacy of the University of Granada. A length of 3 cm from the scalp was retained: assuming an average growth rate of 1 cm/month, a 3 cm segment contains cortisol that has been deposited over approximately the last 3 months. The hair samples were weighed and ground to a fine powder to break up the hair's protein matrix and increase the surface area for extraction using a ball mill (Bullet Blender Storm, Swedesboro NJ). Cortisol from the interior of the hair shaft was extracted into HPLC-grade methanol by incubation of the sample for 72 hours at room temperature in the dark with constant inversion using a rotator. After incubation, the supernatant was evaporated until completely dry using a vacuum evaporator (Centrivac, Heraeus, Hanau, Germany) and the extract was reconstituted in 200 ul of phosphate buffered saline at a pH of 8.0. The reconstituted sample was immediately frozen at -20ºC for later analysis [43, 44]. The cortisol in the hair sample was measured using the Salivary ELISA Cortisol kit© with the reagent provided, following the manufacturer's directions (Alpco Diagnostics®, Windham, NH).

Statistical Methods

In the first step, we checked the comparability of the subset of "waiting" participants and the rest of the sample, and also checked the comparability of the Directive and the Non-Directive groups with respect to socio-demographic data and PB at baseline. χ^2 tests for categorical variables and a t-test for PB measure were employed.

In the second step, we performed Pearson's correlations among all variables at baseline, in order to look at zero-order relations among self-reported, informant-reported, and biological measures.

In the third step, we investigated whether there was any significant predictor of drop-out or data missingness, in order to further determine whether we should control for any specific variable in the subsequent analyses. We conducted a binary logistic regression analysis to explore the effect of nine possible drop-out predictors (gender, age range, marital status, number of children, educational level, employment status, net monthly income range, type of intervention received and PB at baseline) on the drop-out binary (completers vs. non completers) dependent variable.

In the fourth step, we examined whether and to what extent outcome variables spontaneously changed with time in the absence of intervention. This analysis was therefore performed on the 40 "waiting" subjects. We performed on this subsample a repeated measure MANOVAs with Time (baseline vs. 8 weeks later) as the within-subject factor, and the following self-reported dependent variables: PB, the presence of neglectful and violent behaviors towards children, positive and negative emotions, the Balance between Risks and Resources, and job burnout. Given that 36.6% of participants were unemployed – and therefore did not complete the job burnout questionnaire –, we ran two MANOVAs: one on employed parents including the job burnout measure, and the other on the entire sample excluding the job burnout measure. Another repeated measure MANOVA was performed with Time (baseline vs. 8 weeks later) as the within-subject factor and the informant-reported outcomes: partner's perception of PB, positive and negative emotions, and irritability. In addition, we compared "waiting" participants in their first two evaluations (first pre-test vs. second pre-test) with the rest of the sample (including both Directive and Non-Directive conditions) in their first two evaluations (pre-test vs. post-test). We ran paired sample t-tests and estimated within-subject

effect sizes for self-reported and informant reported outcomes in each group. Then, we performed independent sample t-tests and estimated between-subject effect sizes to compare the amount of change (i.e., the difference between the first and the second evaluation) between the two groups. Cohen's d values were computed as the difference between the means of two groups (paired or independent) divided by the pooled standard deviation. We did not include the biological outcome in these analyses, as it had only been collected at pre-test and follow-up.

In the fifth step, we compared the two treatment modalities (Directive vs Non-Directive). First, we ran preliminary analyses in order to assess whether effects varied as a function of whether the intervention was immediate (Wave 1) or delayed (Wave 2). Since repeated measure ANOVAs showed no significant effect in the interaction of Time (within-subject factor) X Wave (between-subject factor), and in order to maximize statistical power, we pooled Wave 1 and Wave 2 together. We first looked at self-reported outcomes. Given that over 30% of the respondents failed to complete the follow-up assessment (T3) and in order to keep maximum statistical power, we analyzed T1-T2 first and then T2-T3 and T1-T3. Note that analyzing T1, T2, T3 together in the same MANOVA leads to similar results. We started by examining the effects immediately after the intervention. A repeated measure ANOVA was performed for each outcome variable with Time (T1 vs. T2) as the within-subject factor and Group (Directive vs. Non-Directive) as the between-subject factor. Then we examined whether the effects observed at post-test on self-reported outcomes held stable three months after the end of the intervention (T3). Several separate repeated measure ANOVA tests were performed with Time (T2 vs. T3, and T1 vs. T3 separately) as the within-subject factor and Group (Directive vs. Non-Directive) as the between-subject factor. Cohen's d effect sizes of Time and Time X Group effects (T1 vs. T3) were calculated for each variable as the ratio of the mean of scores' differences over the standard deviation of the differences. We also

performed repeated measure MANOVAs including all the self-reported outcomes, with Time (T1 vs. T2, T2 vs. T3, and T1 vs. T3) as the within-subject factor and Group (Directive vs. Non-Directive) as the between-subject factor, in order to check if this yielded similar results. In addition, we performed an extra repeated measure MANOVA with the same between-group factor, the same outcome variables, through three testing time points (T1 vs. T2 vs. T3), in order to explore the evolution of outcomes over time.

We then examined the effect of the interventions on informant-reported outcomes, namely the partner's perception of participants' PB, positive and negative emotions, and irritability. Repeated measure ANOVAs were run for each outcome variable, with Time (T1 vs. T2) as the within-subject factor and Group (Directive vs. Non-Directive) as the between-subject factor. Cohen's d effect sizes were calculated for each variable. In addition, a repeated measure MANOVA was performed in order to check if it yielded similar results. We decided not to consider the three-month follow-up of informant-reports because of the excessive missing data and resulting very low statistical power (only 21 informants – 12 for the Directive intervention and only 9 for the Non-Directive intervention – completed the protocol three months after the end of the interventions).

It should be noted that in addition to ANOVA and MANOVA tests, we conducted intention-to-treat analyses using multilevel modeling (MLM) [45], in order to take into account missing data and non-independence in observations. Both approaches – ANOVA (and MANOVA) tests, and intention-to-treat analyses – have pros and cons in the evaluation of treatments. ANOVA tests evaluate the effectiveness of the interventions based solely on participants who actually received the allocated intervention and completed the assessment protocol. However, removing from the analyses participants who have dropped out or failed to complete the evaluations may lead the intervention's effectiveness being overestimated, as participants who persist are usually also more motivated and tend to do better than those who fail to complete

the program or withdraw [46]. On the other hand, the intention-to-treat approach preserves the benefits of randomization, accommodating missing data and taking into account all available information of randomized participants, regardless of their degree of adherence to the protocol. However, if there is a high volume of missing data, the effectiveness of the tested intervention may be underestimated [46]. For these reasons, we chose to conduct our analyses with both approaches, in order to check if they pointed in the same direction, and avoid the risk of over- or underestimating the effects.

We then examined the effects of the tested interventions on the biological measure: the level of hair cortisol. Because of the huge inter-individual variations and in line with usual practice as reported in the hair cortisol literature, hair cortisol scores were log-transformed before analysis, and considered as the outcome variable of a repeated measure ANOVA with Time (T1 vs. T3) as the within-subject factor and Group (Directive vs. Non-Directive) as the between-subject factor.

Lastly, we assessed whether the effects varied according to psychologists' preferences for a specific intervention. Groups were divided into groups led by psychologists who chose to lead only the preferred condition (henceforth "Preference"), and groups in which at least one of the two psychologists leading the sessions had no preference and led either type of intervention (henceforth "No-Preference"). Repeated measure ANOVAs were performed for self-reported, informant-reported, and biological outcome variables, with Time (T1 vs. T2) as the within-subject factor and Psychologists' Preference (Preference vs. No-Preference) as the between-subject factor.

Supplementary Results

Descriptive results first indicated that parents who enrolled in the interventions were not suffering from ordinary parenting stress but that they were really severely distressed. This was

reflected not only in the parents' average score on the Parental Burnout Assessment (i.e., 85, more than 2.5 SD higher than the mean score of Belgian Parents, viz. 25, see Roskam, Brianda & Mikolajczak [33]) but also in the very high levels of hair cortisol exhibited by these parents. Parents enrolling in the current interventions had a mean level of hair cortisol level twice as high as control parents (we recruited 61 demographically matched control parents as part of another study) and even higher than that of patients suffering of severe chronic pain [47].

Comparability between groups at baseline

Preliminary analyses on socio-demographics and PB at baseline confirmed the comparability between "waiting" participants, and between the Directive and Non-Directive groups (see Table S3 for details).

Zero order relations among self-reported, informant-reported, and biological measures at baseline

Globally, Pearson's correlations at baseline showed that self-reported measures were significantly related to one another, as were informant-reported and biological measures (see Table S4). As previously shown, self-reported PB had moderate to high correlations with all other self-reported measures (and especially with parental neglect and violence), and was moderately related to informant-reported PB as well. The latter was highly related to all other informant-reported measures. Biological measures were not significantly related to most self-reported and informant-reported measures.

Participants' flow and analysis of missingness

Of the 151 parents who received and completed the 8-week intervention (82 in the Directive condition and 69 in the Non-Directive condition), 142 (94.04%) completed the protocol assessment at T1, 107 (70.86%) at T2 and 96 (63.58%) at T3 (see Figure S1). None of the

considered predictors explained the likelihood of dropping out or failing to complete the questionnaires (logistic regression model: $\chi^2(8) = 8.41$, $p = 0.40$). Therefore, and taken together with the results of Step 1, no control variable was included in the analyses of the subsequent steps.

Assessing spontaneous changes with time

A repeated measure MANOVA was employed to assess changes during the period between the two assessments prior to the intervention (n = 40 "waiting" participants) for both self-reported measures and informant-reported measures. The results showed no significant effect of time for any of the outcome variables (see Table S5 for details). Observed means and Cohen's d estimations generally showed very small effects of time for self-reported and informant-reported outcomes, with the exception of informant-reported positive and negative emotions (see Table S6 for more details), compared to the rest of the sample (which includes both Directive and Non-Directive conditions). Comparing the amount of change of each outcome in the absence and in the presence of intervention, we observed medium to large between-subject effect sizes for self-reported outcomes, with the exception of the Balance between Risks and Resources and job burnout, and for informant-reported outcomes, with the exception of irritability. Given that there was virtually no change during the waiting time and in order to preserve maximum statistical power for comparing the two interventions, we

focused the subsequent analyses on pre-post intervention test only, and removed the first pre-test ("pre-pre test").

Evolution of parents after the interventions

Means, standard deviations, mean differences across testing times and effect sizes of ANOVA tests are reported in Table 1 (main manuscript).

Self-reported outcomes

96 participants completed measures at pre- and post-test and 68 at pre-test, post-test and follow-up. The results of the repeated measure ANOVA tests performed on each variable are described below. The results revealed a significant main effect of time, and no significant Group x Time interaction, indicating that parents evolved similarly in the two intervention groups. Repeated measure ANOVA tests with Time (T2 vs. T3) as the within-subject factor showed that the effects remained stable three months after the end of the interventions, and even significantly increased in some cases (see below). Note that analyzing all self-reported outcomes together in the same MANOVA yielded similar results to individual ANOVAs. And as shown in Table S7, intention-to-treat analyses also yielded similar results: effects which were significant in the (M)ANOVAs remained significant, and effects which were non-significant remained non-significant.

Parental burnout. There was a significant main effect of time (T1 vs. T2) on PB immediately after the end of the treatment ($F(1,94) = 65.05$, $p < 0.001$, $\eta^2 = 0.40$, Cohen's $d = 0.83$), indicating that both interventions were effective in reducing PB. A significant effect of time was also found between T2 and T3 ($F(1,72) = 7.27$, $p < 0.01$, $\eta^2 = 0.09$, Cohen's $d = 0.29$), showing that PB symptoms continued to decrease significantly three months after the end of the interventions. Overall, we observed a large effect of time between before and three months after the intervention (T1 vs. T3) on PB ($F(1,84) = 64.24$, $p < 0.001$, $\eta^2 = 0.43$, Cohen's $d = 0.86$).

Parental neglect. There was a significant main effect of time (T1 vs. T2) on the frequency of neglectful behaviors immediately after the end of the treatment ($F(1,93) = 17.14$, $p < 0.001$, $\eta^2 = 0.16$, Cohen's $d = 0.43$), indicating that parental neglect decreased in both interventions. At T3 this effect remained stable and a significant Group X Time effect emerged ($F(1,83) = 5.16$, $p < 0.05$, $\eta^2 = 0.05$), indicating a greater reduction in the Directive condition (Cohen's $d_{Directive} = 0.63$) compared to the Non-Directive condition (Cohen's $d_{Non-Directive} = 0.22$).

Parental violence. There was a significant main effect of time (T1 vs. T2) on the frequency of violent behaviors immediately after the end of the intervention ($F(1,93) = 24.73$, $p < 0.001$, $\eta^2 = 0.21$, Cohen's $d = 0.52$), indicating that parental violence decreased in both interventions. This effect remained stable at T3 (Cohen's $d = 0.58$).

Positive and negative emotions. There was a significant main effect of time (T1 vs. T2) on positive emotions ($F(1,92) = 25.31$, $p < 0.001$, $\eta^2 = 0.21$, Cohen's $d = 0.54$) and negative emotions ($F(1,91) = 73.23$, $p < 0.001$, $\eta^2 = 0.45$, Cohen's $d = 0.91$) immediately after the end of the intervention, indicating that positive affectivity increased and negative affectivity decreased in both interventions. Time effect remained stable at T3 both for positive emotions (Cohen's $d = 0.50$) and for negative emotions (Cohen's $d = 0.91$).

The Balance between Risks and Resources. No significant main effects of Time or Group X Time were found immediately after the end of the intervention (T1 vs. T2) on the Balance between Risks and Resources. However, three months after the end of the intervention we found a significant main effect of time ($F(1,84) = 8.17$, $p < 0.01$, $\eta^2 = 0.09$, Cohen's $d = 0.31$), indicating that both interventions were accompanied by a restoration of the balance.

Job Burnout. As expected given that the interventions were focused on *parental* burnout, no significant main effects of Time or Group X Time were found immediately after

the intervention (T1 vs. T2). At T3, the main effect of Time was still not observed, but we found a significant Group X Time interaction ($F(1,60) = 3.88$, $p = 0.05$, $\eta^2 = 0.06$), indicating a reduction of job burnout symptoms in the Directive intervention (Cohen's $d_{Directive} = 0.43$), but not in the Non-Directive intervention (Cohen's $d_{Non\text{-}Directive} = 0.09$).

In order to better visualize the evolution of all self-reported outcomes, we performed a repeated measure MANOVA with Time (T1 vs. T2 vs. T3) as the within-subject factor and Group (Directive vs. Non-Directive) as the between-subject factor. Although we only had 36 completers for the Directive group and 28 for the Non-Directive group (which is the reason why we analyzed T1-T2 differences, T2-T3 differences, and T1-T3 differences separately), bar plots provided an interesting overview, which illustrate the evolution of the outcomes through the three testing times (Figure S2).

Informant-reported outcomes

The statistical power was much lower for informant-reported outcomes, because only 37 respondents (23 for the Directive and 14 for the Non-Directive conditions) completed the informant-reported assessment before and immediately after the intervention. Due to the lack of statistical power necessary to draw valid conclusions, informant-reported data at T3 were not analyzed (only 12 participants from the Directive group and only 9 from the Non-Directive group completed the protocol three months later). Repeated measure ANOVA showed a significant effect of Time (T1 vs. T2) on informant-reported variables, with the exception of irritability. Since no significant effect emerged from the interaction Group X Time, we report only main effects of Time below (see also Table 1 in the main manuscript).

Informant's perception of the participant's symptoms of parental burnout. Informants (partner or someone close to the participant) perceived a significant reduction of PB after both interventions ($F(1,35) = 10.31$, $p > 0.01$, $\eta^2 = 0.23$, Cohen's $d = 0.58$).

Informant's perception of participant's positive and negative emotions.

Informants noted a significant increase in positive emotions ($F(1,35) = 13.16$, $p < 0.01$, $\eta^2 = 0.27$, Cohen's $d = 0.62$) and a significant decrease in negative emotions ($F(1,35) = 24.63$, $p < 0.001$, $\eta^2 = 0.40$, Cohen's $d = 0.88$) after both interventions.

Biological outcomes

Of the 151 participants who received one of the two interventions, 88 provided hair samples. Results of the repeated measure ANOVA on log-transformed hair cortisol scores showed a significant effect of Time ($F(1,86) = 37.09$, $p < 0.001$, $\eta^2 = 0.28$, Cohen's $d = 0.60$) on cortisol concentration, indicating that hair cortisol decreased in both interventions. In terms of raw data (before log-transformation), the mean hair cortisol concentration was reduced by 52% (from 107.06 to 50.91). Moreover, we found a significant Group X Time interaction ($F(1,86) = 8.23$, $p < 0.01$, $\eta^2 = 0.06$), indicating a greater reduction in the Non-Directive group (-65.98%, Cohen's $d_{Non\text{-}Directive} = 0.72$) compared to the Directive group (-36.22%, Cohen's $d_{Directive} = 0.36$).

Supplementary Discussion

PB is a serious condition, with damaging consequences for both parents and children. The main goal of this study was to test the effectiveness of two treatment modalities for PB in a sample of parents suffering from PB (intervention manuals are available upon request from corresponding author). More specifically, we compared a highly standardized and directive group intervention, which aimed to minimize parental stressors and optimize parental resources (based on the PB theory of the Balance between Risks and Resources [6]), to a highly flexible and non-directive experiential support group, mainly focusing on active listening, unconditional positive regard, and group dynamics. A subset of participants assessed two times before the beginning of the intervention allowed us to examine whether the outcomes of interest spontaneously change over time.

First, our results indicate that whereas PB tends to remain stable in the absence of intervention – neither participants nor their partners reported a spontaneous reduction of PB symptoms over a period of two months before the intervention – it decreased significantly after both interventions. Participants in both treatment modalities experienced a considerable decrease in PB symptoms (up to 37% decrease on average) after the intervention, and this improvement was maintained three months after the end of the intervention. Parents' perceptions were in line with their partners' (or adults close to them), who noted a 28% drop in PB symptoms following the intervention. These results dovetail with those observed regarding participants' emotional state, which improved after the intervention and over the follow-up period: parents and their partners observed an increase in positive emotions (reported increases of 18% and 28% respectively) together with a decrease in negative emotions (decreases of 29% and 24% respectively).

The improvement in PB after the intervention was mirrored by an improvement in parents' behaviors. Parents reported on average a 35% decrease in neglect and a 32% decrease in violent acts towards their children following the intervention. The frequency of these behaviors remained stable in the absence of intervention. This improvement in PB and parental behaviors is not surprising considering the strong reduction in hair cortisol concentration following the intervention (-52%): Martorell and Blunt Bugental [48] have shown that increases in cortisol levels — measured via saliva samples — predict an increase in the use of harsh parenting practices and mediate the impact of parent's and child's characteristics (i.e., parental powerlessness and child's difficult temperament) on these behaviors. Given that cortisol levels decreased by 52%, one may wonder why PB symptoms decreased by "only" 37%. One very likely explanation lies in scale differences: contrary to cortisol which is a continuous variable with constant intervals, intervals on the response scale of PB are not constant (i.e., never, a few times a year, once a month or less, a few times a

month, once a week, a few times a week, every day). Namely, while for cortisol, the weight interval between 1 and 2 picograms is exactly the same as the interval between 5 and 6 picograms, for PB, the time interval between 1 and 2 (months of differences) strongly differ from the time interval between 5 and 6 (days of differences). Thus, any decrease on PB scores corresponds to an exponential reduction in burnout symptoms – which is perfectly consistent with the greater reduction of cortisol levels observed here.

As regards the comparison between the two interventions, both approaches were accompanied by an improvement of the psychological and physical state of the parents. Although the evolution of parents was comparable in many ways in the two interventions, it differed regarding a few outcomes. For instance, among parents who attended the highly directive intervention that aimed to "restore the balance", we observed a greater reduction of parental neglect, compared to parents who attended the less directive intervention. Among parents who attended the less directive intervention that was mainly focused on active listening, positive regard, and group dynamics, we observed a more pronounced decrease in cortisol levels. We are reluctant to speculate on possible explanations until these interaction effects are replicated. However, future studies will benefit from further addressing these issues in order to determine not only whether these interventions lead to distinct benefits, but also who benefits the most from each type of intervention, as it is possible that different profiles of parents may be differently susceptible to the methods proposed [49-51].

The current study contributes to the literature on several related subjects in various ways. As regards PB, it provides large-scale evidence that PB can improve following a psychological intervention, as suggested by previous pilot studies [3, 4]. Although further studies are needed to assess the benefits of the group setting compared to individual or couple settings, our findings provide encouraging results about the potential of the group format, which is good news as group settings are more cost-effective and thus more affordable for parents.

Moreover, it seems that a framework of active listening, empathy, and comprehension, along with an invitation to consider selected topics relevant for parents with PB, may be sufficient to achieve positive and lasting effects on parents' well-being. These results suggest that it may not be necessary to be expert in PB to deliver effective support for parents suffering from PB, at least in a group setting. Psychologists trained in active listening and with expertise in group interventions have the necessary skills to help parents suffering from PB, which is also very good news given the high prevalence of this condition.

Another important contribution concerns the use of hair cortisol concentration as a biomarker of the effectiveness of burnout interventions [39, 52, 53]. In hair cortisol literature, most studies have looked at cross-sectional associations with a variety of clinical conditions (e.g. depression or generalized anxiety), and psychosocial sources of stress (e.g. unemployment or major life stressors; see Staufenbiel et al. [40] for a review). Less is known about the extent to which hair cortisol can be changed following a psychological intervention. In our sample, PB symptoms and hair cortisol were not correlated at baseline, which is likely due to the fact that all parents were in burnout. Indeed, when we pool this clinical sample with a sample of control parents who also completed the PBA, a linear correlation comparable in size to that observed here with job burnout emerges. Nevertheless, the current study provides the first evidence that hair cortisol is sensitive to change after psychological interventions for burnout and therefore suggests that it can be reliably used as an objective indicator of effectiveness in other clinical trials on PB.

Beyond its contribution to the literature on PB and hair cortisol, this study also contributes to the literature on children's health and maltreatment. Going a step further than previous cross-lagged longitudinal studies [35], these results confirm the presence of a large, direct, and causal relation between PB and child maltreatment: when PB is experimentally reduced via an intervention, neglectful and violent behaviors decrease in equal proportions to PB

symptomatology. These results provide the first causal evidence between PB and neglectful/violent behaviors towards children. They also confirm the status of PB as a new, important, yet malleable risk factor for child maltreatment. In this way, our study also responds to the call in the child maltreatment literature for further investigation of the prevention of child abuse and neglect through training programs specifically addressing major parental risk factors [54]. Moreover, our findings draw attention to the potential of parenting group-delivered interventions in the field of child maltreatment. While past research seemed to show that the group setting is insufficient to decrease child maltreatment unless accompanied by individual support [55], our results suggest that if child maltreatment is caused by PB, the group setting is in fact effective. Another important contribution of this study concerns the assessment of parental neglect and violence. By showing that neglect and violence scores do not change in the absence of intervention but that they are responsive to treatment and to the decrease of PB and hair cortisol, the current results provide further validation of the Parental Neglect and Violence Scales [34]. While the majority of previous studies have mainly focused on changes of parental *attitudes* towards maltreatment [56], these scales assess the *frequency* of specific neglectful and violent parental behaviors, in order to better capture how parents concretely interact with children.

In spite of the fact that this multi-method study was conducted and reported in compliance with CONSORT guidelines [31, 32], it has some limitations that must be acknowledged. First, this study does not allow to formally test the possibility that the benefits of the intervention lies in very non-specific factors, such as the mere presence of a group, or the mere fact that, by participating to the group, parents take time for themselves. It may also be due to what is called “non-specific factors” in psychotherapy, i.e., attention, empathy, warmth, positive consideration of the therapist, … [57]. Given the comparably positive effect of the directive intervention, it is conceivable that the two groups did not differ in their outcome because of

non-specific factors [57], which may be of primary importance in this population. A fruitful avenue for future research would consist in examining the effect of a combined intervention. If the efficiency of the current interventions are solely due to non-specific factors, this combination should not lead to stronger effects. If their efficiency is at least partially due to other processes, it should be twice as efficient at equivalent duration. A second limitation is the extent of missing data: whereas only seven participants dropped out during the intervention, many participants failed to complete the assessment protocol at the third time of testing. A third limitation is the use of only second-wave participants as WLC. Given the risks associated with PB (and especially the suicide risk), we opted for an immediate cross-over of WLC into the experimental condition, rather than a three-arm randomization. Although this approach was the only one ethically possible here, it made it impossible to conduct independent group comparisons among different conditions (as WLC participants later become experimental group participants), and ruled out the assessment of spontaneous changes over the follow-up period. A fourth limitation is the self-selection of participants in this study. Although self-selection reflects the real-world nature of mental health services and can also be viewed as a strength as it gives clinicians an idea of the exact amount of improvement to be expected in clinical practice (see Albert et al. [58], p. 384) future studies might prefer to rely on participants referred to by General Practitioners or Psychiatrists. A possibly related limitation is the scarcity of fathers in our sample, which is nevertheless in line with previous literature showing that fathers are particularly underrepresented in parenting studies [59, 60]. Future investigations might consider the possibility of organizing groups for fathers only, led by two male psychologists. This could facilitate fathers' participation. A fifth limitation is that our findings are entirely based on mean group scores, and nothing is known about what lies behind these means. Further investigations would benefit from adopting a person-centered approach in order to determine what type of parent benefits most from each type of intervention and who are the parents who do not benefit from any of them [49-51]. A

final limitation is inherent to individual approaches to psychological conditions, which by definition change individuals but not macro-societal factors. However, the current Western parenting culture, and specifically the ideology of "intensive parenting", has significantly increased pressure on parents [61] and the conviction that they are responsible for all their children's outcomes, making them live under the constant threat of negative long-lasting effects of their decisions on their children's lives [62, 63]. Although we have integrated these pressures within the current PB intervention programs, macro-societal – and not only individual – changes are necessary in the long run [61, 64, 65].

Another note of caution concerns the possible side-effects of the interventions. Because all the outcomes that we collected were reported in the paper, and because the impact of the interventions was positive on all of them, we don't have information about outcomes that might have deteriorated after the interventions. However, this does not mean that there are no side effects of the interventions. One possible side-effect concerns couple functioning. While we can easily imagine that some couples function better after the intervention (the parent is less stressed and less irritable and, as a result, the couple relationship is more satisfying), some couples may have deteriorated after the intervention, especially if the participating parent start to judge or criticize the other parent for not adopting the right parental practices, or if the non-participating parent feels excluded from what is experienced during the group sessions.

In summary, the current study meets the need to develop effective interventions for PB, a serious and chronic condition, which affects a significant proportion of parents and leads to a number of harmful consequences for both the parent and the children. Our findings show that PB can be effectively treated through short-term group interventions. Very positive effects were observed – immediately after the intervention and during the following months – in parents (statistically and clinically significant reduction of PB symptomatology, hair cortisol

levels, and frequency of negative emotions, along with an increase of positive emotions) and their children (statistically and clinically significant reduction of neglectful and violent parental behaviors). This study provides large-scale evidence of the effectiveness of the group setting, proposing a cost-effective approach to meet the needs of parents with PB effectively and promptly.

Statements

Acknowledgements

We would like to thank R.G. for her contribution to the hair cortisol analysis and all the psychologists who led the interventions: J.A., N.B., E.B., N.B., L.C., E.D., S.D., C.G., S.G., A.G., A.G., A.G., A.G., S.H., S.H., F.J., L.L., M.R., J.R., R.C., K.S., C.S. and M.T.. Finally, we would like to express our gratitude to the parents for their participation and trust.

Statement of Ethics

The study reported here was approved by the Institutional Review Board and was carried out in accordance with the provisions of the World Medical Association Declaration of Helsinki. All participants gave written informed consent. The study protocol was approved by the research institute's committee on human research.

Disclosure Statement

M.E.B., J.J.G., A.F., F.K. and F.G. have no conflicts of interest to declare. M.M. and I.R. have now founded the *Training Institute for Parental Burnout* which delivers training on parental burnout to professionals. The institute was founded after the completion of the study (including the analysis of the results). Thus, the institute did not participate in the funding of this study nor did it influence the process or the results in any manner.

Funding Sources

This study was co-funded by a public center for the promotion of health, Infor Santé (Mutualité Chrétienne) and an FSR-2016 Grant of UCLouvain to M.M. and I.R.

Author Contributions

M.E B., M.M., I.R., A.F., F.K., and F.G. developed the study concept and the study design. M.E.B., I.R., A.F., F.K. and F.G. set up the study and handled the logistics. M.E.B., A.F., F.K., and F.G collected the data. M.E.B. performed the data analysis, under the supervision of M.M. and I.R. M.E.B., M.M., I.R. and J.G. interpreted the data. M.E.B. drafted the manuscript under the supervision of M.M., I.R. and J.J.G. M.M., I.R. and J.J.G. provided critical revisions. All authors approved the final version of the manuscript for submission.

Open Practices Statement

The experiment reported in this article was not formally preregistered. However, the data have been made available on a permanent third-party archive: Open Science Framework. The database is available at https://osf.io/kn7bw/?view_only=None.

References

1 Bilgin S, Gozum S: Reducing burnout in mothers with an intellectually disabled child: an education programme. J Adv Nurs 2009;65:2552-5261.

2 Kessler CS, Eisenmann C, Oberzaucher F, Forster M, Steckhan N, Meier L, Stapelfeldt E, Michalsen A, Jeitler M: Ayurvedic versus conventional dietary and lifestyle counseling for mothers with burnout-syndrome: A randomized controlled pilot study including a qualitative evaluation. Complement Ther Med 2017;34:57-65.

3 Lindström C, Åman J, Anderzén-Carlsson A, Lindahl Norberg A: Group intervention for burnout in parents of chronically ill children–a small-scale study. Scand J Caring Sci 2016;30:678-686.

4 Anclair M, Lappalainen R, Muotka J, Hiltunen AJ: Cognitive behavioural therapy and mindfulness for stress and burnout: a waiting list controlled pilot study comparing treatments for parents of children with chronic conditions. Scand J Caring Sci 2018;32:389-396.

5 Hubert S, Aujoulat I: Parental Burnout: When Exhausted Mothers Open Up. Front Psychol 2018;9:1021-1021.

6 Mikolajczak M, Roskam I: A theoretical and clinical framework for parental burnout: The Balance between Risks and Resources (BR²). Front Psychol 2018;9:886.

7 Mikolajczak M, Raes M-E, Avalosse H, Roskam I: Exhausted parents: sociodemographic, child-related, parent-related, parenting and family-functioning correlates of parental burnout. J Child Fam Stud 2018;27:602-614.

8 Sánchez-Rodríguez R, Perier S, Callahan S, Séjourné N: Revue de la littérature relative au burnout parental. Can Psychol 2019

9 Barlow J, Coren E: The Effectiveness of Parenting Programs: A Review of Campbell Reviews. Res Social Work Prac 2018;28:99-102.

10 Bennett C, Barlow J, Huband N, Smailagic N, Roloff V: Group-based parenting programs for improving parenting and psychosocial functioning: A systematic review. J Soc Social Work Res 2013;4:300-332.

11 Löfgren HO, Padyab M, Ghazinour M, Nilsson K, Petersen S, Hägglöf B: Healthier Parents-Effects of Parent Training Programs on Mental Health. International Journal of Social Work and Human Services Practice 2017;5:70-79.

12 Sanders MR, Kirby JN, Tellegen CL, Day JJ: The Triple P-Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. Clin Psychol Rev 2014;34:337-357.

13 Hembree-Kigin TL, McNeil CB: Parent—child interaction therapy. New York: Springer Science & Business Media, 2013.

14 Webster-Stratton C, Reid MJ: The incredible years parents, teachers and children training series: A multifaceted treatment approach for young children with conduct problems; Evidence-based psychotherapies for children and adolescents. New York: The Guilford Press, 2003, pp 224-240.

15 Gross D, Garvey C, Julion W, Fogg L, Tucker S, Mokros H: Efficacy of the Chicago Parent Program with low-income African American and Latino parents of young children. Prev Sci 2009;10:54-65.

16 Feinberg ME: Coparenting and the transition to parenthood: A framework for prevention. Clin Child Fam Psychol Rev 2002;5:173-195.

17 Schulz MS, Cowan CP, Cowan PA: Promoting healthy beginnings: A randomized controlled trial of a preventive intervention to preserve marital quality during the transition to parenthood. J Consult Clin Psychol 2006;74:20.

18 Feinberg ME, Kan ML: Establishing family foundations: intervention effects on coparenting, parent/infant well-being, and parent-child relations. J Fam Psychol 2008;22:253.

19 Halford WK, Petch J: Couple psychoeducation for new parents: Observed and potential effects on parenting. Clin Child Fam Psychol Rev 2010;13:164-180.

20 Nelis D, Kotsou I, Quoidbach J, Hansenne M, Weytens F, Dupuis P, Mikolajczak M: Increasing emotional competence improves psychological and physical well-being, social relationships, and employability. Emotion 2011;11:354.

21 Van Den Hoofdakker BJ, Van der Veen-Mulders L, Sytema S, Emmelkamp PM, Minderaa RB, Nauta MH: Effectiveness of behavioral parent training for children with ADHD in routine clinical practice: a randomized controlled study. J Am Acad Child Adolesc Psychiatry 2007;46:1263-1271.

22 Loop L, Mouton B, Stievenart M, Roskam I: One or many? Which and how many parenting variables should be targeted in interventions to reduce children's externalizing behavior? Behav Res Ther 2017;92:11-23.

23 Treacy L, Tripp G, Baird A: Parent stress management training for attention-deficit/hyperactivity disorder. Behav Ther 2005;36:223-233.

24 Lambert MJ, Barley DE: Research summary on the therapeutic relationship and psychotherapy outcome. Psychother-Theor Res 2001;38:357.

25 Rogers CR, Wood JK: Client-centered theory. In: Burton A, editor. Operational theories of personality. Oxford: Brunner/Mazel, 1974

26 Page RC, Weiss JF, Lietaer G: Humanistic group psychotherapy. In: Cain DJ, editor. Humanistic psychotherapies: Handbook of research and practice. Washington: American Psychological Association, 2002, pp 339-368.

27 Bernard H, Burlingame G, Flores P, Greene L, Joyce A, Kobos JC, Leszcz M, MacNair-Semands RR, Piper WE, McEneaney AMS: Clinical practice guidelines for group psychotherapy. Int J Group Psychother 2008;58:455-542.

28 Kane G, Wood V, Barlow J: Parenting programmes: a systematic review and synthesis of qualitative research. Child Care Health Dev 2007;33:784-793.

29 Scholl MB, Ray DC, Brady-Amoon P: Humanistic counseling process, outcomes, and research. The Journal of Humanistic Counseling 2014;53:218-239.

30 Klein MJ, Elliott R: Client accounts of personal change in process–experiential psychotherapy: A methodologically pluralistic approach. Psychother Res 2006;16:91-105.

31 Montgomery P, Grant S, Mayo-Wilson E, Macdonald G, Michie S, Hopewell S, Moher D: Reporting randomised trials of social and psychological interventions: the CONSORT-SPI 2018 Extension. Trials 2018;19:407.

32 Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P: Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Intern Med 2008;148:295-309.

33 Roskam I, Brianda M-E, Mikolajczak M: A step forward in the conceptualization and measurement of parental burnout: The Parental Burnout Assessment (PBA). Front Psychol 2018;9

34 Mikolajczak M, Brianda ME, Avalosse H, Roskam I: Consequences of parental burnout: Its specific effect on child neglect and violence. Child Abuse Negl 2018;80:134-145.

35 Mikolajczak M, Gross JJ, Roskam I: Parental Burnout: What Is It, and Why Does It Matter? Clin Psychol Sci 2019;7:1319-1329.

36 Watson D, Clark LA, Tellegen A: Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54:1063.

37 Maslach C, Jackson SE, Leiter MP, Schaufeli WB, Schwab RL: Maslach burnout inventory. Palo Alto: Consulting Psychologists Press, 1986.

38 Craig KJ, Hietanen H, Markova IS, Berrios GE: The Irritability Questionnaire: a new scale for the measurement of irritability. Psychiatry Res 2008;159:367-375.

39 Russell E, Koren G, Rieder M, Van Uum S: Hair cortisol as a biological marker of chronic stress: current status, future directions and unanswered questions. Psychoneuroendocrinology 2012;37:589-601.

40 Staufenbiel SM, Penninx BW, Spijker AT, Elzinga BM, van Rossum EF: Hair cortisol, stress exposure, and mental health in humans: a systematic review. Psychoneuroendocrinology 2013;38:1220-1235.

41 Stalder T, Steudte-Schmiedgen S, Alexander N, Klucken T, Vater A, Wichmann S, Kirschbaum C, Miller R: Stress-related and basic determinants of hair cortisol in humans: a meta-analysis. Psychoneuroendocrinology 2017;77:261-274.

42 Sauvé B, Koren G, Walsh G, Tokmakejian S, Van Uum SH: Measurement of cortisol in human hair as a biomarker of systemic exposure. Clin Invest Med 2007;30:183-191.

43 Meyer J, Novak M, Hamel A, Rosenberg K: Extraction and analysis of cortisol from human and monkey hair. J Vis Exp 2014;83:e50882.

44 Russell E, Kirschbaum C, Laudenslager ML, Stalder T, De Rijke Y, van Rossum EF, Van Uum S, Koren G: Toward standardization of hair cortisol measurement: results of the first international interlaboratory round robin. Ther Drug Monit 2015;37:71-75.

45 Singer JD, Willett JB, Willett JB: Applied longitudinal data analysis: Modeling change and event occurrence. New York: Oxford university press, 2003.

46 McCoy CE: Understanding the Intention-to-treat Principle in Randomized Controlled Trials. West J Emerg Med 2017;18:1075-1078.

47 Van Uum SHM, Sauvé B, Fraser LA, Morley-Forster P, Paul TL, Koren G: Elevated content of cortisol in hair of patients with severe chronic pain: A novel biomarker for stress. Stress 2008;11:483-488.

48 Martorell GA, Bugental DB: Maternal variations in stress reactivity: Implications for harsh parenting practices with very young children. J Fam Psychol 2006;20:641.

49 Belsky J, van IJzendoorn MH: What works for whom? Genetic moderation of intervention efficacy. Dev Psychopathol 2015;27:1-6.

50 van Ijzendoorn MH, Bakermans-Kranenburg MJ: Genetic differential susceptibility on trial: Meta-analytic support from randomized controlled experiments. Dev Psychopathol 2015;27:151-162.

51 van Ijzendoorn MH, Bakermans-Kranenburg MJ: Differential susceptibility experiments: Going beyond correlational evidence: Comment on beyond mental health, differential susceptibility articles. Dev Psychopathol 2012;48:769-774.

52 Goldberg SB, Manley AR, Smith SS, Greeson JM, Russel E, Van Uum S, et al: Hair Cortisol as a Biomarker of Stress in Mindfulness Training for Smokers. J Altern Complement Med 2014;20:630-634.

53 Greff MJE, Levine JM, Abuzgaia AM, Elzagallaai AA, Rieder MJ, van Uum SHM: Hair cortisol analysis: An update on methodological considerations and clinical applications. Clin Biochem 2019;63:1-9.

54 Barth RP: Preventing child abuse and neglect with parent training: Evidence and opportunities. Future Child 2009;19:95-118.

55 MacMillan HL, Wathen CN, Barlow J, Fergusson DM, Leventhal JM, Taussig HN: Interventions to prevent child maltreatment and associated impairment. Lancet 2009;373:250-266.

56 Lundahl BW, Nimer J, Parsons B: Preventing child abuse: A meta-analysis of parent training programs. Res Social Work Prac 2006;16:251-262.

57 Guidi J, Brakemeier E-L, Bockting CL, Cosci F, Cuijpers P, Jarrett RB, Linden M, Marks I, Peretti CS, Rafanelli C: Methodological recommendations for trials of psychological interventions. Psychother Psychosom 2018:1-9.

58 Albert U, Aguglia A, Bogetto F, Cieri L, Daniele M, Maina G, Necci R, Parena A, Salvati L, Tundo A: Effectiveness of Cognitive-Behavioral Therapy Addition to Pharmacotherapy in Resistant Obsessive-Compulsive Disorder: A Multicenter Study. Psychother Psychosom 2012;81:383-385.

59 Fabiano GA: Father participation in behavioral parent training for ADHD: Review and recommendations for increasing inclusion and engagement. J Fam Psychol 2007;21:683.

60 Lundahl BW, Tollefson D, Risser H, Lovejoy MC: A meta-analysis of father involvement in parent training. Res Social Work Prac 2008;18:97-106.

61 Faircloth C: Intensive parenting and the expansion of parenting. In: Lee E, editor. Parenting culture studies. London: Palgrave Macmillan, 2014, pp 25-50.

62 Hays S: The cultural contradictions of motherhood. New Haven: Yale University Press, 1998.

63 Füredi F: Paranoid parenting: Abandon your anxieties and be a good parent. London: Allen Lane, 2001.

64 Nelson MK: Parenting out of control: Anxious parents in uncertain times. New York: NYU Press, 2010.

65 Wolf JB: Is breast best?: Taking on the breastfeeding experts and the new high stakes of motherhood. New York: NYU Press, 2010.

Table S1. Groups' location, composition, type of intervention, and psychologists' preference for the type of intervention

Group	Location	N			Type of intervention	Pairs	Psychologists' preference[1]
		Females	Males	Tot			
1	Anderlues	6	0	6	Directive	1	Yes
2	Anderlues	7	1	8	Non-Directive		Yes
3	Arlon	6	0	6	Directive	2	No
4	Arlon	11	2	13	Non-Directive		No
5	Bruxelles	11	0	11	Directive	3	No
6	Bruxelles	7	1	8	Non-Directive		Yes
7	Liège	8	1	9	Directive	4	No
8	Liège	10	1	11	Non-Directive		No
9	Namur	11	4	15	Directive	5	Yes
10	Wavre	13	1	14	Non-Directive		Yes
11	Jambes	6	2	8	Directive	6	No
12	Louvain-la-Neuve	10	1	11	Non-Directive		No
13	Verviers	11	0	11	Directive[a]	7	Yes
14	Tournai	9	2	11	Directive[a]	8	Yes
Tot	-	126	16	142	-		

[1]For "Psychologists' preference", yes = psychologists who led the group had a preference for a specific condition and led that group alone; no = psychologists had no preference and led two groups (one in each condition). [a]The original design envisaged 16 groups: two groups could not be included since psychologists had to withdraw from the project (for either personal or professional reasons). Hence, in two cases (groups 13 and 14), we do not have the corresponding paired group.

Table S2. Descriptive statistics of medication intake

	n	%
Corticoids	3	2.2%
Pituitary or steroid hormones	19	13.7%
Antidepressants	24	17.3%
Tranquilizers	11	7.9%
Mood regulators	5	3.6%
Antipsychotics/neuroleptics	1	0.7%
Sleeping pills	4	2.9%

Table S3. Sociodemographics and PB at baseline for the Waiting participants, and for the Directive and Non-Directive groups

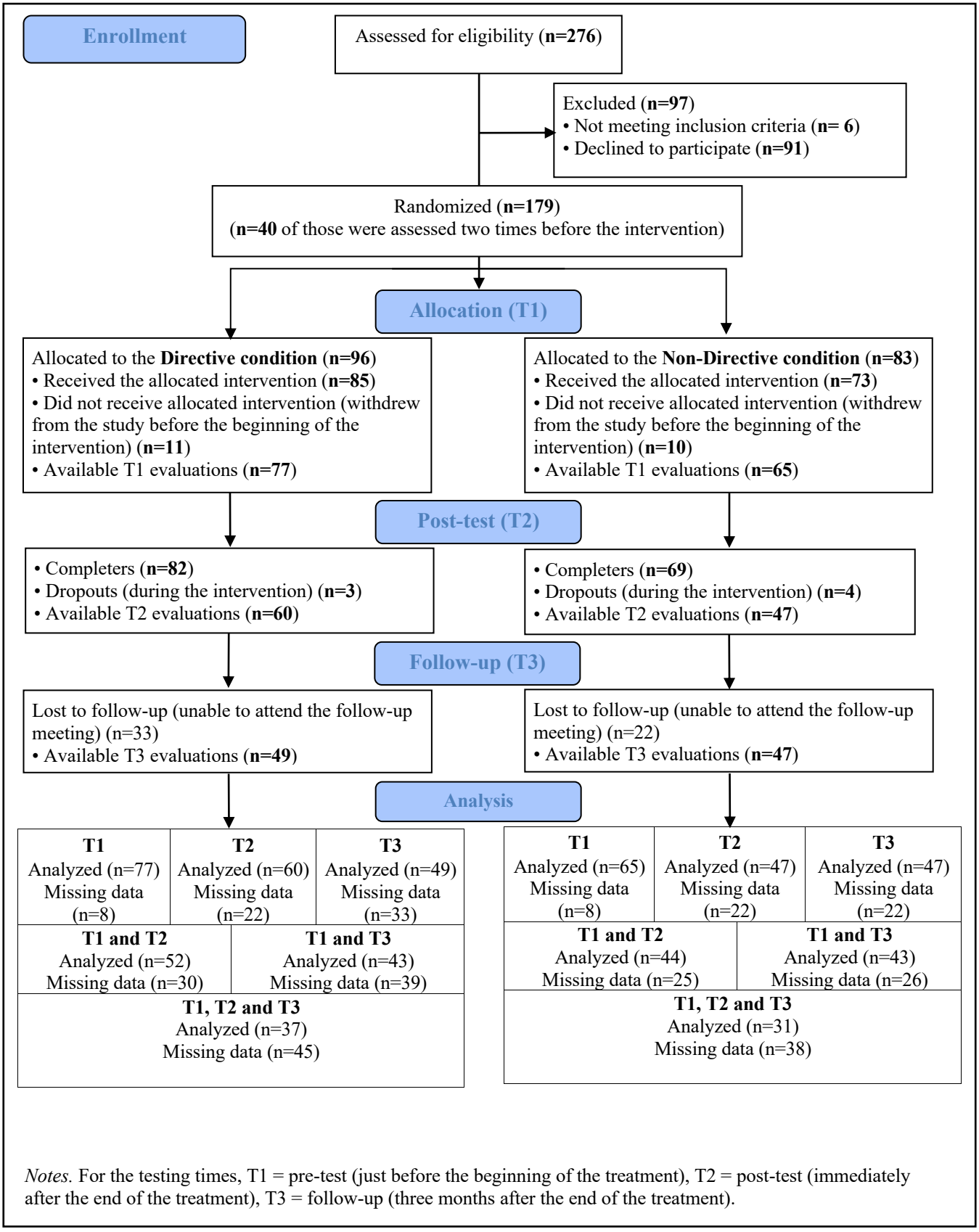
		Waiting participants	Experimental group (Directive+Non-Directive)	*p*	Directive	Non-Directive	*p*
Gender	Female, %	90.0%	88.4%	0.79	88.3%	89.2%	0.86
Age	25-34 years, %	12.5%	23.1%	0.16	23.4%	16.9%	0.13
	35-44 years, %	67.5%	52.1%		57.1%	55.4%	
	45-54 years %	15.0%	23.1%		15.6%	27.7%	
	Over 54 years, %	5.0%	1.7%		3.9%	-	
Number of children	One child, %	17.5%	9.1%	0.25	10.4%	13.8%	0.38
	2 or 3 children, %	72.5%	74.4%		77.9%	67.7%	
	More than 3 children, %	10.0%	16.5%		11.7%	18.5%	
Marital status	In couple, %	92.5%	81.0%	0.09	84.4%	86.2%	0.77
Highest degree obtained	Middle school or high school, %	17.5%	16.5%	0.74	16.9%	15.4%	0.25
	Bachelor's or master's degree, %	67.5%	72.7%		75.3%	67.7%	
	Tertiary education degree, %	15.0%	10.7%		7.8%	16.9%	
Net monthly income	Between €1000 and €2500, %	7.5%	12.4%	0.31	15.9%	20.5%	0.33
	Between €2500 and €4000, %	32.5%	20.7%		45.5%	34.1%	
	Between €4000 and €5500, %	22.5%	16.5%		31.8%	25.0%	
	Over €5500, %	2.5%	9.9%		6.8%	20.5%	
Employment status	Full-time employed, %	25.0%	36.4%	0.28	27.3%	36.9%	0.15
	Part-time employed, %	40.0%	39.7%		40.3%	44.6%	
	Unemployed, %	35.0%	24.0%		32.5%	18.5%	
PB at baseline	Mean PBA (± standard deviation)	85.53 (±30.09)	85.42 (±30.33)	0.98	88.64 (±28.47)	79.48 (±30.89)	0.07

* p values are for χ^2 tests for categorical variables and t-test for parental burnout score.

Table S4. Zero order relations among self-reported, informant-reported, and biological measures at baseline

		Self-reported measures							Informant-reported measures				Biological measure	
		1	2	3	4	5	6	7	8	9	10	11	12	13
Self-reported measures	PB (1)	-	.43**	.48**	-.52**	.50**	-.31**	.29**	.36**	-.10	.21	.22	-0.03	-0.04
	Parental Neglect (2)		-	.51**	-.24**	.03	-.02	.16	.11	.08	.02	.10	-.15	-.19*
	Parental Violence (3)			-	-.24**	.27**	-.15	.02	.01	.01	.14	.26*	.01	.08
	Positive Emotions (4)				-	-.47**	-.33**	-.23**	-.24*	.21	-.30**	-.15	.11	.04
	Negative Emotions (5)					-	-.31**	.20*	.30**	-.13	.39**	.32**	-.03	-.01
	Balance R-R (6)						-	-.12	-.29*	.08	-.30**	-.16	-.02	-.06
	Job Burnout (7)							-	.11	.00	.09	.09	-.22*	-.23*
Informant-reported measures	PB_i (8)								-	-.35**	.57**	.55**	.15	.16
	Positive $Emotions_i$ (9)									-	-.18	-.31**	-.03	-.14
	Negative $Emotions_i$ (10)										-	.71**	.17	.18
	$Irritability_i$ (11)											-	.23	.19
Biological measure	Hair Cortisol (12)												-	.88**
	Hair Cortisol_log (13)													-

Notes. ** = $p < 0.01$; * = $p < 0.05$. Balance R-R = score of the balance between risks and resources questionnaire.



Notes. For the testing times, T1 = pre-test (just before the beginning of the treatment), T2 = post-test (immediately after the end of the treatment), T3 = follow-up (three months after the end of the treatment).

Figure S1. Participants' flow through the three testing times

Table S5. MANOVA univariate effects of Time (baseline vs 8 weeks later) on self-reported and informant-reported outcomes for the Waiting participants

Outcomes		T0 M (SD)	T1 M (SD)	MS	F	*p*
Self-reported outcomes (N=40; N=28 when Job Burnout is considered)	PB	85.53 (30.09)	84.10 (29.70)	40.61	0.29	.60
	Parental Neglect	21.35 (14.53)	19.80 (14.21)	48.05	1.91	.18
	Parental Violence	18.70 (13.07)	16.88 (12.02)	66.61	3.89	.06
	Positive Emotions	10.05 (4.61)	10.72 (4.99)	9.11	2.02	.16
	Negative Emotions	19.75 (5.35)	18.55 (5.98)	28.80	2.95	.09
	Balance R-R	0.50 (52.20)	4.25 (54.64)	281.25	0.32	.58
	Job Burnout	34.75 (18.63)	35.75 (18.48)	14.00	0.23	.63
Informant-reported outcomes (N=12)	PB_i	62.75 (27.77)	61.58 (26.11)	8.17	0.08	.79
	Positive $Emotions_i$	7.58 (2.23)	6.50 (2.65)	7.04	2.82	.12
	Negative $Emotions_i$	13.75 (2.22)	12.33 (4.19)	12.04	1.96	.19
	$Irritability_i$	26.50 (13.69)	27.08 (15.92)	2.04	0.04	.84

Notes. For self-reported outcomes, df1 = 1 and df2 = 38 (df1 = 1 and df2 = 26 when considering job burnout); for informant-reported outcomes, df1 = 1 and df2 = 11. Balance R-R = score of the balance between risks and resources questionnaire.

Table S6. Means and SDs of the first and second evaluations for the Waiting List (pre-pre-test vs. pre-test), and the experimental group (pre-test vs. post-test), within-subject and between-subject effect sizes

Outcomes	Waiting List		Experimental group (Directive+Non-Directive)		Waiting List vs. Experimental group
Self-reported outcomes	n	M (SD)	n	M (SD)	Effect sizes Cohen's d^2 (95% CI)
PB					
First evaluation	40	85.53 (30.09)	70	87.92 (30.02)	
Second evaluation	40	84.10 (29.70)	70	63.54 (10.96)	
Cohen's d^1 (95% CI)		0.05 (-0.13 – 0.22)		0.75 (0.50 – 1.00)	
Mean difference and Cohen's d^2 (95% CI)		1.43 (16.80)		24.38 (29.71)	0.88 (0.48 – 1.29)
Parental Neglect					
First evaluation	40	21.35 (14.53)	69	19.19 (13.94)	
Second evaluation	40	19.80 (14.21)	69	15.03 (10.96)	
Cohen's d^1 (95% CI)		0.11 (-0.05 – 0.26)		0.33 (0.12 – 0.54)	
Mean difference and Cohen's d^2 (95% CI)		1.55 (7.10)		4.16 (10.71)	0.27 (-0.12 – 0.66)
Parental Violence					
First evaluation	41	18.70 (13.07)	69	19.70 (13.94)	
Second evaluation	41	16.88 (12.02)	69	14.47 (9.98)	
Cohen's d^1 (95% CI)		0.14 (-0.01 – 0.29)		0.48 (0.26 – 0.70)	
Mean difference and Cohen's d^2 (95% CI)		1.83 (5.85)		5.23 (9.45)	0.41 (0.01 – 0.80)
Positive Emotions					
First evaluation	40	9.93 (4.63)	69	8.64 (3.89)	
Second evaluation	40	11.05 (5.65)	69	10.96 (4.86)	
Cohen's d^1 (95% CI)		-0.14 (-0.34 – 0.06)		-0.23 (-0.78 – -0.27)	
Mean difference and Cohen's d^2 (95% CI)		-0.68 (3.01)		-2.29 (4.53)	-0.42 (-0.82 – -0.03)
Negative Emotions					
First evaluation	40	19.75 (5.35)	69	20.41 (5.20)	
Second evaluation	40	18.55 (5.98)	69	14.80 (6.12)	
Cohen's d^1 (95% CI)		0.21 (-0.04 – 0.46)		0.99 (0.68 – 1.29)	
Mean difference and Cohen's d^2 (95% CI)		1.20 (4.42)		5.58 (6.36)	0.77 (0.37 – 1.17)
Balance R-R					
First evaluation	40	0.50 (51.20)	71	-7.42 (51.61)	
Second evaluation	40	4.25 (54.64)	71	-0.72 (54.19)	
Cohen's d^1 (95% CI)		-0.07 (-0.31 – 0.18)		-0.13 (-0.38 – 0.12)	
Mean difference and Cohen's d^2 (95% CI)		-3.75 (42.14)		-6.77 (57.44)	-0.06 (-0.44 – 0.33)
Job Burnout					
First evaluation	28	34.75 (18.63)	56	36.44 (14.95)	
Second evaluation	28	35.75 (18.48)	56	35.95 (18.53)	
Cohen's d^1 (95% CI)		-0.05 (-0.27 – 0.17)		0.03 (-0.18 – 0.24)	
Mean difference and Cohen's d^2 (95% CI)		-1.00 (10.97)		0.49 (13.79)	0.11 (-0.34 – 0.57)
Informant-reported outcomes					
PB_i					
First evaluation	11	62.75 (27.77)	38	74.42 (27.62)	

Second evaluation	11	61.58 (26.11)	38	54.05 (33.04)	
Cohen's d^1 (95% CI)		0.04 (-0.26 – 0.35)		0.66 (0.31 – 1.02)	
Mean difference and Cohen's d^2 (95% CI)		1.17 (14.52)		18.37 (30.40)	0.19 (-0.44 – 0.83)
Positive Emotions$_i$					
First evaluation	11	7.58 (2.23)	39	6.41 (2.65)	
Second evaluation	11	6.50 (2.65)	39	8.41 (3.35)	
Cohen's d^1 (95% CI)		0.41 (-0.11 – 0.98)		-0.65 (-1.00 – -0.31)	
Mean difference and Cohen's d^2 (95% CI)		1.08 (2.23)		-2.00 (2.99)	-0.83 (-1.49 – -0.18)
Negative Emotions$_i$					
First evaluation	11	13.75 (2.22)	39	14.95 (4.10)	
Second evaluation	11	12.33 (4.19)	39	11.59 (4.23)	
Cohen's d^1 (95% CI)		0.36 (-0.20 – 1.03)		0.79 (0.47 – 1.13)	
Mean difference and Cohen's d^2 (95% CI)		1.42 (3.50)		3.36 (3.63)	0.27 (-0.36 – 0.90)
Irritability$_i$					
First evaluation	11	26.50 (13.69)	37	36.38 (16.01)	
Second evaluation	11	27.08 (15.92)	37	32.81 (17.86)	
Cohen's d^1 (95% CI)		-0.04 (-0.41 – 0.33)		0.21 (-0.10 – 0.52)	
Mean difference and Cohen's d^2 (95% CI)		-0.58 (9.65)		3.57 (15.94)	-0.11 (-0.75 – 0.52)

[1]Within-subject effect sizes, first evaluation vs. second evaluation in Waiting List and Experimental groups.
[2]Between-subject effect sizes, mean difference between the first and the second evaluation in the Waiting List vs. mean difference between the first and the second evaluation in the Experimental group.
Notes. Balance R-R = score of the balance between risks and resources questionnaire.

Table S7. Results of intention-to-treat analyses using multilevel modeling on self-reported and informant-reported outcomes

	Fixed effects			
	Coefficient	SE	t	
Self-reported outcomes				
PB				
Intercept (β_{00})	107.24***	4.79	(107) 22.41	
Time effect (β_{10})	-13.98***	1.58	(194) -8.84	
Time X Group effect (β_{11})	-1.06	1.75	(194) -0.61	
Deviance				2866.84
Parental Neglect				
Intercept (β_{00})	23.44***	2.09	(107) 11.21	
Time effect (β_{10})	-2.61***	0.59	(192) -4.45	
Time X Group effect (β_{11})	0.31	0.56	(192) 0.56	
Deviance				2238.21
Parental Violence				
Intercept (β_{00})	21.55***	1.60	(107) 13.43	
Time effect (β_{10})	-2.34***	0.40	(192) -5.85	
Time X Group effect (β_{11})	0.08	0.42	(192) 0.20	
Deviance				2140.81
Positive Emotions				
Intercept (β_{00})	7.58***	0.70	(107) 10.80	
Time effect (β_{10})	1.31***	0.22	(187) 6.08	
Time X Group effect (β_{11})	0.02	0.25	(187) 0.08	
Deviance				1680.09
Negative Emotions				
Intercept (β_{00})	23.73***	0.86	(107) 27.53	
Time effect (β_{10})	-2.69***	0.28	(187) -9.60	
Time X Group effect (β_{11})	-0.23	0.31	(187) -0.73	
Deviance				1821.99
Balance R-R				
Intercept (β_{00})	-11.16	7.49	(107) -1.49	
Time effect (β_{10})	6.53**	2.42	(196) 2.69	
Time X Group effect (β_{11})	0.10	2.83	(196) 0.03	
Deviance				3198.52
Job Burnout				
Intercept (β_{00})	40.37***	2.40	(97) 16.81	
Time effect (β_{10})	-1.16	0.69	(145) -1.70	
Time X Group effect (β_{11})	0.41	0.99	(145) 0.41	
Deviance				1958.21
Informant-reported outcomes				
PB_i				
Intercept (β_{00})	86.25***	11.21	(32) 7.69	
Time effect (β_{10})	-8.89*	4.28	(32) -2.08	

Time X Group effect (β_{11})	0.05	3.86	(32) 0.01	
Deviance				621.39
Positive Emotions$_i$				
Intercept (β_{00})	4.85***	1.02	(32) 4.76	
Time effect (β_{10})	0.98*	0.47	(33) 2.08	
Time X Group effect (β_{11})	0.54	0.41	(33) 1.33	
Deviance				329.12
Negative Emotions$_i$				
Intercept (β_{00})	19.59***	1.42	(32) 13.76	
Time effect (β_{10})	-2.44***	0.53	(33) -4.57	
Time X Group effect (β_{11})	-0.28	0.54	(33) -0.52	
Deviance				369.98
Irritability$_i$				
Intercept (β_{00})	34.77***	6.17	(32) 5.63	
Time effect (β_{10})	0.74	2.41	(31) 0.31	
Time X Group effect (β_{11})	0.72	2.21	(31) 0.33	
Deviance				537.50

Notes. *** = p <.001. ** = p <.01. * = p < .05. Time effect for self-reported outcomes has been analyzed across three testing times: before the intervention (T1), after the intervention (T2), and at three months follow-up (T3). For informant-reported outcomes, time effect has been analyzed exclusively between T1 and T2, because of the excessive missing data at T3 and the consequent lack of statistical power.

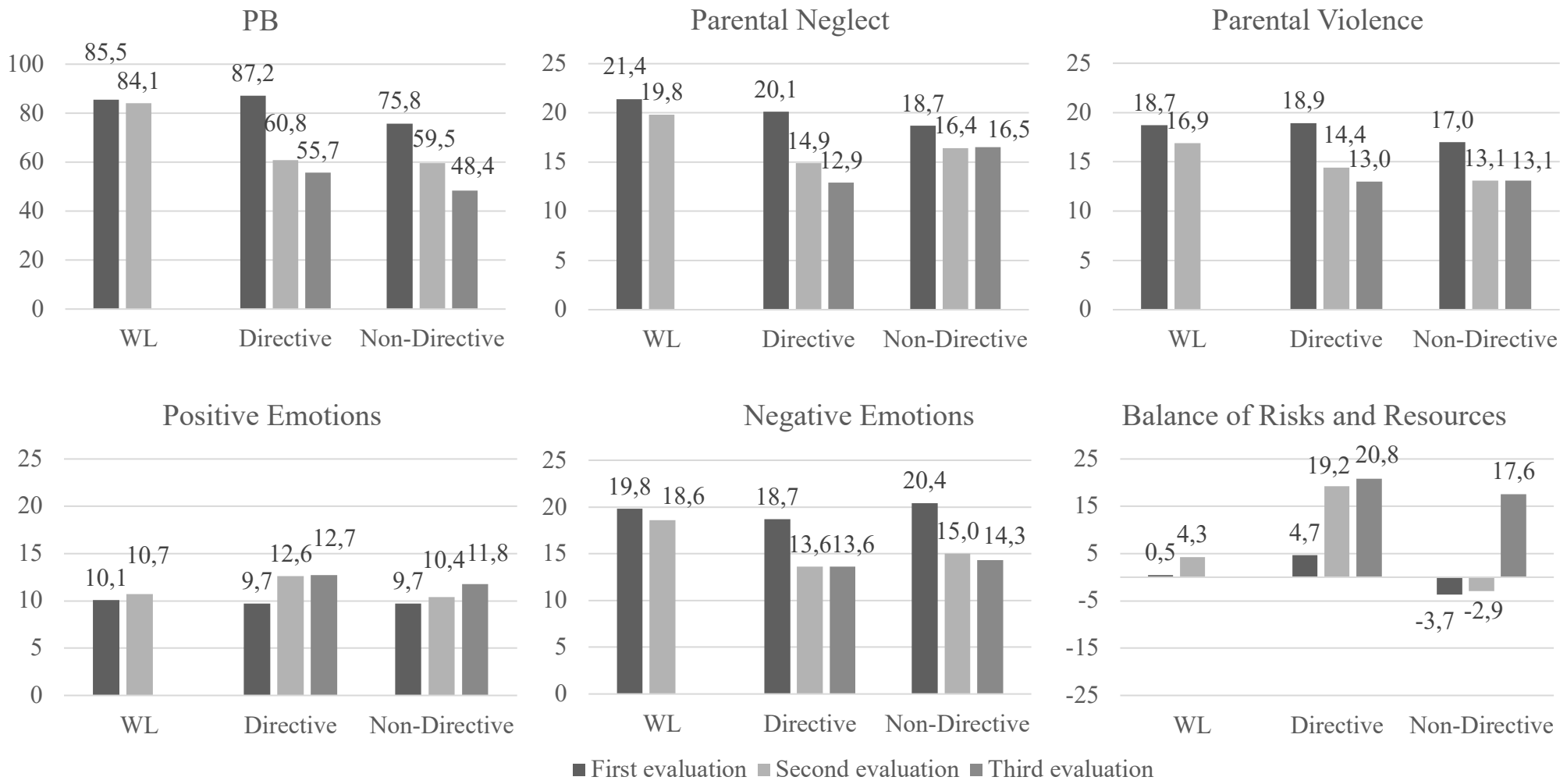


Figure S2. Evolution of PB, parental neglect and violence, positive and negative emotions and the balance of risks and resources through the three testing time points in the Waiting List (WL) and the two experimental groups