The temporal dynamics of parental burnout:

Extending the network approach to the family system

The temporal dynamics of parental burnout:

Extending the network approach to the family system

M. Annelise BLANCHARD

Promoteur

Alexandre Heeren (UCLouvain/FNRS)

Thèse présentée en vue de l'obtention du grade de Docteur en sciences psychologiques et de l'éducation

Supervisory Panel

Isabelle ROSKAM (UCLouvain)
Richard J. McNALLY (Harvard University)

Collection de thèses de l'Université catholique de Louvain, 2024

Supervisory panel and jury

Isabelle ROSKAM (UCLouvain, president of the jury)
Alexandre HEEREN (UCLouvain/FNRS, secretary of the jury)

Baptiste BARBOT (UCLouvain)
Eva CEULEMANS (KULeuven)
Dominique MACIEJEWSKI (Tilburg University)
Richard J. McNALLY (Harvard University)

Cover image: John Paul Blanchard
Page formatting and layout: Yorgo Hoebeke

Abstract

Parenting can be joyful, but also difficult. Some parents, faced with chronic unresolvable stress and insufficient resources, burn out: unlike their previous excitement and passion for parenting, they now feel emotionally exhausted, emotionally distant from their children, and fed up with parenting. Parenting entails interactions within a family context—and so to understand parental burnout requires situating it in its context of the family. This thesis proposes using a network approach to examine parental burnout and the family system: exploring how parenting experiences (reciprocally) relate to interactions with children, the co-parent, and the wider environment. Moreover, parental burnout has only been investigated as a unitary construct, but this thesis examines the distinct roles of each of its components. Using both cross-sectional and temporal network analyses, we found that emotional exhaustion likely contributes to the development of parental burnout, and emotional distance to its maintenance. I discuss these findings as grounded in specific cultural parenting norms and call for further development of network models to adequately assess family systems.

Résumé

Être parent peut être un plaisir, mais aussi une source de difficultés. Certains parents, confrontés à un stress chronique persistant et à des ressources insuffisantes, tombent en burn-out : contrairement à la passion qu'ils éprouvaient, ils se sentent désormais émotionnellement épuisés, distants de leurs enfants, et frustrés dans leur rôle de parents. Par définition, la parentalité implique des interactions avec un contexte familial—alors pour comprendre le burnout parental, il y est nécessaire de l’étudier dans son contexte. Cette thèse propose d'utiliser une approche en réseau pour examiner le burnout parental et le système familial : explorer comment les expériences parentales sont (réciproquement) liées aux interactions avec les enfants, l'éventuel co-parent et l'environnement familial. En outre, le burnout parental n'a été étudié que comme un concept unitaire, mais cette thèse examine les rôles distincts de chaque composante du burnout parental. Sur base de nos analyses de réseau transversales et temporelles, il semblerait que l'épuisement émotionnel contribue au développement du burnout parental, et la distance émotionnelle à son maintien. Je propose que ces résultats reflètent des normes parentales culturelles spécifiques, et je discute de l’urgence d’un développement plus approfondi des modèles en réseaux, afin d'évaluer de manière plus appropriée la nature complexe et dynamique des systèmes familiaux.

Table of Contents

Acknowledgements

This dissertation came to exist thanks to a broad constellation of support and care. First, Alexandre: thank you for your ever-present support and guidance, for teaching me how to write while keeping the audience and structure in mind, and for always both supporting my ideas but also looking for flaws and discussing alternatives. Through working with you, I truly learned how to be a researcher. To my committee members, Profs. Isabelle Roskam and Richard McNally: thank you both for your advice and support throughout this project. Through your commentary, I learned how to center the parent in my research. To my jury members, Profs. Baptiste Barbot, Dominique Maciejewski, and Eva Ceulemans: thank you all for your careful review and comments regarding my dissertation, I sincerely appreciate your time and effort.

This research project only took place due to the help from many collaborators. First, to Jordan Revol and Rana Begum Kalkan, thank you for helping me prepare the studies and their materials; and to Eline Kerkhofs, Léa de Clippele, and Charlotte Monin, thank you for your help in collecting data. Your enthusiasm and curiosity about the research project inspired me even more, for both the project itself and for collaborating with others. I also received advice and support from the Parental Burnout ARC: Profs. Isabelle Roskam, Moïra Mikolajczak, Aurore Francois, Isabelle Aujoulat, as well as Margaux, Anne-Catherine, and Gao-Xian: thank you for your advice and questions throughout many meetings. I appreciated working with researchers from many different domains on a same topic, as it enriched my understanding of parental burnout and how it is embedded in place and time, while also helping me to see my own biases as a psychology researcher. I was also fortunate enough to receive multiple datasets from other researchers to reanalyze with other methods: thank you to Profs. Isabelle Roskam, Moïra Mikolajczak, Margaret Kerr, Rachel Barr, and Heather Kirkorian, who were all kind enough not only to share their data, but also to provide insightful feedback and suggestions on the shared projects.

And most crucially, thank you to all the parents who agreed to participate in these studies and share their daily experiences of parenting. You helped advance the understanding of parental burnout a little further, and I appreciate the time, care, and consideration each and every one of you took in participating.

Within the faculty, I had substantial support from those around me. To all who have been part of the lab, Charlotte, Yorgo, Camille, Mariah, Doris, Alba, and Clara: thank you for sharing lunches, stories, ideas, projects, and kind but constructive feedback. My research and my time at the faculty is better for it. Thank you also to so many others around the faculty, for the many conversations during breaks and over food about research and life and more. Un grand merci également à Nadine Didier, Isabelle Legrain, et Dominique Hougardy, pour leur soutien régulier tout au long de ces années.

Many of you also shaped how I view research. Yorgo: I appreciate our many conversations of promises and flaws of various research ideas (along with all the entertaining digressions), as well as your willingness to always share code and ideas. Alba: You taught me to bring a clinical perspective to my research, particularly when working with parents in burnout. I treasure our many conversations on research and life, the many plates of good food we've shared, and your confidence in my research skills. Elise: Your presence on Teams has accompanied almost every workday of this last year or so, and you brought me such positive energy and motivation for the last push. Sam: I truly learned so much about research with you, from discussing research methods, brainstorming about journals, and nitpicking each other's studies. I am so happy that we became friends and shared so many discussions on academia and research ethics over many, many cups of coffee. Marbella: I am so lucky to have shared an office with you these past four and more years. From stretching breaks, to asking each other writing questions while eating snacks, to learning about mindfulness both professionally and personally, you have accompanied my PhD journey from its very start with such joyful presence. I strive to emulate how you incorporate compassion and attentive listening into your research and personal life.

And to the great number of people who kept me going: I appreciate your support so much. Thank you to all the friends and family—in Belgium, the U.S., and elsewhere—who have heard snippets of my research at various times and showered me with such sincere support and encouragement.

And lastly, to my family: thank you. Mom, you taught me to love reading and value kindness, and both still permeate and improve my daily life. I remember telling you I would never go into any field related to either science or teaching (whoops), but you have embodied such a good role model for both. Pop, you taught me so much of what I know about writing, from editing my high school essays together and asking 'why' about every change. Thank you for your clinical insights, your help with translations, and all the delicious food you've made for me these past years. Also—what a cover, thank you! Brigitte, sharing about my daily work tasks and hearing how different yours were has always been a delightful part of my week. Your questions asking me to summarize my research in one sentence genuinely helped me learn how to quickly give a snapshot of my work (even if I sometimes still find it a struggle…). Teele, dearest, you have brought joy and lightness into my days since I met you. Thank you for the support and love these past few years, and for all the snacks to munch on while working, coffee to fuel my days, and time to unwind together. You have all believed in me so sincerely and offered such caring support : thank you.

Section 1.

General Introduction

Chapter 1.

Theoretical background

Everyone acknowledges parenting as difficult, while asserting that the joys of parenting are worth all the associated troubles. For some parents, however, parenting it not worth it: these parents feel exhausted and worn down, and parenthood becomes a source of suffering instead of joy. In other words, they burn out as parents. How might this happen?

Research into parental burnout directly stemmed from research into job burnout. In 1975, Freudenberger described in a near-autoethnographic account his own experience and that of his colleagues working at free clinics via the term 'burnout,' which he characterizes as a syndrome of emotional depletion, loss of commitment, and a cynical attitude toward work that once held passion and zeal. The phrase 'burn out' was also previously used informally in workplace settings (Heinemann & Heinemann, 2017), conveying intense imagery: someone who burns brightly at the beginning but gives too much of themselves in the face of overwhelming stress, until all that remains is a charred husk. The first qualitative study examining individuals' experiences of burnout (Maslach, 1976) soon followed, and shortly afterwards a coherent conceptualization of burnout as a multidimensional syndrome emerged, with accompanying quantitative measurement tools (Maslach & Jackson, 1981). This view of burnout includes three dimensions: **emotional exhaustion** (i.e., feeling "emotionally extended and exhausted by one's work"), **depersonalization** (i.e., "unfeeling and impersonal response towards recipients of one's care or service"), and **low personal accomplishment** (i.e., feeling incompetent and low achievement in one's work; Maslach & Jackson, 1981, p. 101).

Although some researchers argue that burnout can only apply to work contexts (Maslach et al., 2001; Schaufeli & Taris, 2005), others assert that burnout can arise in any sphere of life that offers people a sense of meaning (Pines & Aronson, 1988). Indeed, Bianchi and colleagues (2014) deftly argue that each dimension of burnout proposed by Maslach and Jackson (1981) can easily apply to other contexts. First, emotional exhaustion can occur wherever there is unresolvable chronic stress, which is the fundamental cause of burnout (Maslach et al., 2001). Second, cynicism—or in other words, detachment and withdrawal—can occur wherever an individual emotionally invests in a specific domain (be it work or elsewhere) only to eventually find this investment thwarted and frustrated: they will then defensively disinvest. In Maslach and Jackson's conception, this domain takes the form of client-facing work context, and cynicism results in depersonalization toward clients. Third, whenever someone undertakes an action, they can feel incompetent and ineffective if they do not accomplish it to their standards.

Burnout particularly fits the context of parenting, as parenthood is a sphere of life often imbued with overwhelming meaning. Since early on in academic conversations about burnout, researchers discussed burnout not only in the context of work, but also in the context of parenting (Pines & Aronson, 1988; Procaccini & Kiefaber, 1984). Indeed, Pelsma and colleagues (1989) directly administered the Maslach Burnout Inventory to a sample of stay-at-home mothers, and found that both the dimensions of emotional exhaustion and of low personal accomplishment suited the parenting context—only the dimension of depersonalization did not. Further research focused directly on parents facing substantial stress in the realm of parenting: parents of children with illnesses or other special needs (e.g., Lindahl Norberg, 2007; M. J. Weiss, 2002). These studies consistently found evidence of burnout in parents, using either a burnout questionnaire developed to be occupation-neutral (e.g., the Shirom-Melamed Burnout Questionnaire focusing on physical fatigue, emotional exhaustion & cognitive weariness; Lundgren-Nilsson et al., 2012; Melamed et al., 1999) or adapting a job burnout questionnaire to focus specifically on parental burnout (e.g., M. J. Weiss, 2002).

More recently, researchers have also investigated burnout applying to any parent, since parenting itself can bring overwhelming and unresolvable stress even when one's children do not have special needs. Part of this chronic stress comes with the intensifying pressures parents increasingly face when raising their children. Parenting involves not just fulfilling the basic needs of children but also socializing children to acceptable standards within their culture (Dallos & Draper, 2015). As Masten and Palmer (2019, p. 160) put it: "stakeholders in childrearing, including societies, cultures, communities, children, and parents themselves, expect parents to execute a multifaceted set of roles with little or no training for a long period." Adding on to the many responsibilities and little preparation involved in parenting comes incredible pressure. This is exemplified by the very same authors later stating:

> Parents will play a central role in preparing families for […] challenges and helping their children to navigate through and recover from the inevitable adversities of life. They need knowledge and support to enhance their own resilience as parents and foster resilience in their families **because the future of children and societies depends on their success** (Masten & Palmer, 2019, p. 179, emphasis mine).

With this view of their children's future livelihood resting heavily on parents' shoulders, is it any wonder that parenting can engender such inescapable stress for so many? This pressure is exacerbated by the "rapid sociocultural change [in many cultures] from their [own] cultural traditions toward more individualistic and 'modern' family configurations" (Super & Harkness, 2020, p. 189), which can also translate into less family, community and institutional support for parents. It therefore seems likely that any parent, investing emotionally in parenthood, could exhaust themselves in the face of this unresolvable chronic stress resulting from overwhelming pressure, endless duties, and little support.

Recently, the question of parental burnout occurring amidst the general population—and not only among parents of children with special needs—has gained traction. This began with the direct adaptation and empirical investigation of the Maslach Burnout Inventory for parental burnout (i.e., the Parental Burnout Inventory; Roskam et al., 2017), which involved replacing the term "work" in some items with "parental role," as well as adapting the dimension of "depersonalization" into "emotional distance" (which is how focus groups described the core characteristics of depersonalization in a parental context[1]). This study not only found good psychometric properties of the adapted questionnaire, but also low correlations with professional burnout, depression, and parenting stress, suggesting that parental burnout is a distinct syndrome. Moreover, out of the 1,723 parents from a community sample who completed the Parental Burnout Inventory, around 2-12% (depending on conservativeness of the cut-off point) scored high enough that the authors consider them "burned out," emphasizing that parental burnout can occur in the general population.

The same researchers further refined quantitative assessment of parental burnout by designing a questionnaire via inductive methods using the testimonies of burned out parents (the Parental Burnout Assessment; Roskam et al., 2018). This procedure led to four main dimensions: emotional exhaustion and emotional distance remain, with the addition of a feeling of being fed up with parenting and a sense of contrast with the previous parental self. This questionnaire thus deviates from the Parental Burnout Inventory by removing the more functional aspect of "low accomplishment" and replacing it with a further emphasis on affective experience with "feeling fed up." In addition, the dimension of contrast adds further emotional nuance (feelings of guilt and shame arising from the discrepancy between the previous parental self and the current one), while also emphasizing that parental burnout does not apply to parents who were always neglectful and distant but specifically to parents who at one point held passion and joy for parenting that they have since lost. This notion of contrast, although central to the concept of burnout, is not formally measured in the Maslach Burnout Inventory nor in the Parental Burnout Inventory (Roskam et al., 2018). The four-dimensional conceptualization of the Parental Burnout Assessment also had good psychometric properties, with similar results to (and close correlations with) the Parental Burnout Inventory. It has become the standard questionnaire in parental burnout research, not only because it stems directly from parents' experiences, but particularly because it can be used at no cost to researchers and participants (unlike questionnaires adapted from the Maslach Burnout Inventory).

The psychometric properties of the Parental Burnout Assessment have been validated in many countries and samples, including Poland (Szczygieł et al., 2020), Japan (Furutani et al., 2020), Brazil (Matias et al., 2020), Lebanon (Gannagé et al., 2020), Turkey (Arikan et al., 2020), China (Cheng et al., 2020), Finland (Aunola et al., 2020), Iran (Mousavi et al., 2020), and Romania (Stănculescu et al., 2020). Analyses in all of these studies support a four-factor structure of parental burnout across languages and cultures. In addition, a large consortium of

[1] While one can "dehumanize" clients in a work context and consider them as objects, this is less suitable to describe the emotional reaction toward one's own children when severely emotionally depleted (Roskam et al., 2017). The term "emotional distance" captures the same essence of emotional disengagement but applied to the parental context.

researchers in 42 countries have investigated the prevalence of parental burnout, finding that parental burnout occurs in many countries (Roskam et al., 2021). Prevalence rates, however, vary widely around the world: from near 8% in Poland, Belgium, Egypt and Canada to near 1% in Thailand, China, Turkey and Vietnam, with many countries somewhere in the middle. Culture and norms around parenting and family thus likely play a role in parental burnout, perhaps in what constitutes parental stress or suffering (Bornstein, 2012).

Regardless of the exact form of stress, researchers agree that burnout arises from unresolvable chronic stress, which alters physiological states and leaves biological markers (Melamed et al., 1999). One such biomarker is cortisol, particularly hair cortisol concentrations as a biological correlate for exposure to chronic stress: hair grows at a rate of 1 cm per month, and so examining cortisol concentration in several centimeters of hair near the scalp provides a robust retrospective reflection of cortisol secretion over the past several months (Stalder & Kirschbaum, 2012). High levels of cortisol concentration in hair have previously been associated with occupational burnout (Penz et al., 2018). Recently, high hair cortisol concentration has also been found for parents with parental burnout, even when controlling for occupational burnout (Brianda, Roskam, & Mikolajczak, 2020). Therefore, this strongly emphasizes that parental burnout is a chronic stress syndrome specific to the parenting domain.

Given that parental burnout is distinct from occupational burnout, the specific context of parental burnout—the family environment of the parent—must play a fundamental role. Indeed, as Bianchi outlines, burnout can occur where there is a significant emotional investment (later thwarted) in a specific sphere of life, accompanied by unresolvable stress. For parental burnout, both the emotional investment and the unresolvable chronic stress must stem from the parent's engagement and interactions with their children and family. Research into parental burnout therefore must also examine the family context, since parenting does not occur in isolation but by definition involves interactions with children. As Kerig (2019, p. 3) puts it: "Given that the term 'family' inherently refers to a constellation including parents and children, it might be thought that all parenting theory and research concerns the family system." This is because parenting processes do not only impact the children unidirectionally, but also impact the parent, via reciprocal processes and interactions with the children and environment.

To investigate such reciprocal processes between parent and child, and to firmly anchor investigations of parental burnout in the family context, I turn to a family systems approach. Investigating the family as a system has a long history, beginning in the 1950s when researchers and clinicians applied dynamic systems theory—a framework to investigate phenomena, from physics to biology, as a system with reciprocal interactions (von Bertalanffy, 1950)—to the family sphere (Dallos & Draper, 2015; Granic & Hollenstein, 2015). A key aspect of a systemic approach involves a close focus on contexts; when applying this to families, it means that to understand an individual requires understanding them as part of a family system that impacts how they feel, think, and behave (Dallos & Draper, 2015). A system is interactive, and as such does not involve linear sequences with one component directly impacting another, but instead cycles and spirals of feedback between many components. In other words, causality can be understood as circular (Dallos & Draper, 2015). This can be difficult to integrate within psychology (or indeed any scientific) research, which so often seeks to simplify in order to understand (Barabási, 2012): linear processes are easier to investigate than circular and multidimensional ones. However, when simplifying too much, important nuances and context can be erased; to ignore the family context when seeking to understand parenting experiences discards fundamental information.

Since a systems-based approach involves cycles and feedback spirals—because a system implies multiple components that influence one another in circular patterns—it requires dynamic investigation that examines how a system changes *over time*. When applying this to the parenting context, consider how no parent always acts the same way, since the situations and contexts they find themselves in with their children change. In other words, every daily experience as a parent occurs via interaction with the other components of the system, including their children, potential co-parent, and wider environment. Consider how these different scenarios might impact your parenting experience on a particular day: Is your child sick today, or anxious because of a school test, or excited to go to a birthday party? Is your co-parent shouldering most of the household tasks that day, or out of the country for a week? Are grandparents or neighbors able to care for your children that evening, are there extra expenses this week that make it difficult to pay for all necessities for your children, or are there physical dangers in the environment? All of these present different situations that will impact how a parent experiences their day and interacts with their children, and these situations can change from day to day. As Masten and Palmer say (2019, p. 157): family processes "will vary as a function of changing family membership and relationships, the development and health of individual members, the resources and experiences of the family, the supports the family can garner from the environment, and other circumstances that affect the whole family, such as economic capital."

This fits with the concept that although parenting styles are often viewed as stable, they can shift and adapt to the context and situation (Masten & Palmer, 2019). In other words, parenting experiences and actions builds up over time based on individual interactions with children—and no parent is constant in how they act, since situations and contexts are not constant—to form patterns, consolidating into a parenting style. These patterns built over time can be thought of as distributions with an average (i.e., common experiences and actions) as well as variability (i.e., how varied these experiences and actions are over time; Fleeson & Jayawickreme, 2015). Or from another perspective, experiences and actions (occurring via interactions with other system components) eventually consolidate into emergent patterns that form grooves more easily followed in the future (Granic & Hollenstein, 2015; Kerig, 2019). Investigating parental burnout from a family systems perspective thus requires two things: **1) Investigating experiences of parental burnout along with the family environment** (i.e., interactions with children, potential co-parent, and wider environment), and **2) doing so over time**.

However, there are many challenges to studying the family as a system. Indeed, although it has long been accepted that families can usefully be understood as a system with reciprocal processes, this is rarely how family and parenting processes are investigated through research (Dallos & Draper, 2015). This is due to a number of factors. First, although family systems theory grew out of a research framework from systemic approaches, it has grown more clinical in nature since the 1970s and is currently implemented mainly via postmodern therapies. These reject notions of an objective or expert therapist, and emphasize that meaning between family members is constructed through their interactions and always in flux; these perspectives do not always mesh well with traditional research strategies prioritizing quantitative measures often held to be "objective" (Lebow, 2020). Following this more clinical bent to family systems theories, most recent developments have focused on "strengthening the evidence base supporting family treatments" (Kerig, 2019). This is important and relevant work, but there should always be a bidirectional translation process flowing between clinical applications and research: research into family systems theory, including optimal ways to investigate and build its theory, should keep pace with clinical applications.

A second key challenge involves limitations in methods: "Research methodology is somewhat stretched in its ability to examine complex reciprocal patterns of interaction" (Lebow, 2020, p. 4). Fortuitously, many advances to methodology have come with accelerating computing capabilities. As Kerig puts it: "Family systems theory also promises to enjoy a new resurgence, with the emergence of sophisticated dynamic systems observational, methodological, and statistical procedures that allow for modelling the complex processes as transactions and reciprocal effects among family members as they interact and influence one another over time" (2019, p. 26). Two such sophisticated procedures recently developed permit the work in this dissertation.

First, a network approach has emerged from physics research as a method to examine many interactions between multiple components using data-based mathematical models (Barabási, 2016). This has evolved to many fields, including biology and medicine (Barabási et al., 2011), economics and social interactions (Barabási, 2012), and—central to this dissertation—psychology (Borsboom, 2017; Borsboom & Cramer, 2013; Bringmann et al., 2013). The core of this approach is to allow complexity in research, investigating many components and their complicated interactions, instead of continuously attempting to reduce research to the simplest component and investigate only that simplified part. A network approach to psychology, meanwhile, specifically focuses on how symptoms—or patterns of behavior and thoughts (Cramer et al., 2012), emotions (Lange et al., 2020) or other psychological phenomena (Schmittmann et al., 2013)—interact, and through their interactions create emergent patterns that form psychological constructs. To apply this approach to parental burnout, consider how parenting experiences and interactions with the family environment interrelate over time and, when experienced as stressful, lead to the parent feeling exhausted and overwhelmed.

The second methodological advance permitting this dissertating concerns experience sampling methodology, or ESM. Since reciprocal processes and interactions characterize the family system, any investigation into parenting experiences requires longitudinal data. This longitudinal data should reflect the frequency of parenting experiences: it cannot be measured over months but should reflect day-to-day changes in the family context and the parent's lived experience. Although paper-and-pen daily diary methods have long existed, the advent and popularization of cell phones make such frequent methods both more robust to collect data (e.g., no possibility to fill in responses later on, exact timestamp of response) and more convenient for participants, especially over long time-frames (Palmier-Claus et al., 2011). Such fine-grained data, collected day by day, allows for a micro-level investigation of how experiences of parental burnout fluctuate and interact with the family context.

Thanks to these methodological advances, this dissertation is thus part of the resurgence of family systems research described by Kerig (2019), by focusing on family dynamics, experiences, and affect in relation to a particular form of family dysfunction: parental burnout. However, these methods, and particularly network analyses as adapted for psychology research, had only rarely been used to study family processes when this dissertation project began (e.g., Bodner et al., 2018), and to this day, few family-related studies use such network approaches (e.g., Veenman et al., 2022). It is thus still unknown how well network analyses will be able to capture the reciprocal processes and dynamic nature of parenting experiences as embedded within the family system.

Another unknown concerns the distinct role of the different components of parental burnout. Parental burnout is a multidimensional construct, but at the start of this project, it had only been investigated as a unitary construct via sum scores. It can generally be useful to examine whether the separate components of a psychological construct have distinct roles and associations with other variables—in other words, if the components interact with one another and are not exchangeable. In such a case, useful information can be lost when investigating a construct only as a unitary whole and ignoring any possibility of the different features interacting with each other or their context in distinct ways (Schmittmann et al., 2013). Moreover, the different components of parental burnout might interact with the family context in different ways: for example, increased emotional exhaustion when faced with sibling spats and decreased emotional distance when a child comes to cuddle. As such, this focus on individual components and their distinct roles, and not just the overarching construct, is essential to both network analyses (Borsboom & Cramer, 2013) and family systems theory (Dallos & Draper, 2015). I therefore sought to investigate the possible distinct role of each component of parental burnout, not only in the development and maintenance of parental burnout, but also in relation to different aspects of the family environment.

When I began my thesis, I expected emotional exhaustion to play a central role, as researchers broadly view it as the core dimension of burnout (Halbesleben & Bowler, 2007) and it unites different conceptualization of burnout (Demerouti et al., 2001; Maslach et al., 2001; Pines & Aronson, 1988). In addition, multiple studies on occupational burnout suggested it was the first step in the burnout process (Lee & Ashforth, 1993; Leiter, 1993; Leiter & Maslach, 1988), and researchers theoretically proposed it played the same role for parental burnout (Mikolajczak & Roskam, 2018). However, as research on parental burnout had not yet distinguished between the components, no further evidence on emotional exhaustion sparking the development parental burnout existed.

1.1 Thesis objectives

The focal point of this thesis holds that to understand parenting—and therefore also parental burnout—we must examine it within its context: the family system. Although many challenges hinder researching the family via a systemic approach, I incorporated intensive longitudinal data collections and network analysis methods to closely examine parenting experiences of burnout and their reciprocal interactions with the family context. These methods allowed me to examine how the different components of parental burnout interrelate with interactions with children, the potential co-parent, and the wider environment, and whether the different components play distinct roles. For example, is emotional exhaustion central to parental burnout, as it is viewed to be core to burnout generally? Moreover, network methods had not yet been applied to research family systems, and so I also sought how well this analytical approach suited the investigation of the family system.

Thus, the main objectives of this dissertation can be summarized via two central aims:

1. What distinct role (if any) do the different components of parental burnout play, either in the development or maintenance of parental burnout, or in interactions with the family environment?
2. How well does a network approach suit the investigation of parenting experiences and the family system?

To answer these questions, this dissertation includes both theoretical and empirical chapters, organized in incremental steps as follows:

In the remainder of the **General Introduction**, **Chapter 2** provides a deep dive into network methods, including vocabulary and main concepts, the theoretical background for a network approach to psychology, and extensions and critiques. **Chapter 3** further argues why a network approach particularly suits investigations into parental burnout.

In **Section 2**, I set the stage for a network approach to family systems. First, **Chapter 4** provides a proof of concept that parental burnout and its family context can be investigated via a network approach, via a cross-sectional snapshot. In this study, a large dataset (N=1551) of unselected parents was reanalyzed via two types of cross-sectional network analyses. Measures include questionnaires on parental burnout, partner conflict and estrangement, and maltreatment toward children, and analyses investigate the distinct associations between the components of parental burnout, child maltreatment, and partner conflict. Further insight—

particularly about reciprocal interactions in the family system—required moving to longitudinal data collections and analyses. As no measures to examine fluctuations of parental burnout or the family context existed, I developed daily diary items for this purpose: **Chapter 4** details the development procedure and initial validation of these items. Next, in **Chapter 5**, I explore suitable analysis methods for conducting temporal network analyses via a scoping review.

In **section 3**, I present three studies that examine parental burnout and its family context from a temporal perspective, using intensive longitudinal data (i.e., daily diaries completed over 8 weeks). **Chapter 6** contains a temporal network analysis focusing on parents (N=50) from the general population, while **Chapter 7** contains a temporal network analysis of parents in burnout (N=40). In **Chapter 8**, I then reanalyze both datasets, as well as data from research collaborators, to examine the affective dynamics of parenting: do specific temporal trajectories of parental affect predict parental burnout severity?

Lastly, in the **General Discussion**, I discuss what the results from these varied studies imply for our theoretical understanding of parental burnout, and particularly what role the different components might play. I also discuss clinical implications for prevention and treatment of parental burnout. Next, I provide an in-depth evaluation of the strengths and flaws of using network analyses to examine family systems, and also discuss broader measurement and methodological considerations.

Chapter 2.

An introduction to networks in psychology

For decades, science has attempted to understand the world by boiling concepts down to their simplest components in a quest for reductionism; but, there is now a move to map the world's complexities via a network perspective. Clinical psychology has embraced this perspective as well, and researchers have started to conceptualize mental disorders as network systems. In this chapter, we review and discuss the theoretical conjectures, levels of evidence, current challenges and future directions, as well as the potential translational value of this new approach to psychopathology.

Reference

Blanchard, M.A. & Heeren, A. (2022). Ongoing and future challenges of the network approach to psychopathology: From theoretical conjectures to clinical translations. In G.J.G. Asmundson (Ed.), *Comprehensive Clinical Psychology* (2nd ed., vol. 11, pp. 32-46). https://doi.org/10.1016/B978-0-12-818697-8.00044-3

"I think the next [21st] century will be the century of complexity."

— Stephen Hawking (2000)

2.1 Embracing complexity within psychopathology

Over the past few decades, many scientific fields have changed the way they have sought to understand the world. Instead of attempting to reduce the complexities of their field into easy-to-understand models, they have instead embraced these complexities by mapping them out using networks (Barabási, 2012). This pursuit of complexity was enabled by both growing computational power (allowing more complex calculations on ever-larger datasets) and the development of network science. Network science contains key benefits that permitted its adoption by research fields as diverse as ecology, sociology, and epidemiology: (1) network science can apply to any type of system, with nodes representing the item under study (e.g., people, viruses, computers) and edges representing the connections between these nodes (e.g., correlations, temporal dependencies, probabilistic dependencies); and (2) network science grew out of graph theory, allowing researchers to understand and predict networks using mathematical rules. Psychology researchers have also recently recognized the potential that network science generates for understanding psychopathology (Contreras et al., 2019; for systematic reviews, see Robinaugh, Hoekstra, et al., 2019).

In this chapter, we first discuss how a network approach to psychopathology emerged and then review relevant network vocabulary and concepts. Afterwards, we discuss the main tenets of a network perspective of psychopathology, as well as recent extensions of this perspective. We then evaluate network theory by discussing its current main challenges and future directions. Lastly, we consider the current and possible future clinical applications of a network perspective of psychopathology.

2.2 From a latent disease model to a system approach

The understanding of mental disorders has for decades followed the medical disease model, where a single entity (usually conceptualized as latent, or unobserved, such as a still-undiscovered biological component) is assumed to be the common cause of symptom emergence and covariance (see Figure 1A; for a discussion, see Fried, 2015). This parallels how medical diseases are understood: A malignant lung tumor explains the emergence and co-occurrence of symptoms such as chest pain and coughing blood, for example, in the same way that depression would supposedly explain why anhedonia, fatigue, and other symptoms co-occur (Borsboom & Cramer, 2013; McNally, 2016).

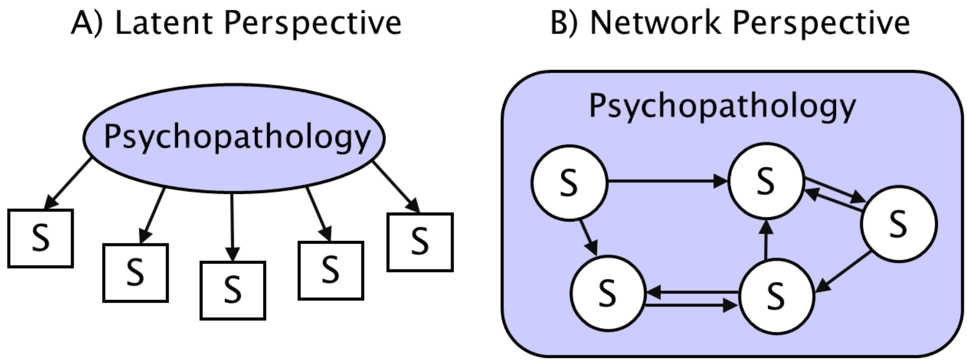


Figure 1. A latent vs. a network perspective of psychopathology

Note. Panel A: A latent perspective of psychopathology assumes that there is one main variable that causes all psychopathology symptoms. This model assumes that symptoms are interchangeable and independent. Panel B: A network perspective defines psychopathology through the causal interactions of its symptoms, where the symptoms activate and sustain one another. S = Symptoms. Inspired from "Commentary: A Network Theory of Mental Disorders," by P.J. Jones, A. Heeren. & R.J. McNally, 2017, *Frontiers in Psychology, 8,* p. 2.

This common cause model of psychopathology contains important fallacies. First, in a medical model, the underlying causal disease can exist without symptoms (e.g., asymptomatic lung cancer). However, the idea of someone having a mental disorder without any symptoms is baffling, because symptoms are precisely how we define and diagnose mental disorders. Second, a crucial assumption of the common cause model is that all symptoms are interchangeable; specifically, because all symptoms arise from a common entity, they all have equal importance in determining a diagnosis (Fried, 2015). This assumption is not realistic for most mental disorders. Take depression as an example, wherein specific symptoms impact functioning in different ways and are differentially linked with various risk factors and biological correlates (Fried, 2015; Fried & Nesse, 2014; Kendler, 2005). As another example, in the context of social anxiety disorder where fear and avoidance of social evaluative situations are the two hallmark symptoms, fear and avoidance are not interchangeable (for a discussion, see Heeren et al., 2020; Heeren & McNally, 2016). They are, instead, functionally distinct processes that merit different interventions, and merely relying on a sum score from a social anxiety questionnaire that aggregates fear and avoidance together would occlude this distinction (Heeren et al., 2020). Third, another important assumption of the common cause model is that all symptoms are independent from one another. Because a common entity causes them to arise, they do not influence each other (for further discussion of this assumption of local independence, see Borsboom, 2008; Holland & Rosenbaum, 1986). This assumption is also unrealistic for psychopathological symptoms. Both concentration difficulties and fatigue are symptoms of depression, for example, and it is clear that fatigue can lead to concentration difficulties and that insomnia may cause fatigue (Fried, 2015). Likewise, in anxiety disorders, fear may trigger avoidance, and vice versa (Heeren et al., 2020; Hofmann & Hay, 2018).

More generally, causal relationships between symptoms of mental disorders are consistently found, and these causal connections are at times even part of the diagnosis (Borsboom & Cramer, 2013). For instance, drawing on Mowrer's (1960) prominent two-stage learning theory, most diagnostic conceptualizations of anxiety disorders assume that avoidance symptoms—and use of safety behaviors in situations that the person cannot overtly avoid—directly trigger fear symptoms (e.g., Hofmann & Hay, 2018; for a discussion on the clinical relevance of distinguishing fear from avoidance, see also the seminal work of Rachman, 1978). All of these examples illustrate that a model assuming one common cause does not adequately explain the complexity of a mental disorder.

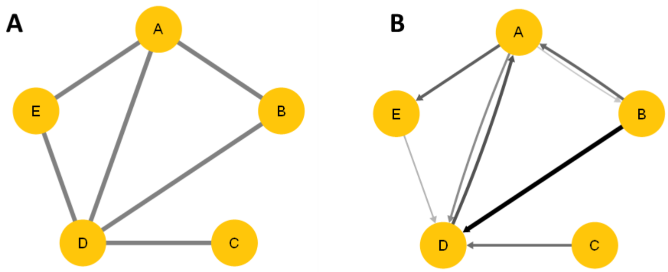
An alternate way of conceptualizing mental disorders is to view them as clusters of symptoms that tend to co-occur and co-vary across individuals. If similar clusters of symptoms co-occur regularly in different individuals, these clusters likely are not random but arise because their symptoms share a number of different causal mechanisms—a notion known as mechanistic property clusters (Kendler et al., 2011). Researchers and clinicians have grouped symptom clusters and named them as syndromes including, for example, major depressive disorder, social anxiety disorder, and bipolar disorder. Nonetheless, it is still crucial to examine these syndromes at a symptom level, because different people with the same diagnosis often exhibit different specific symptoms (Fried & Nesse, 2015), and because the symptoms often interact and reinforce one another, potentially becoming a vicious cycle (Borsboom & Cramer, 2013). Thus, to be able to understand how people develop and experience psychopathology, it is necessary to understand how symptoms influence and interact with one another. A key way to investigate symptom interactions involves visualizing and modeling mental disorders as networks of symptoms.

2.3 A network perspective of psychopathology

Psychology researchers have leveraged network science to build a network theory of psychopathology that addresses the fallacies of the latent disease model and extends our understanding of mental disorders. Instead of focusing on mental disorders as something concrete and specific, a network approach visualizes symptoms as nodes and maps out the associations between them. This emphasis on the interactions of symptoms shifts the focus from how to precisely define the boundaries of a disorder to how symptoms activate and sustain each other, and ultimately provides clues about how the network is self-sustaining. This reveals how network theory defines psychopathology—as (self-sustaining) networks of causally interacting symptoms (see Figure 1B). Psychology seeks to understand what causes symptoms, and this also is true of network theory. Network theory posits that symptoms cause (or, in other words, activate or inhibit) each other, through feedback loops and other interactions. To assess whether the interactions between nodes are in fact causal requires meeting many assumptions, such as including all relevant variables in the network, establishing temporal precedence, or investigating conditional independence (i.e., the association remaining between two nodes after controlling for their association with any other nodes; Epskamp & Fried, 2018). Most network analyses, therefore, investigate interactions between nodes that are indicative of causality, requiring either experimental research designs or extensive replications with full adherence to assumptions to make firm conclusions regarding causality (Epskamp, van Borkulo, et al., 2018).

Network models: A brief vocabulary lesson

Before discussing how network analyses can be applied to psychopathology, we will begin by reviewing some of the essential vocabulary and concepts of network science (for a thorough coverage of network science, see Barabási, 2016). A network is a graph that, together with a function, assigns a real number to each connection. A network analysis, for its part, issues a network to visualize and understand a system; that is, it traces out the connections between the system components (see Figure 2).



C

Nodes	A	B	C	D	E
A	0	1	0	1	1
B	1	0	0	1	0
C	0	0	0	1	0
D	1	1	1	0	1
E	1	0	0	1	0

D

Nodes	A	B	C	D	E
A	0	0.22	0	0.45	0.62
B	0.59	0	0	0.98	0
C	0	0	0	0.57	0
D	0.67	0	0	0	0
E	0	0	0	0.29	0

Figure 2. Example networks and their adjacency matrices

Note. Panel A: A network that is undirected (bidirectional connections) and unweighted (edges are binary, either present or absent). The circles represent nodes, and the lines represent edges. Panel B: A network that is directed (with edges that point from one node to another) and weighted (the edges vary in strength between 0 and 1, and this strength is reflected in width of the edge. Panel C: The adjacency matrix that accompanies Figure 2A. Each value records the edge weight that connects the specified two nodes. All edge weights are either 0 or 1 since Figure 2A is an unweighted network. Panel D: The adjacency matrix that accompanies Figure 2B. The edges vary from 0 to 1, with larger values represented in Figure 2B with thicker lines. The adjacency matrix is not symmetrical (unlike Figure 2C) since if an edge points from node A to node E, it is just represented once in the adjacency matrix. The R code used to estimate and visualize these two networks has been made publicly available via the Open Science Framework and can be accessed at https://osf.io/njsf6/.

The components (which can be any object, person, etc.) are called nodes, and the connections between nodes are called edges. A vast variety of systems can be conceptualized as a network. Examples range from a power grid (with generators as nodes and power supply lines as edges), to the world wide web (with webpages as nodes and their links towards one another as edges), to neural networks (with neurons as nodes and their connections as edges), to social networks (with people as nodes and their personal or professional ties as edges).

Networks have a number of different characteristics. First, networks can be either directed or undirected. In an undirected network (see Figure 2A), edges represent bidirectional links, whereas in a directed network (see Figure 2B), the edges point from one node to another. Second, networks can also be weighted or unweighted. In unweighted networks (see Figure 2A), edges are binary (i.e., either present or absent), while in weighted networks (see Figure 2B), each edge has a specific strength or weight, which is usually reflected by the width of the edge (i.e., thicker edges represent stronger connections). An adjacency matrix encodes all edge values within a network such that each node is represented by a column and a row in the matrix, and the weight of each edge can be found at the intersection of the two relevant nodes (see Figure 2D). If the network is unweighted, the edge weights will be encoded as 1 for present and 0 for absent (see Figure 2C).

A network representing a mental disorder is typically weighted, since symptoms form the nodes and the connections between symptoms are not known but estimated (such as through correlations, partial correlations, or other methods). Psychopathology networks can be based on either cross-sectional or time-series data. If the network is estimated using cross-sectional data, this data is collected at one time point, usually made up of self-report questionnaire responses. Cross-sectional psychopathology networks allow researchers to visualize and better understand how symptoms are generally related to each other, by mapping the correlations (or other statistical relationships) between symptoms. These networks are typically undirected, because cross-sectional data yields no information on how symptoms influence each other over time. A dominant estimation method is the Gaussian graphical model (GGM), which visualizes partial correlations between nodes. GGMs usually employ a regularization method to render the model sparse or, in other words, to only include edges that likely represent true connections (Epskamp, Waldorp, et al., 2018). There are a number of tutorials that explain how to estimate cross-sectional psychopathology networks using R code (e.g., Costantini et al., 2015; Epskamp, Waldorp, et al., 2018; Epskamp & Fried, 2018). More advanced methods, such as directed acyclic graphs (DAGs) and relative importance networks, allow researchers to model directed networks using cross-sectional data. Although DAGs (for illustrations, see Heeren et al., 2020; McNally, Mair, et al., 2017) and relative importance networks (for illustrations, see Heeren & McNally, 2016; Robinaugh et al., 2014) have occasionally been used in clinical psychology, they have increasingly become popular in other fields, such as epidemiology and environmental sciences.

Psychopathology networks can also be constructed from time-series data, collected, for example, through experience sampling methodology where participants answer short questions once or multiple times a day over weeks. These temporal networks are typically estimated using multilevel autoregressive (VAR) models (Bringmann et al., 2013, 2016), although there are other possibilities, such as group iterative multiple model estimation (Beltz & Gates, 2017) or sparse VAR models (de Vos et al., 2017). Networks estimated from temporal data result in directed networks that represent how symptoms influence one another over time. They require many timepoints but fewer participants than cross-sectional networks. With sufficient timepoints, a directed network can even be estimated using the data from just one person, yielding a personalized (or idiographic) network (e.g., Fisher et al., 2017).

Centrality

Using network theory to model psychopathology allows the adoption of network tools and graph theory concepts to understand mental disorders. One of these network concepts is node centrality, which posits that some nodes in the network are more important to the network structure than others. Nodes are considered more central if they are especially connected to the rest of the network (i.e., connected to many nodes), because they can then influence the entire network. As an example, picture a network tracing out how people gossip. If someone frequently gossips with other people, they form a central node in the network. If they share a specific story, it will quickly spread throughout the network. On the other hand, if someone gossips only with one or two people, they are represented by a node with few connections. If this person shares a story, it will only gradually spread throughout the network and might even remain limited to a small part of the network. Another example involves virus propagation: someone who sees many people in a day (i.e., very connected to other people) will, if infected, spread the virus to many more people than someone who sees only one or two people in a day.

A variety of centrality metrics exist, all using slightly different methods to indicate how connected a node is to the rest of the network (for an overview of different centrality indices, see Borgatti, 2005; McNally, 2016). Commonly used centrality metrics in psychopathology networks include strength (or degree) centrality (McNally, 2016), which computes the sum of absolute edge weights connecting to a node, and expected influence, which operates similarly while also taking into account the negative or positive valence of the edge (Robinaugh et al., 2016). There are many other common centrality metrics, including betweenness centrality (how many times a node lies on the shortest path connecting two other nodes) and closeness centrality (the average distance from one node to all other nodes), both of these indices involve assumptions that are very unlikely for psychological variables (Bringmann et al., 2019).

When interpreting node centrality through the lens of psychopathology, researchers posit that central symptoms are especially important to the course of the disorder, since if a central symptom is activated, it is more likely to activate and influence other symptoms (see Figure 3). This also adheres to the mathematical theory of preferential attachment, which posits that when new nodes are activated, they connect to central nodes first (Barabási & Albert, 1999). The concept that central symptoms are key to the course of a disorder is termed the centrality hypothesis (Borsboom & Cramer, 2013). The centrality hypothesis offers new avenues for research into how disorders develop and persist because it posits that central symptoms activate and sustain other symptoms, thus playing a key role in the etiology and maintenance of disorders. This also proposes central symptoms as candidates for prevention and treatment efforts, since if a clinician can “turn off” or downregulate a central symptom, then other symptoms are less likely to be activated or sustained (McNally, 2016).

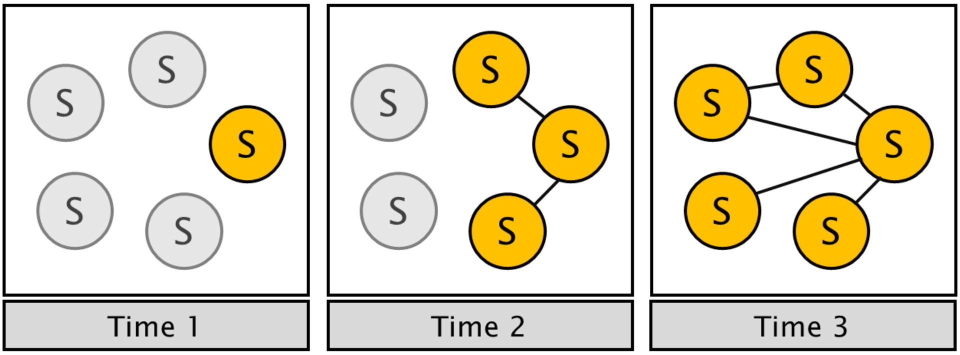


Figure 3. A central node activating other nodes

Note. Central symptoms have many connections to the rest of the network, and the mathematical theory of preferential attachment (Barabási & Albert, 1999) posits that when new nodes are activated, they connect to central nodes first. This figure provides an illustration of how this might occur. S = Symptoms.

On an empirical level, multiple studies have investigated whether centrality indices actually identify symptoms that play a role in the course of a disorder. In their systematic review, Robinaugh, Hoekstra, and colleagues (2019) report that such studies have found that central symptoms are more predictive of a later diagnosis than peripheral symptoms and that central symptoms are highly correlated with change in the rest of the network. But Robinaugh and colleagues also caution that the results from these studies—which all include the same measures in both the baseline and the subsequent timepoint—could also be explained by a common cause model. In this case, nodes that are identified as central symptoms in a network analysis would represent more reliable indicators (i.e., they would contain less measurement error), which would be more predictive of future psychopathology than less reliable indicators. A few recent studies present evidence that cannot be accounted for by a common cause model, however. Elliot and colleagues (2020) found that central anorexia symptoms at baseline were strongly correlated with recovery status and posttreatment clinical impairment prognostic utility. Since clinical impairment was measured independently from the symptoms included in the network, this presents more robust evidence for central symptoms influencing the course and recovery of a disorder. Likewise, Spiller and colleagues (2020) found that in a large sample of PTSD patients, only expected influence (and not other centrality estimates) predicted the association between changes in a node and changes in the entire network (presumably indicating a node’s importance to the network’s activation) in two different PTSD measures. However, other non-centrality metrics such as mean symptom severity yielded stronger predictions than expected influence. Evidence regarding the causal role of central nodes remains therefore decidedly mixed.

All studies presented above suggest only correlational associations between central symptoms and later psychopathology outcomes. As Spiller and colleagues (2020) emphasize, to assume a causal role for central nodes implies that these nodes causally interact with other nodes over time and within individuals; investigating the centrality hypothesis mainly with cross-sectional between-subject networks is therefore a major limitation. To obtain robust evidence for the centrality hypothesis and its implications requires an experimental paradigm, one targeting and manipulating specific central nodes (vs. peripheral nodes) and then examining how the network flow changes. If the centrality hypothesis (and its implications for clinical treatment) are supported, then downgrading a central node through some sort of intervention should lead to the entire network becoming less active (and vice versa for upgrading a central node). Targeting a peripheral node, on the other hand, should have a much lesser effect on the network.

These strong clinical predictions arising from the centrality hypothesis implicate many assumptions, such as whether one can even target a specific node on its own without also affecting other nodes (Bringmann & Eronen, 2018). There also is no universally accepted centrality metric, and the implications and interpretations of centrality metrics are not straightforward (Bringmann et al., 2019). Most centrality metrics were developed for social networks, which have people as nodes and their relationship, such as gossip partners as in the example above (or sharing an article authorship, being classmates, etc.) as edges (Bringmann et al., 2019). These centrality metrics therefore were developed under several assumptions that are problematic for psychopathology networks. One example is that the edges in social networks are often already known: the raw data already includes the relationships between people. With psychopathology networks, however, the edges are estimated, since the raw data does not include information on how symptoms relate to one another (Bringmann et al., 2019). Another unlikely assumption involves nodes being exchangeable: one person that has X number of connections plays the same role in a gossip network as another person who has the same number of connections. In psychopathology networks, however, nodes could have the same number of connections but represent symptoms of varying severity (e.g., suicidal thoughts vs. fatigue), and it could be problematic to assume they have the same impact on network functioning (Bringmann et al., 2019). Despite these limits, methods now exist to estimate the stability of centrality indices (Epskamp, Waldorp, et al., 2018), but only for cross-sectional networks and not temporal ones.

Connectivity

Another important network concept is connectivity, which refers to whether the overall network includes many connections (meaning its nodes are closely connected) or few connections (meaning its nodes are relatively independent and isolated). In general, if a network is more connected, then activation or information will spread through the network more quickly. Considering the previous example of a gossip network again, if most people in the network frequently talk to many other people, then gossip will rapidly spread throughout the network. If, however, the network is not strongly connected, and most people only gossip seldomly and with few individuals, then it will take longer for specific stories to reach everyone in the network. Different researchers have used different ways to measure connectivity, including the number of edges (as a ratio over the number of possible edges; de Vos et al., 2017) and the average of the absolute values of edge weights (Bringmann et al., 2016).

Within psychopathology networks, researchers have posited that the overall network connectivity is linked with severity. If a network is more densely connected, then it is more likely to be or to become pathological (Borsboom, 2017). This relies on the assumption that in a more connected network, activity will spread throughout the network more quickly, and symptoms will activate each other more easily and influence each other more. Borsboom and Cramer (2013) likened a network with high connectivity to a set of dominoes lined up close together, such that if one domino goes down (or one node is activated), it sparks a cascade of dominoes falling down (or a cascade of node activation). The notion that symptom networks that are highly connected are likely to be more severely pathological (or more at risk for developing psychopathology) is termed the connectivity hypothesis.

At the empirical level, there has thus far been mixed support for the connectivity hypothesis. In their systematic review, Robinaugh and colleagues (2019) identified a number of cross-sectional analyses that investigated connectivity through group comparisons. Different methods include comparing people with a mental disorder to those without (e.g., Heeren & McNally, 2018), comparing the baseline measures for groups whose symptoms either persisted or remitted (e.g., Van Borkulo et al., 2015), and investigating connectivity before and after treatment (e.g., Bryant et al., 2017). Robinaugh and colleagues (Robinaugh, Hoekstra, et al., 2019) concluded that greater connectivity usually accompanies greater psychopathology severity (e.g., a diagnosis); but they found mixed results for the other comparisons. Investigating connectivity in cross-sectional networks presumes that connectivity is linked with severity at a group level, while it is possible this only occurs on an intra-individual level. Temporal networks can solve this uncertainty, since they allow investigations into intra-individual network dynamics. A number of studies investigating connectivity in mood state networks have found that increased connectivity accompanied increased severity (Bringmann et al., 2016; Hasmi et al., 2017).

The connectivity hypothesis rests on the assumption that increased connectivity indicates the presence (or likelihood) of feedback loops which sustain symptoms; but, many questions still remain regarding how best to examine connectivity. For example, psychopathology networks often include both positive and negative edges. If symptoms (or other contextual components) inhibit one another (as negative edges indicate), does connectivity still imply a self-sustaining vicious cycle? There are also many questions about how to compute connectivity (for discussion, see de Vos et al., 2017). For example, should researchers include only positive edges, or only negative edges, or both (e.g., Bringmann et al., 2016)? Should they compute either inter-node connections, or intra-node connectivity (i.e., autoregressive feedback loops), or both (e.g., Groen et al., 2019)? And, should they include only significant edges, or all edges (e.g., de Vos et al., 2017)? These questions are currently unanswered, and more research is required to determine how these considerations affect the connectivity hypothesis. In light of this uncertainty, researchers should follow the example of De Vos and colleagues (2017) and include multiple data preprocessing and estimation choices to examine how these affect their results. Future research can also go even further and conduct multiverse analyses, where a multitude of datasets are constructed from the many alternative data processing choices and analyses are conducted on each dataset (Steegen et al., 2016).

The literature on the connectivity hypothesis is still rather limited, especially regarding studies that investigate connectivity using temporal networks. Through this brief overview of that literature, it should be obvious that substantial work remains to determine how preprocessing and analytic choices affect results and interpretations. Nonetheless, the connectivity hypothesis proposes a novel approach to understand and investigate psychopathology severity, by focusing on how symptoms relate and interact with one another. Some researchers also advocate for additional investigations into connectivity that go beyond the connectivity-as-vulnerability hypothesis (Groen et al., 2019), such as by pursuing questions that arise from the complex dynamical systems theory literature (for an overview, see section above on Network changes: Critical transitions).

Comorbidity

With its focus on symptom-to-symptom connections, network theory proposes an intuitive understanding of comorbidity (i.e., dual or multiple diagnoses). Comorbidity has long interested and puzzled psychologists, especially because the rates of comorbidity are high. For people who have one diagnosis, around 45% meet the criteria for another diagnosis as well (Kessler et al., 2005). Under a disease model of psychopathology, which conceptualizes disorders as distinct entities with strict boundaries, attempting to understand why and how people have multiple disorders is a complicated endeavor. Theories of comorbidity range from statistical artefacts arising from the construction or measurement of disorders (i.e., large overlaps in symptoms for many disorders in the Diagnostic and Statistical Manual of Mental Disorders), to actual underlying connections between two latent disorders (such as through a common etiological mechanism), to the existence of a 'p factor' (i.e., a general factor predisposing people to psychopathology analogous to the general g factor for intelligence; Caspi & Moffitt, 2018; Cramer et al., 2010). However, investigating comorbidity by focusing on diagnoses, as suggested by the disease model of psychopathology, makes little sense because the boundaries between disorders are rather blurred (Cramer et al., 2010).

Network theory, on the other hand, offers a straightforward explanation for how people develop comorbid disorders. Because network theory defines a disorder through its symptom-to-symptom connections, comorbidity would occur through common connections that bridge disorders. These common connections can operate through overlapping symptoms or symptom-symptom connections that exist in the co-occurring disorders, but also through symptoms that connect them to the same process (Cramer et al., 2010). Symptoms which build a path from one disorder to another are also called bridge symptoms. Borsboom and Cramer (2013) give the example of how chronic worry (a main symptom of generalized anxiety disorder) can lead to sleep problems and thus fatigue, both of which are simultaneously symptoms of generalized anxiety disorder and depression. Sleep problems and fatigue, in turn, can then lead to depressed mood, which is a main symptom of depression. This example demonstrates how pathways that connect symptoms of a disorder to one another can also lead to the activation of symptoms from a comorbid disorder; specifically, symptoms do not stop interacting just because they reach a theoretical diagnostic threshold. The focus of network theory on symptoms and their connections allows intuitive investigations into comorbidity that explicitly recognize the blurry boundaries separating mental disorders.

On an empirical level, studies regularly find that symptoms of one disorder are differentially related to symptoms of another disorder (e.g., Bekhuis et al., 2016), supporting a perspective of comorbidity that focuses on symptoms rather than on syndromes. In line with this, studies utilizing network approaches to investigate comorbidity often focus on identifying symptoms that bridge two disorders. This can be done with centrality indices that specifically measure bridge centrality (i.e., how well-connected a node is with symptoms of another disorder; Jones et al., 2019). A specific example includes an investigation into the comorbidity of social anxiety disorder and depression (Heeren, Jones, et al., 2018). This investigation identified suicidal ideation, loss of interest, and anhedonia as depression symptoms closely connected with social anxiety disorder, while avoidance of participating in small groups and avoidance of attending parties were symptoms of social anxiety disorder closely connected with depression. Another study investigated comorbid anxiety and eating disorders (ED) and found that avoidance of social eating had the highest bridge centrality among ED symptoms, while low self-confidence had the highest bridge centrality among anxiety symptoms (Forrest et al., 2019). Another way to investigate comorbidity with network analyses uses directed acyclic graphs, a Bayesian approach that generates directed networks from cross-sectional data (see section above on Network Models: A Brief Vocabulary Lesson). Researchers using this approach have identified nodes that seemingly drive activation of both disorders and have labeled these as probable bridge symptoms (McNally, Mair, et al., 2017).

In all of these studies, the researchers discuss how the identified bridge symptoms could be ideal targets of prevention efforts that seek to prevent someone with one disorder from developing a comorbid disorder, or targets of treatment efforts that seek to successfully consider both (or multiple) comorbid disorders. To give a specific example, Levinson and colleagues (2018) identified difficulty with public eating and drinking as a symptom bridging social anxiety and eating disorders. They suggested that future clinical investigations should examine whether exposure therapy that specifically targets public eating and drinking beverages (vs. targeting other non-bridge symptoms) can decrease symptoms for both social anxiety and eating disorders.

These clinical possibilities are intriguing; but extensive research is needed to establish the validity and effectiveness of any clinical applications that target bridge symptoms. For one thing, there are still relatively few network investigations into comorbidity, and no replications have yet been conducted to verify the results of any of the above studies in different samples. For another, clinical trials are required to evaluate whether targeting bridge symptoms is actually effective (and/or more effective than existing interventions) in treating and preventing comorbid disorders. Moreover, there is not yet a concrete understanding of how exactly bridge symptoms might affect comorbidity (Cramer et al., 2010). For example, do specific bridge symptoms increase the risk of developing a comorbid disorder when compared to other (non-bridge) symptoms? Does intervening on bridge symptoms in people with comorbid disorders lead to slower symptomatology for both disorders, or prevent the development of further comorbidity? Answering these types of questions would further clinical knowledge regarding comorbidity, and arriving at these answers requires experimental study designs. Nonetheless, the network approach, even in its current form, is useful for investigating comorbidity, as it allows researchers to examine how symptoms interact regardless of the boundaries of a diagnosis (Fried, 2015). Specific recent developments in bridge centrality metrics (e.g., Jones et al., 2019) also allow researchers to easily identify potential bridge nodes, which can help researchers piece together which symptoms seem especially relevant in connecting two (or more) comorbid disorders.

Network changes: Critical transitions

Prior to the development of network theory, there were already branches of systems science that conceptualized mental disorders as complex dynamic systems (e.g., Hayes & Strauss, 1998). Network theorists have thus adopted some of the approaches and analytic methods of system science to investigate psychopathology networks (for a brief overview, see Robinaugh, Hoekstra, et al., 2019). One specific example involves examining how and predicting when a complex system transitions from one state to another. This has been researched in climate science, ecosystems, and other fields (Scheffer et al., 2009). Investigating and predicting transitions between stable states can also be highly useful to psychopathology research, such as when viewing "healthy" functioning as one stable state that can transition into a state of psychopathology, or vice versa (Borsboom, 2017; Hayes et al., 2015). Complex systems usually provide early warning signals when they are on the brink of tipping into a new state. Recent research has identified one such warning signal as critical slowing down—the system as a whole becomes more synchronized, with its components showing high variance and high autocorrelations (for a more detailed explanation, see Scheffer et al., 2009). Psychology researchers have hypothesized that when individuals approach a tipping point from a healthy state to a depressed state, this transition is also preceded by a critical slowing down of the psychological system (van de Leemput et al., 2014).

On an empirical level, few studies to date have investigated critical transitions in psychopathology. Researchers have found support for critical slowing down preceding a transition from a healthy state to a depressed state, in both group-level (van de Leemput et al., 2014) and individual (Wichers et al., 2016) data. Others have identified critical slowing down in two specific affect variables (i.e., feeling 'alert' and 'inspired') that predicted greater changes in depression symptoms over time (Curtiss et al., 2019). Phase transitions can also be investigated within a disorder, for example to clarify the transition pathways from a depressive episode to a manic episode in bipolar disorder. Nakamura and colleagues (2013) found evidence that critical slowing down of mood fluctuations and physical activity predicted a transition from depression to hypomania in one patient.

Of the few studies that have thus far investigated critical transitions in psychopathology networks, most have focused on depression. Therefore, extensive further research is required to draw robust conclusions about whether early warning signals can reliably indicate a transition into (or out of) a state of psychopathology, and whether this is limited to specific psychopathology states (such as depression) or applies more widely to other forms of psychopathology. More research is also required to ensure that the indices used to investigate early warning signals (e.g., autocorrelation or variance) are reliable. With additional theoretical validation and methodological robustness, important clinical implications arise. For one thing, this will allow psychology researchers to better identify general early warning signs that signal a tipping point into psychopathology or recovery (e.g., Curtiss et al., 2019), and then strategically use these signs in targeted treatment approaches. Being able to detect early warning signals for transitions involving psychopathology states could also improve when and how treatment (or prevention) efforts are implemented, on both a general (van de Leemput et al., 2014) and an individualized level (e.g., Wichers et al., 2016).

2.4 Extensions of network theory

Beyond symptoms

The original iterations of network theory mostly focused on networks comprised solely of psychiatric symptoms (e.g., Borsboom, 2017); but the emphasis on causal interactions between symptoms propelled other researchers to include any variables that had a plausible causal role within the system, whether formally recognized as symptoms or not (Jones et al., 2017). To draw conclusions about causal connections, a key assumption involves including all relevant variables in the model (Epskamp, van Borkulo, et al., 2018), and it is likely that relevant variables are not limited to DSM-recognized symptoms. This has led some researchers to include psychological processes of mental disorders, as studied with laboratory measures in experimental psychopathology, as nodes in psychopathology networks. For example, Heeren and McNally (2016) included as nodes theory-driven cognitive and behavioral mechanisms (i.e., attentional bias for threat, attention control, emotional reactivity to a laboratory-induced stressor, behavioral avoidance) thought to be involved in the maintenance of social anxiety disorder. They found that the orienting component of attention as well as avoidance were central nodes in the network structure. As another example, Everaert and Joormann (2019) included nodes representing repetitive negative thinking and positive reappraisal, two emotion regulation strategies, in a network of depression and anxiety symptoms. They found that repetitive negative thinking had high bridge centrality, and so played a large role in in the comorbidity between depression and anxiety symptoms. In today’s literature, the use of network analysis to examine the “bridges” between psychopathology symptoms and psychological processes thought to be involved in psychopathology has become one of the most fruitful research agendas, given that it generates an especially strong enthusiasm by the community of clinical psychological researchers (e.g., Heeren et al., 2020; Kraft et al., 2019; N. H. Weiss et al., 2020).

Allowing networks to include more than just symptoms also paves the way for networks that go beyond symptoms, and researchers have started to investigate how features of different transdiagnostic phenomena thought to be involved in psychopathology interact together. For instance, Bernstein and colleagues (2017) relied on a network approach to investigate how features of executive control interact with features of rumination. Likewise, Greene and colleagues (2020) recently examined the interplay between features of maladaptive daydreaming and emotion regulation strategies. Furthermore, others have even started to embrace a network perspective to examine how the constitutive components of a given transdiagnostic phenomenon interact together. For instance, Bernstein and colleagues (2019) proposed reexamining trait rumination—a core transdiagnostic process thought to be involved in the etiology and maintenance of several mental disorders—as a network system of interacting components (i.e., nodes). In the same vein, Heeren and colleagues (2018) offered a network approach to trait anxiety—a key transdiagnostic mechanism of anxiety and related disorders.

Other psychological phenomena

A network approach has not been limited to psychopathology; indeed, it has also been applied more broadly to a variety of psychological phenomena. Adopting a network approach encourages novel theoretical conceptualizations that focus on component-level interactions, as well as new statistical tools and visualizations. To go into detail about how a network approach can bring new theoretical understanding to a field of study, we will focus on personality research (Cramer et al., 2012). A traditional latent view of personality posits that personality dimensions are latent constructs which cause the items they measure; for example, being extraverted causes a person to like attending parties. A network perspective instead posits that the way people act and feel influences the way they think about themselves, and some of these behaviors and feelings naturally cluster together (such as "feeling comfortable starting conversations" and "liking parties"). When modelling these various personality components in a network, the network architecture then reveals clusters that can be interpreted as personality dimensions. Previous theorists have critiqued the latent model of personality as incompatible with the way specific personality components dynamically interact with their environments. A network approach offers an alternative theory of personality (with accompanying methodological and analytical techniques), where personality dimensions consist of the gradually consolidated and subsequently maintained interactions between the way people act, think, and feel within specific environments.

Researchers have similarly applied network analysis to intelligence, using it to unite models of cognition and intelligence by positing that the cognitive system develops mutually with different facets of intelligence (Van Der Maas et al., 2017). Network science has also been used to formalize a theory of attitudes, tying together the interactions between beliefs, feelings, and behaviors in a way that promotes testable predictions (Dalege et al., 2016). Network analyses have even been proposed as a useful tool for explaining functional disorders (e.g., fibromyalgia, chronic fatigue) because a focus on symptom interactions makes sense given the lack of unique pathophysiology and large individual differences in symptom patterns (Hyland, 2019). In addition, using network analysis to conceptualize functional disorders allows intuitive metaphors that patients appreciate, such as viewing functional disorders as "programming errors" that arise from a complex biological and behavioral system (Hyland, 2019).

In general, a network approach to psychological concepts discards a common cause explanation and instead proposes that psychological phenomena arise from an interactive system of components. As such, a network approach always prioritizes investigations at the component-level instead of focusing on the phenomenon as a whole (Bringmann & Eronen, 2018), which can be useful to reconceptualize research and theory for a variety of psychological fields of study.

Extending inwards to biology and outwards to social context

When considering disorders as causally interacting networks, it is clear that symptoms and psychological phenomena do not interact in isolation but always within a context (Borsboom & Cramer, 2013). This context is first and foremost located within a biological system, our human bodies. When someone feels lethargic, or cries, or feels anxious, a whole host of biologically based components are at play, from hormones to neurotransmitters to protein chains to genes. Extensive research has concluded that a singular biological component is never solely responsible for how a person experiences feelings, emotions, or behaviors (Kendler, 2005). This is also true on a broader level. Even phenomena that are thought as purely biological—such as diseases—are typically much more complicated than a single genetic abnormality (Barabási et al., 2011). Indeed, medical research has embraced network science, because it allows investigating diseases as complex intracellular networks that interact with other (internal and environmental) systems (Barabási et al., 2011). This has prompted Guloksuz and colleagues (2017) to argue that the strict dichotomy between medical diseases and mental disorders, which motivated the adoption of a network approach to psychopathology, should be relaxed. They propose that the main difference between medical diseases and mental disorders is that researchers have identified more of the biological components involved in disease networks, while psychopathology networks still remain mostly at the level of behavioral and affective signs and symptoms (Guloksuz et al., 2017).

These psychological symptoms are also located within an external environment. Humans are embedded within a social context and interpret their experiences through that social context. The most immediate social context is an individual's family, and family systems theory asserts that an individual's behaviors and feelings, as well as any psychological symptoms, are always in conversation with an interpersonal context (Bavelas & Segal, 1982). Indeed, humans do not exist in isolation but are constantly acting and reacting to their environment—not only immediate environments such as the family, workplace, or school, but also broader environments such as the neighborhood, media, and social services, and macro-level contexts such as culture (Bronfenbrenner, 1979). Bronfenbrenner's varied ecological contexts have even been recently theorized as overlapping and interconnected networks, where various social interactions in different contexts together shape an individual's development and expression (Neal & Neal, 2013), including the way they express psychopathology symptoms.

Oude Maatman (2020) discusses in depth how specific network connections can vary according to the cultural and historical context within which they are embedded. As an example, he considers the connection between loss of honor and suicidal ideation. In WWII-era Japan, a loss of honor was a reason to consider committing suicide, as suicide was the only way to regain honor. For a devout Catholic, however, the connection is traditionally the opposite, with suicide representing a gateway to hell and something to avoid at any cost. In both cases, the cultural context can determine what connections are made between mental states. These connections could be automatic (formed through neural plasticity) or cognitive (such as if thinking about loss of honor naturally leads to thinking about suicide). In either case, this example demonstrates that culture can structure how mental states are causally connected within individuals. As such, researchers should consider including cultural contexts as factors in network analyses when relevant to the network in question.

Symptoms—and, more broadly, any psychological processes—are inextricable from their contexts, from the biological level up to the cultural level. This supports extending psychopathology networks to include the contexts of disorders. A few network studies have already begun to combine psychological networks with contextual components. For example, a network analysis of positive and negative affect in bipolar disorder included actigraphy measures of physical activity (Curtiss et al., 2019), while another study investigated the connections between affective states, genetic vulnerability to psychopathology, and childhood trauma (Hasmi et al., 2017). One network study investigated how psychotic experiences and momentary affect interacted with minor daily stress and being alone (Klippel et al., 2018), while another study examined how social media use co-varied with depressive symptoms (Aalbers et al., 2019). A third study examined how COVID-19-related worries and beliefs in the exaggeration of its threat interacted with avoidance and coping behaviors (Taylor et al., 2020), and another investigated how parental burnout interacted with family-related variables (Blanchard et al., 2021; for a discussion, see Blanchard & Heeren, 2020). These examples highlight how the larger, ongoing efforts of the field to incorporate contextual variables and investigate how symptoms vary in relation to their inner and outer environments fulfill an especially valuable and useful niche in clinical psychology. The future possibilities of contextual networks expand even further. Guloksuz and colleagues (2017) propose the concept of layers of networks, with each layer situated at a different level of analysis. For example, a layer of biological components would be embedded within a layer of signs and symptoms, embedded within a layer of social interactions and environmental factors, embedded within a layer of cultural and historical context (see Figure 4). While nesting network analyses within different layers is not (yet) computationally possible (although multiplex networks are getting close; Shang, 2015), Borsboom and Cramer (2013) foresee the simultaneous analysis of social, symptom, and physiological networks as both a major challenge and a potential key merit of network models. We hope that with future improvements of network methodology and computational tools, layering networks to discover the connections between components at many different levels will be possible.

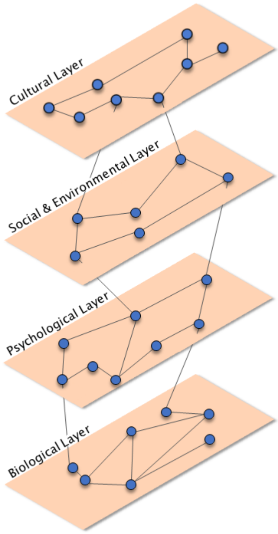


Figure 4. Example of a layered network

Note. A series of nested networks, with a biological layer embedded within a psychological layer, embedded within a social and environmental layer, embedded with a layer or cultural and historical context. Inspired from "Application of network methods for understanding mental disorders: Pitfalls and promise," by S. Guloksuz, L.K. Pries, & J. Van Os, 2017, *Psychological Medicine, 74*(16), p. 3.

2.5 Evaluating network theory

Ongoing challenges and future directions

As we have covered already in different sections (see sections above on Centrality, Connectivity, and Comorbidity), network theory still lacks validation for its main tenets (Wichers et al., 2017). The network approach to psychopathology is in many ways still in its infancy. Although network theory comes with many tools stemming from graph theory (such as centrality indices) and systems science (such as early warning signals for phase transitions), these tools still need to be adapted and validated for use with psychopathology networks. Substantial work is also needed to improve the methodological robustness of network theory. There currently exists considerable variation in terms of data (pre)processing, handling of model assumption violations, model estimation techniques, and so on, all of which can affect results (Bulteel et al., 2016; de Vos et al., 2017). We currently know little about the reproducibility of network analysis results; indeed, there are very few studies that cross-validate their results by splitting their sample, and even fewer exact replications using different samples, despite the recent exponential growth in (cross-sectional) network studies (Guloksuz et al., 2017). There are a few examples of efforts to evaluate the replicability of network analyses, particularly focusing on PTSD network analyses: a systematic review and a meta-analysis concluded that the results were relatively replicable (despite large study heterogeneity), but that centrality estimates might not be that informative for the relevant networks (Birkeland et al., 2020; Isvoranu et al., 2021). Such evaluations of the replicability are needed for other domains within psychopathology network analyses, to allow a more comprehensive view of the field. Another critical challenge concerns the uncertainty regarding how to estimate adequate sample sizes for network analyses. For cross-sectional networks, post-hoc power estimations exist (Epskamp, Waldorp, et al., 2018), and rules of thumb for sample sizes can be borrowed from structural equation modeling (e.g., minimum 10 participants per parameter; Schreiber et al., 2006). Estimating power a priori to plan an adequate sample size is much more complicated and currently requires simulations based on a similar network model, which is rarely available (Epskamp & Fried, 2018). For temporal networks, even less is known about the optimal sample size and number of timepoints.

Another important consideration concerns whether the network approach should be empirically-driven or theory-driven, or both. Most network analyses currently work to uncover patterns in data, and are thus exploratory and empirically-driven; this is to be expected, considering the recency of network analyses to psychopathology (Robinaugh, Hoekstra, et al., 2019). A number of researchers recommend moving toward confirmatory network analyses relying on formal theories, potentially using computational models of specific disorders (for further information, see Haslbeck et al., 2022; for an illustration with panic disorder, see Robinaugh, Haslbeck, et al., 2019).

What network theory brings now

Even in its current form, however, the network perspective of psychopathology is useful. It offers a new model to understand mental disorders, with a focus on the components and their interactions instead of on one unitary syndrome. The different hypotheses stemming from a network perspective of psychopathology (e.g., the centrality hypothesis and the connectivity hypothesis) require further validation, in addition to the general methodological improvements still needed to ensure the robustness of network analyses. However, although the network perspective can be viewed as a realist model of how psychopathology truly functions (which is what its theoretical founders adopt; Kalis & Borsboom, 2020), it can also be viewed as a tool we can leverage to better understand psychopathology (van Loo & Romeijn, 2019). Following this instrumentalist interpretation, which views models as instruments to investigate the world instead of assuming they strictly represent the world, researchers can use a network perspective to better understand and predict psychopathology, without necessarily "committing to the idea that [it] provide[s] a picture of the world" (van Loo & Romeijn, 2019).

According to this instrumentalist interpretation, network models should then be used whenever they are especially suited to a research goal. For example, network models allow intuitive visualizations of how multiple symptoms interact with each other and with other components. Relatedly, network analyses can be used as a data-driven tool to investigate patterns in how symptoms cluster together and co-vary. Guloksuz and colleagues (2017) suggest investigating how symptoms hang together when ignoring DSM-based category diagnoses, such as by performing network analyses on broad transdiagnostic or early vulnerability data. Network models also provide intuitive ways to conceptualize and investigate comorbidity, by focusing on bridge symptoms instead of diagnostic categories (Cramer et al., 2010). Finally, with a focus on symptoms and components, network theory privileges investigating processes of change. Indeed, when researchers investigate many variables with dense time series data that allows bidirectional associations, as with temporal networks, they can investigate how symptoms change over time (Hofmann et al., 2020). Wichers and colleagues (2017) suggest that this is where the promise of a network perspective of psychopathology truly lies—in moving toward a dynamic view of psychopathology. Indeed, a network perspective encourages dynamic questions: Why do specific groups of people develop psychopathology? How and when do they develop or recover from psychopathology? Trying to answer these questions is essential not only for improving our understanding of psychopathology, but also for treating it.

2.6 Current and future clinical applications of the network perspective

Even in its current formulation—still requiring theoretical validation and methodological advances—the network perspective of psychopathology already can inform clinical practice, by providing new ways to think about mental disorders. The focus on symptom interactions is generally more useful than focusing only on disorder categories, since there is often large variation of symptomatology within a diagnosis (Fried & Nesse, 2014). Many clinicians already focus on symptoms and not diagnoses when deciding on treatment options (Waszczuk et al., 2017), and the network perspective encourages research to follow this path. Research that adopts a network perspective can also yield new information on general group-level mechanisms of psychopathology, which can inform clinical treatment. It is important to note, however, that although theoretical development often requires the use of a group-level approach, these group results may not generalize to network systems as they occur within an individual, and it is these individualized networks that allow personalized clinical recommendations for a specific client (Fisher et al., 2018).

As a result, ongoing computational efforts further the growing possibilities of personalized network models. These models allow individualized insight into one person's symptom dynamics, by using intensive time series data from an individual participant to build an idiographic temporal network (for an illustration, see Fisher et al., 2017). This can allow the clinician, but also the patient, to have a personalized visual window into how their symptoms interact and feed off one another (Epskamp, van Borkulo, et al., 2018). Personalized networks could also provide a more fine-grained approach to best identify targets for meaningful prevention and intervention (Fisher et al., 2017). With sufficient additional research, clinicians could even use insight from phase transitions and early warning signals (see section above on Network changes: Critical transitions) to predict when a patient is about to enter a disordered state or return to a healthy one. This can allow more efficient and targeted treatment for specific individuals, while also improving general understanding about risk and resilience requisite for fine-grained clinical recommendations for a specific client.

If the network perspective and its related hypotheses (see sections above on Centrality and Connectivity) were validated, then network interventions would become a possibility. The centrality and connectivity hypotheses both posit that to modify a network, turning a central node off (or on) can affect the entire network system. This opens up possibilities for very targeted interventions on a few isolated symptoms that could have large impacts. This also applies to comorbid disorders, where targeting bridge symptoms could have therapeutic effects for both/all comorbid disorders. Network interventions are the most speculative clinical application of network theory, but their specific targeting also offers an innovative way to think about and plan treatment.

We hope that future clinical developments dovetailing with the network perspective of psychopathology will follow the research branch in openly sharing data, code, methodology, and tutorials that currently characterize this field (for an example of an open textbook, see Barabási, 2016; for examples of tutorials, see Costantini et al., 2015; Epskamp & Fried, 2018; for examples of empirical studies sharing code and data, see Aalbers et al., 2019; McNally, Mair, et al., 2017). For clinical applications, open materials could include open-source phone applications for patients and software to automatically summarize and visualize data. This would allow clinicians to easily and inexpensively incorporate network applications within their practice, generating treatment-related insights for both them and their patients. Incorporating a network approach within clinical work meshes with the main goals of network theory: attempting to understand psychopathology in all of its complexities and iterations.

2.7 Conclusion

We demonstrated how a network perspective of psychopathology acknowledges the complexity of mental disorders by viewing them as self-sustaining systems of symptoms. A network approach to psychopathology still requires substantial development, including theoretical validation and more sophisticated statistical analyses. Nonetheless, a network perspective offers not only new theoretical avenues for research on psychopathology but also crucial translational opportunities for adapting a network perspective to inform clinical practice.

Chapter 3.

Why a network approach to parental burnout?

Network science has allowed varied scientific fields to investigate and visualize complex relations between many variables, and psychology research has begun to adopt a network perspective. In this paper, we consider how leaving behind reductionist approaches and instead embracing a network perspective can advance the field of parental burnout. A network approach allows simultaneous investigations (and clear visualizations) of many variables and their interactions, integrates smoothly with prior family systems theories, and prioritizes dynamic research questions.

Note: the original version of this article was written early on in my thesis project and contains substantial overlap with the previous chapter on a network approach to psychology. As such, this is a summarized version containing the key aspects relevant to the next chapters. For readers interested in the complete version, they can find further details in the reference below.

Adapted from:

Blanchard, M. A., & Heeren, A. (2020). Why we should move from reductionism and embrace a network approach to parental burnout. *New directions for child and adolescent development, 2020*(174), 159-168. https://doi.org/10.1002/cad.20377

For decades, science has attempted to understand the world by boiling concepts down to their simplest components, such as searching for a singular biological element or straightforward model to account for complex psychological processes (Barabási, 2012, 2016). Yet, this quest for reductionism has fallen short. Scientists from different fields have repeatedly discovered that complex systems work together to generate ecological, social, human, and biological processes, and that singular components or simple models cannot adequately explain these complex systems. There is now a move to map the world's complexities using a network approach, visualizing any system as a network to allow scientists to understand how a system's many components interact. This network takeover has only become possible in the last few decades, with growing computational power, and many fields have adopted this network science to map their systems of interest. Recently, psychology research has begun to embrace this network perspective as well, particularly to investigate psychopathology (Borsboom, 2017) but also other constructs, from personality (Cramer et al., 2012) to emotions (Lange et al., 2020).

However, this approach has only rarely been used to investigate families. From our understanding, only study before this dissertation used this network analyses, specifically to model affect over time between mother, father, and adolescent during a recorded interaction (Bodner et al., 2018). This paucity is surprising, because family systems theory—derived from general systems theory (von Bertalanffy, 1950)—specifically conceptualized the family as a network of interacting components. Indeed, the commonly termed "systemic" approach (e.g., Jackson, 1965; Minuchin, 1974; for an overview, see Dallos & Draper, 2015) has theoretically prevailed for years in family research and therapy. This approach holds that understanding an individual's cognitions and behaviors requires also understanding the individual as embedded within a context and examine how that individual affects and responds to others in their family system (e.g., Bavelas & Segal, 1982; Watzlawick et al., 1967).

This approach grew out of general systems theory, formulated to apply widely across scientific domains—from physics to biology, and from sociology to psychology (von Bertalanffy, 1950). Early conceptualizations of families as systems (particularly in the 1950s) hewed close to how systems worked in other fields: mechanistic, with an assumption that if all interaction patterns could be identified and mapped out, then the family network would be wholly understood (Dallos & Draper, 2015). As the field of family systems developed over the next decades, particularly in clinical practice, the way of understanding families as systems shifted, with increasing focus placed on how meaning and identities are built and exchanged and ever-changing in family systems (Dallos & Draper, 2015). With this shift to more qualitative understandings of the family as a system came increasing importance placed on the broad context surrounding and shaping the family, particularly cultural and gender norms. Although viewing family processes as a dynamic system has remained popular through shifting lenses in both psychological theory and therapy (Bavelas & Segal, 1982; Sydow et al., 2013; Watzlawick et al., 1967), the ability to perform analogous statistical analyses remained limited. However, with the recent intense development in network science, there now exists both the computational capability and the mathematical tools to study the family as an actual dynamic network system.

Parental burnout is a prime candidate to represent and investigate using a network approach, as parental burnout is inextricable from its family system. Parental burnout involves emotional exhaustion, emotional distance from one's children, feeling fed up as a parent, and a sense of contrast with the previous ideal parental self (Roskam et al., 2018). It occurs when parents chronically lack the parenting resources needed to cope with their specific parenting-related stressors (Mikolajczak & Roskam, 2018). What resources parents can access vary: from adequate financial resources to emotional resources, including social support, a cooperative co-parent, and sufficient leisure time (Mikolajczak, Raes, et al., 2018). Parenting stressors can include anything from having multiple very young children, striving to be a perfect parent, or living with family disorganization (Mikolajczak, Raes, et al., 2018). The particularities of each family and each parent combine to form their specific parenting risks and resources (which can vary culturally; Mikolajczak & Roskam, 2018), and so by its very nature, parental burnout cannot be separated from its specific family context.

On a conceptual level, a network approach to parental burnout prioritizes the view that parental burnout emerges through the interactions of its constitutive elements (Schmittmann et al., 2013). This view dovetails with the leading theoretical framework explaining parental burnout, which specifically emphasizes that a variety of family-specific factors (categorized as either "risks" or "resources," as mentioned above) combine over time and eventually result in a full-blown episode of parental burnout (Mikolajczak & Roskam, 2018). Therefore, a network model suits this theoretical framework of parental burnout exceptionally well, since network theory similarly posits that parental burnout arises through the interactions of its constitutive components (e.g., symptoms, family context, and external support). Moreover, most current research assesses parental burnout as a unified latent construct, which loses valuable information by collapsing the different features of parental burnout into one and ignoring any possibility that these distinct features could interact with the family system in different ways. Interestingly, practitioners know from experience that the different components of parental burnout interact with one another (e.g., the frequency of experiencing emotional exhaustion may increase the frequency of emotional distance, and vice versa; the frequency of feeling inefficient as a parent may foster emotional distance, and vice versa).

In conclusion, a network approach to parental burnout focuses on the features of parental burnout as embedded within a specific family system and encourages dynamic investigations using intensive time-series data. Reducing parental burnout to one isolated syndrome—as done by most current analyses—does not fit the theoretical framework of parental burnout, largely ignores the specific family system, and neglects how parental burnout features interact with contextual family factors. We therefore urge researchers to adopt a network approach to parental burnout, particularly when their research question involves the family context or the specific influences of the different features of parental burnout. We believe that this perspective can lead to more nuanced insight regarding how parental burnout affects the entire family, with all its contextual specificities.

Section 2.

Proof of concept & delving into methods

Chapter 4.

Proof of concept: A cross-sectional network approach to parental burnout

---

Background: The use of network analyses in psychology has increasingly gained traction in the last few years. A network perspective views psychological constructs as dynamic systems of interacting elements.

Objective: We present the first study to apply network analyses to examine how the hallmark features of parental burnout — i.e., exhaustion related to the parental role, emotional distancing from children, and a sense of ineffectiveness in the parental role — interact with one another and with maladaptive behaviors related to the partner and the child(ren), when these variables are conceptualized as a network system.

Participants and setting: In a preregistered fashion, we reanalyzed the data from a French-speaking sample (N = 1551; previously published in Mikolajczak, Brianda, Avalosse, & Roskam, 2018), focusing on seven specific variables: the three hallmark parental burnout features, partner conflict, partner estrangement, neglectful behavior toward children, and violent behavior toward children.

Methods: We computed two types of network models, a graphical Gaussian model to examine network structure, potential communities, and influential nodes, and a directed acyclic graph to examine the probabilistic dependencies among the different variables.

Results: Both network models pointed to emotional distance as an especially potent mechanism in activating all other nodes.

Conclusions: These results suggest emotional distance as critical to the maintenance of the parental burnout network and a prime candidate for future interventions, while affirming that network analysis can successfully expose the structure and relationship of variables related to parental burnout and its consequences related to the partner and the child(ren).

Reference

Blanchard, M. A., Roskam, I., Mikolajczak, M., & Heeren, A. (2021). A network approach to parental burnout. *Child Abuse & Neglect, 111*, 104826. https://doi.org/10.1016/j.chiabu.2020.104826

4.1 Introduction

A growing number of fields, including ecology, biology, and physics, have within the last decade embraced a network perspective, in line with both a growing interest and an increasing computational capability to model the complexity of systems (Barabási, 2012). Psychology subfields, including personality (e.g., Cramer et al., 2012) and psychopathology research (for systematic reviews, see Contreras et al., 2019; Robinaugh, Hoekstra, et al., 2019), have also recently embraced a network perspective, viewing psychological constructs as dynamic systems of interacting elements. Instead of investigating a construct as if it were a unitary phenomenon measured by multiple variables, a network approach examines and visualizes how the variables themselves interact with one another (for a detailed and accessible introduction to network analysis, readers can refer to the reviews by Borsboom, 2017; Borsboom & Cramer, 2013; Contreras et al., 2019; McNally, 2016; Robinaugh, Hoekstra, et al., 2019).

Exploring the structure and topology of a network of psychological variables can thus grant insight into the organization and mechanisms of psychological constructs. For example, nodes that are central to a network (and so especially connected to other nodes) work to maintain the network system, since once activated they can quickly spread that activation (Borgatti, 2005; Costantini et al., 2015). Within network perspectives of psychopathology, central nodes have been theoretically (Borsboom & Cramer, 2013) and empirically (Elliott et al., 2020; Papini et al., 2020) linked with the prognosis of a disorder. Assessing network structure also allows researchers to investigate whether nodes cluster together into communities and whether specific nodes act as bridges to connect these potential communities (Jones et al., 2019). For psychopathology networks, understanding community structures and identifying potential bridge nodes is useful when examining comorbidity, but also more broadly to understand how specific symptoms might be crucial to the network system (McNally, 2016). Bringing a network perspective to psychology research thus allows researchers to identify key variables that work to galvanize other variables or that connect multiple constructs.

A psychological phenomenon that is optimal for a network perspective is parental burnout. Parental burnout occurs when parents chronically lack the parenting resources needed to cope with their specific parenting-related stressors and involves three hallmark features, namely: 1) an overwhelming exhaustion related to the parental role, 2) an emotional distancing from one's children, and 3) a sense of ineffectiveness and loss of accomplishment in the parental role (Mikolajczak & Roskam, 2018; Roskam et al., 2017). Although epidemiological studies have not yet been published, recent assessments estimate that 5% of parents experience at least two of these three hallmark features daily (Roskam et al., 2018). However, although the concept of parents experiencing burnout has previously been investigated with parents of ill children (e.g., Lindström et al., 2011), it has only recently been examined in general populations of parents (Mikolajczak et al., 2020; Roskam et al., 2017). Still, the notion of parental burnout has gained traction (e.g., Cheng et al., 2020; Sorkkila & Aunola, 2020) and increasingly widespread recognition (e.g., Griffith, 2022).

Parental burnout presents itself as a prime candidate for a network analysis for two important reasons (Blanchard & Heeren, 2020). First, uncertainty persists regarding how the three hallmark features of parental burnout interact together. Indeed, in most of the previous parental burnout studies that have been published so far, researchers have tallied items of the three features to achieve a sum-score as an index of parental burnout (Blanchard & Heeren, 2020). This is unfortunate as one may easily envision how these features might interact with one another (e.g., a sense of ineffectiveness and loss of accomplishment in one's parental role may foster emotional distancing from one's children, and *vice versa*; an overwhelming exhaustion related to one's parental role may cultivate a sense of ineffectiveness and loss of accomplishment in one's parental role, and *vice versa*). A critical step would thus be to elucidate the functional relations among these features.

Second, parental burnout can easily be conceptualized as a dynamic system of interacting variables, with the three hallmark features of parental burnout impacting the entire family system (Blanchard & Heeren, 2020). Indeed, parental burnout impacts more than just the affected parent. This is because one parent feeling burned out is linked with family-wide dysfunction, including within the couple's relationship (e.g., conflict and estrangement) but also including severe negative consequences for the children, such as neglect and violence (Mikolajczak, Brianda, et al., 2018). Research investigating the effects of parental burnout is still in its beginning stages, but what literature does exist supports that parental burnout is linked with harmful parenting behaviors (Hansotte et al., 2021; Mikolajczak et al., 2019). These findings are not surprising, as previous research has found that with an increase in parental stress comes an increase in abusive behaviors (Berger et al., 2011; Curtis et al., 2000). However, as of yet no studies have untangled how the main features of parental burnout impact each other and dysfunctional family functioning, and vice versa. We thus propose a network analysis approach, to investigate how the hallmark features of parental burnout interact with each other and with variables implicated in family dysfunction – namely, a couple's relationship with each other (including conflict and estrangement) and their behavior toward their children (including neglect and abuse).

In this project, we have three primary goals. First, we endeavor to test whether the hallmark features of parental burnout—i.e., an overwhelming exhaustion related to the parental role, an emotional distancing from the child(ren), and a sense of ineffectiveness and lack of accomplishment in the parental role—cohere as a network system. Relatedly, we aim to elucidate the pairwise connections between these features and maladaptive behaviors relating to the partner (i.e., marital conflict, partner estrangement) and the children (neglect and violence toward children). Of critical interest will be the computation of centrality metrics that disclose the potential most influential nodes that maintain the network system (Borgatti, 2005; Costantini et al., 2015).

Second, we will test whether these nodes cohere as a unitary network of interacting elements or whether they constitute distinct communities (or subnetworks) of nodes serving different functions. And, if we find evidence for distinct communities, we will investigate whether certain nodes function as "bridges," i.e., processes that connect or are shared by communities. These analyses will offer viable heuristics to identify especially potent nodes that may kindle mutually reinforcing interactions between the features of parental burnout and maladaptive behaviors toward the family.

Finally, we will use Bayesian methods to compute a directed acyclic graph (DAG) to estimate a directed, potentially causal model of the interplay among the key features of parental burnout and maladaptive behaviors toward the partner and children. A DAG is a directed network whereby each edge has an arrow tip on one end, signifying the direction of prediction (and possibly causation), or, more conservatively, the direction of probabilistic dependency (e.g., Heeren et al., 2020; McNally, 2016; McNally, Heeren, et al., 2017; Pearl et al., 2016; Scutari, 2010). These complementary approaches can provide both a novel perspective and potential new hypotheses about plausible causal connections that merit more rigorous follow-up research. All of these computations will be performed as a re-analysis of an existing dataset (Mikolajczak, Brianda, et al., 2018) that includes questionnaires relating to parental burnout, animosity between the parents, and child maltreatment.

4.2 Method

Preregistration

As this is an exploratory study using secondary data, we followed the guidelines of Weston, Ritchie, Rohrer and Przybylski (2019) and preregistered the analysis plan, methodology, and prior knowledge regarding the dataset at https://osf.io/32vhg/. De-identified data as well as R code are publicly available via the Open Science Framework and can be accessed at https://osf.io/dxtgz/. We did not deviate from our preregistration.

Participants

We reanalyzed data from a previous study (Mikolajczak, Brianda, et al., 2018), using only the data from individual participants who had completed full questionnaires regarding their own parental burnout, their own partner-related behaviors, and their own violent and neglectful behaviors toward their own children. This yielded a sample of 1551 French-speaking participants (75.8% women). Participants were recruited from the general community (i.e., unselected sample) via websites, social media, schools, and pediatricians inviting people to participate to an online survey. All participants were parents with at least one child living at home, and participants were rewarded with the opportunity to participate in a lottery. Further details about the data collection can be found in the original article (Mikolajczak, Brianda, et al., 2018). The initial study was approved by the institutional review board of the Psychological Sciences Research Institute (UCLouvain, Belgium) and conducted according to the Declaration of Helsinki. Participants' characteristics appear in Table 1.

Table 1. Mean, Standard Deviation (SD), and Range for Sample Characteristics

Demographic Variable	Mean	SD	Range
Age (Women)	38.23	7.28	23 – 63
Age (Men)	42.61	8.47	27 – 66
Number of Children	2.30	1.38	1 – 7
Age of Children	8.49	6.70	0 – 35

Measures

Parental burnout

We used the Parental Burnout Inventory (PBI), a 22-item self-report questionnaire designed to measure the severity of parental burnout (Roskam et al., 2017). It includes three subscales that map directly onto the three hallmark features of parental burnout: a) eight items measuring emotional exhaustion (e.g., "*I feel emotionally drained by my parental role*"*);* b) eight items measuring emotional distancing (e.g., "*I sometimes feel as though I am taking care of my children on autopilot*"*);* and c) six items measuring parental accomplishment and efficacy (e.g., "*I accomplish many worthwhile things as a parent*"*).* Participants rated each item on a 7-point Likert-type scale, from 0 (*never*) to 6 (*every day*). Items denoting the absence of the feature of interest (e.g., "*I look after my children's problems very effectively*"*)* were reverse scored prior to computing each subscale's total by summing the relevant item scores.

Scores on the PBI have excellent internal consistencies and good validity (Roskam et al., 2017). In the present dataset, the internal reliability of PBI was high, with a Cronbach's alpha of .93 for the global scale score (.94 for the emotional exhaustion, .88 for the emotional distancing, and .86 parental accomplishment and efficacy). Given our interest in distinguishing between the three features, we computed separate scores for the three subscales. For each subscale, higher scores indicate greater endorsement of that feature. The scores for the emotional exhaustion subscale could range from 0 to 48; from 0 to 48 for the emotional distance subscale; and from 0 to 36 for the inefficacy subscale.

Conflicts with partner

Conflicts were measured with two self-report items as reported in Mikolajczak, et al. (2018): "*How often do you quarrel with your partner?*" and "*How often do you quarrel with your partner in front of your child(ren)?*" Participants rated items on a 7-point Likert-type scale, from 1 (*never*) to 7 (*every day*). For each participant, we computed a sum score: scores could range from 2 to 14. In the present dataset, the internal reliability of this scale was high, with a Cronbach's alpha of .85.

Partner estrangement

Partner estrangement was assessed with a five-item self-report questionnaire, as reported in in Mikolajczak, Brianda, et al. (2018). These items include: "*I sometimes fantasize about someone other than my partner,*" "*I am sometimes unfaithful to my partner,*" "*We have separate rooms* *or I go to sleep elsewhere,"* *"I sometimes think of leaving my partner,"* and *"I threaten my partner with leaving."* Participants rated each item on 8-point Likert-type scale ranging from 1 (*never*) to 8 (*several times a day).* For each participant, we computed a sum score: scores could range from 5 to 40. In the present dataset, the internal reliability of this scale was adequate, with a Cronbach's alpha of .73.

Neglectful behaviors toward children

Neglectful behavior toward children was assessed via a 17-item parent-reported questionnaire (for details on how this questionnaire was derived, see Mikolajczak, Brianda, et al., 2018). Sample items include: *"I sometimes neglect my child when s/he is sad, frightened or distraught"* and *"I sometimes don't take my child to the doctor when I think it would be a good idea."* The parent rated each item on an 8-point Likert-type scale ranging from 1 (*never*) to 8 (*several times a day*). In the present dataset, the internal reliability of this scale was good, with a Cronbach's alpha of .82. Accordingly, we relied on a sum score, and scores could range from 17 to 136.

Violent behaviors toward children

Violent behavior toward children was assessed with a 15-item parent-reported questionnaire (for details on how this questionnaire was derived, see Mikolajczak, Brianda, et al., 2018). Sample items include: *"I sometimes hurt my child with a belt, a hairbrush, a stick or some other object"* and *"I sometimes tell my child that I will abandon him/her if s/he is not good."* Items were rated with the same scale as used for the neglectful behaviors scale. The internal reliability of this scale was good, with a Cronbach's alpha of .82. We computed a sum score for each participant, and scores could range from 15 to 120.

Network analyses

Data preparation

Following recent guidelines in the application of network analysis in clinical psychology (e.g., Epskamp & Fried, 2018), we applied the nonparanormal transformation via the R package *huge* (Jiang et al., 2019). This transformation ensures that the data are multivariate normal and thus meet relevant assumptions for use with the estimated network analyses (Epskamp, Waldorp, et al., 2018).

To ensure that none of our variables overlapped conceptually, we implemented a data-driven method to identify potentially redundant nodes among our seven variables (as reported in Bernstein, Curtiss, et al., 2019). We first confirmed that our correlation matrix was positive definite, reflecting that nodes were not linear combinations of other nodes). We then searched for potential pairs of redundant nodes; that is, nodes that are highly inter-correlated ($r > 0.50$) and that correlate to the same degree with other variables (i.e., > 75% of correlations with other variables did not significantly differ for a given pair). To do so, we implemented the Hittner method for comparing dependent correlations (Hittner et al., 2003) via the *goldbricker* function of the R package *networktools* (Jones, 2018). This method did not identify any redundant variables.

Graphical LASSO network

We used a graphical Gaussian model (GGM) to estimate the undirected network. In this network, edges signify conditional independence relationships between nodes, controlling for the effects of all other nodes (Epskamp & Fried, 2018). We present GGMs that were regularized via the graphical LASSO (Least Absolute Shrinkage and Selection Operator) algorithm, which served two primary functions (Friedman et al., 2011). First, it computed regularized partial correlations between pairs of nodes, thereby eliminating spurious associations (edges) attributable to the influence of other nodes in the network. Second, it shrunk trivially small associations to zero, thereby removing potentially "false positive" edges from the graph and producing a sparse graph comprising only the strongest edges. We used the R package *qgraph* (Epskamp et al., 2012), which automatically implements the graphical LASSO regularization in combination with the Extended Bayesian Information Criterion (EBIC) model selection (Foygel & Drton, 2011). This procedure estimates 100 models with varying degrees of sparsity; a final model is selected according to the lowest EBIC value, given a specific hyperparameter gamma (γ) which controls the trade-off between including false-positive edges and removing true edges. The hyperparameter γ is usually set between 0 (favoring a model with more edges) and 0.5 (favoring a simpler model with fewer edges). Following recommendations based on stimulation studies (for details, see Epskamp & Fried, 2018), we set γ to 0.5 to be confident that our edges are genuine. To estimate the stability of edges, we bootstrapped the confidence regions of the edge weights with 1,000 bootstrapped samples using the R package *bootnet* (Epskamp, Borsboom, et al., 2017).

To quantify the importance of each node in the resulting graphical LASSO network, we computed expected influence centrality (Robinaugh et al., 2016). The expected influence of a node is the sum of the edge weights incident on a given node, including positive and negative values. Higher expected influence values indicate greater centrality and so greater importance in the network. The plot depicts the raw expected influence values for each node.

Finally, we tested whether the nodes denoting the key features of parental burnout and those denoting maladaptive behaviors relating to the partner and the children cohere as one network or as multiple subnetworks ("communities"). Nodes within a community are more strongly interconnected than they are with nodes outside that community. Following prior network research (e.g., Heeren & McNally, 2018; Robinaugh et al., 2014), we implemented the spin glass algorithm (Reichardt & Bornholdt, 2006), a modularity-based community detection procedure suitable for uncovering the structure of relatively small networks with negative edge values (e.g., Traag & Bruggeman, 2009). We used the *spinglass.community* function ($\gamma = 1$, start temperature = 1, stop temperature = .01, cooling factor = .99, spins = 7) of the R package *igraph* (Csardi & Nepusz, 2006; for more information, see Reichardt & Bornholdt, 2006; Traag & Bruggeman, 2009; Z. Yang et al., 2016). Following previous studies (e.g., Everaert & Joormann, 2019; Heeren et al., 2020; Heeren, Jones, et al., 2018), we also identified important nodes that serve as bridges between communities by computing the bridge expected influence index via the *bridge* function of the R package *networktools* (Jones et al., 2018). Nodes with high bridge expected influence values are especially likely to activate nearby communities. Bridge expected influence is the sum of the edge weights connecting a given node to all nodes in the other community or communities (Jones et al., 2019). The plot depicts the raw bridge expected influence values for each node.

Directed acyclic graph (DAG)

We then computed a DAG to estimate the directed structure of the system (Pearl et al., 2016). We used a Bayesian hill-climbing algorithm, implemented via the R package *bnlearn* (Scutari, 2010; Scutari & Denis, 2015). As implemented by *bnlearn*, the bootstrap function computes the structural aspects of the network model by adding edges, removing them, and reversing their direction to ultimately optimize the goodness-of-fit target score (i.e., the Bayesian Information Criterion). This involves an iterative procedure of randomly restarting this process with different possible edges linking different node pairs, perturbing the system, and using 50 different random restarts to avoid local maxima. Following previous studies (Bernstein et al., 2017; Heeren et al., 2020; McNally, Heeren, et al., 2017), we performed 100 perturbations (i.e., attempts to insert, delete, or reverse an edge) for each restart. As this iterative procedure unfolds, the function returns the best fitting network based on this random restart/perturbation process.

To ensure the stability of the resultant DAG, we then bootstrapped 10,000 samples (with replacement), computed a network for each sample, and averaged across the resulting networks to produce a final network structure (e.g., Heeren et al., 2020; McNally, Heeren, et al., 2017), involving a two-step procedure. First, we determined how frequently a given edge appeared in the 10,000 bootstrapped networks. We then used the optimal cut-point method of Scutari and Nagarajan (2013) for retaining edges, which yields networks with both high sensitivity and high specificity. Second, we ascertained the direction of each surviving edge in the 10,000 bootstrapped networks. To do so, we followed Scutari and Nagarajan (2013)'s recommendations and only represented the direction of a surviving edge in the final network if this edge pointed from node X to node Y in at least 51% of the 10,000 bootstrapped networks. In summary, we first determined the structure of the network (i.e., whether an edge is present or not), and then determined the direction of each surviving edge.

4.3 Results

Descriptive information regarding the seven variables (before nonparanormal transformation), including mean, standard deviation, range, skewness, and kurtosis, can be found in Table 2.

Table 2. Mean, Standard Deviation (SD), Minimum (Min), Maximum (Max), Skewness, and Kurtosis of Each Variable (Before Nonparanormal Transformation)

Variable	Mean	SD	Min	Max	Skewness	Kurtosis
Exhaustion	14.97	11.39	0	48	0.87	-0.05
Distance	9.34	8.51	0	47	1.47	2.23
Inefficacy	8.23	6.90	0	36	1.24	1.40
Neglect	26.91	9.17	17	84	2.00	5.91
Violence	22.75	7.58	15	81	2.57	11.19
Partner Estrangement	7.81	3.76	5	35	2.50	8.38
Partner Conflict	5.38	2.50	2	14	0.97	0.35

Note. Exhaustion = Emotional exhaustion; Distance = Emotional distance; Inefficacy = Loss of parental accomplishment and efficacy; Neglect = Neglectful behaviors toward children; Violence = Violent behaviors toward children

Pearson product-moment correlations of nonparanormal-transformed variables are reported in Figure 5 (for a discussion of why we present Pearson correlations of transformed variables and not Spearman correlations of the raw data, see the supplementary materials).

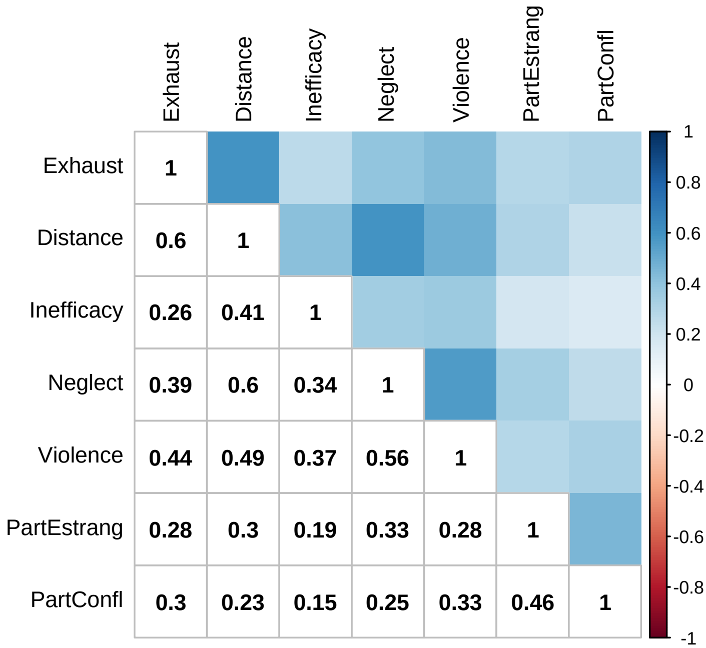


Figure 5. Pearson product-moment correlations between each nonparanormal-transformed variable

Note. Exhaust = Emotional exhaustion; Distance = Emotional distance; Inefficacy = Loss of parental accomplishment and efficacy; Neglect = Neglectful behaviors toward children; Violence = Violent behaviors toward children; PartEstrang = Partner estrangement; PartConfl = Conflicts with partner.

Graphical LASSO network

The graphical LASSO is represented in Figure 6. The edges represent regularized partial correlations between variables. All of the associations between variables are positive, as all of the edges are green (any negative edges would be represented in red). A few pairwise connections stand out. First, the largest edge weight is between emotional exhaustion and emotional distance ($r = .43$). Second, whereas emotional exhaustion has direct connections with violent behaviors toward children ($r = .15$) and conflict with the partner ($r = .13$), emotional distance is strongly connected to neglectful behaviors toward children ($r = .36$) and loss of parental accomplishment and efficacy to violent behaviors toward children ($r = .21$). To estimate the certainty and precision of the edge weights, we bootstrapped confidence intervals for each of the edge weights (Epskamp, Waldorp, et al., 2018) and also performed bootstrapped difference tests. Results support that the edges are stable, and that the strongest and weakest edges are significantly different from one another (see Figure S5 and Figure S6 in Appendix B or https://osf.io/q56sn/).

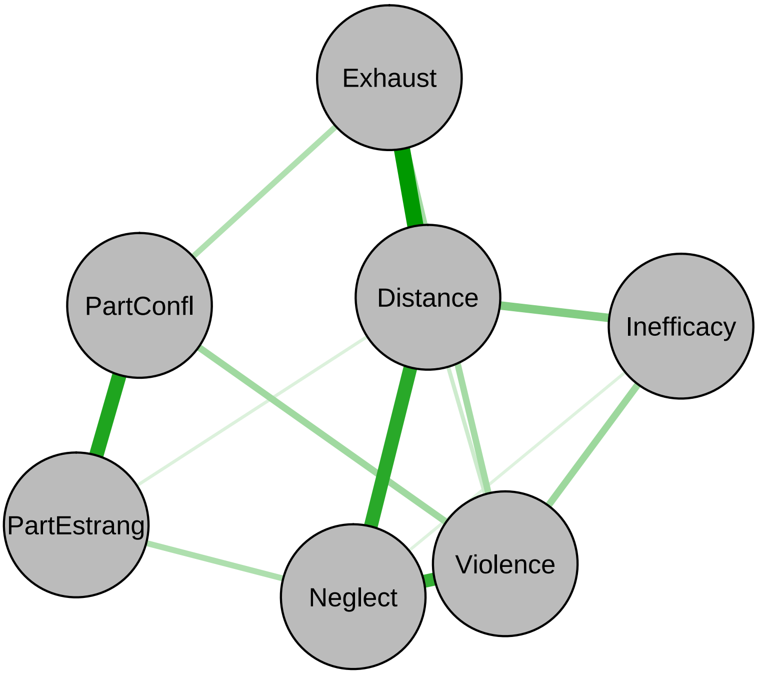


Figure 6. Graphical Gaussian Model Network Constructed via the Graphical LASSO

Note. An edge's thickness represents the magnitude of the association between the two relevant nodes (with the thickest edge representing a value of ….). Green lines represent positive regularized partial correlations. Exhaust = Emotional exhaustion; Distance = Emotional distance; Inefficacy = Loss of parental accomplishment and efficacy; Neglect = Neglectful behaviors toward children; Violence = Violent behaviors toward children; PartEstrang = Partner estrangement; PartConfl = Conflicts with partner.

Expected influence values are reported in Figure 7. Emotional distance, neglect towards children, and violence toward children had the highest expected influence values. To ensure the stability of these centrality estimates, we performed a person-dropping bootstrap procedure (Costenbader & Valente, 2003), which confirmed that these expected influence values are highly stable (see Figure S7 in Appendix B). We also performed a bootstrapped different test, which revealed that the most central nodes (i.e., emotional distance, neglect, and violence) have significantly higher expected influence estimates than less central nodes (including the partner-related variables and the remaining parental burnout features; see Figure S8 in Appendix B).

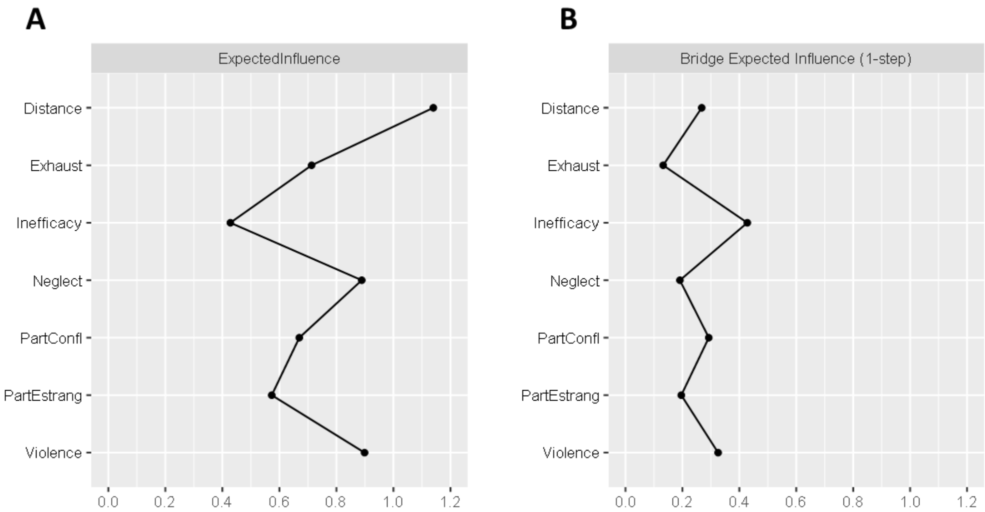


Figure 7. Expected Influence and Bridge Expected Influence Estimates of the Graphical LASSO

Note. Exhaust = Emotional exhaustion; Distance = Emotional distance; Inefficacy = Loss of parental accomplishment and efficacy; Neglect = Neglectful behaviors toward children; Violence = Violent behaviors toward children; PartEstrang = Partner estrangement; PartConfl = Conflicts with partner.

Lastly, the spin glass algorithm detected three communities of nodes. The first community included emotional distance, emotional exhaustion, neglect, and violence; a second community solely encapsulated loss of parental accomplishment and efficacy; and a third community included partner-related behaviors (conflict and estrangement). Figure 7 shows the bridge expected influence values for all nodes, revealing that inefficacy has an especially high bridge expected influence value. We performed a person-dropping bootstrap, which indicated that bridge expected influence values were reasonably stable (see Figure S7 in Appendix B). A bootstrapped different test confirmed that inefficacy had a significantly higher bridge expected value than the remaining variables (see Figure S9 in Appendix B).

Directed acyclic graph

Figure 8A and Figure 8B show directed acyclic graphs (DAGs) resulting from 10,000 bootstrapped samples. For both Figure 8A and Figure 8B, edges that are present in the graph were retained because their strength was greater than the optimal cut resulting from the method of Scutari and Nagarajan (2013). Figure 8A illustrates the importance of each edge to the overall network structure. Specifically, edge thickness reflects the change in the Bayesian Information Criterion (BIC; a relative measure of a model's goodness-of-fit) when that edge is removed from the network. Greater thickness thus signifies that an edge is more crucial to model fit (McNally, 2016). The edges most important to the network structure connect emotional distance to exhaustion (with a change in BIC of -338.29) and emotional distance to neglect (with a change in BIC of -167.37). The edges least important to the network structure, meanwhile, connect emotion exhaustion to partner conflict (with a change in BIC of -24.10) and emotional exhaustion to violence (with a change in BIC of -29.57). The change in BIC value for each edge can be seen in Table S2 in Appendix B.

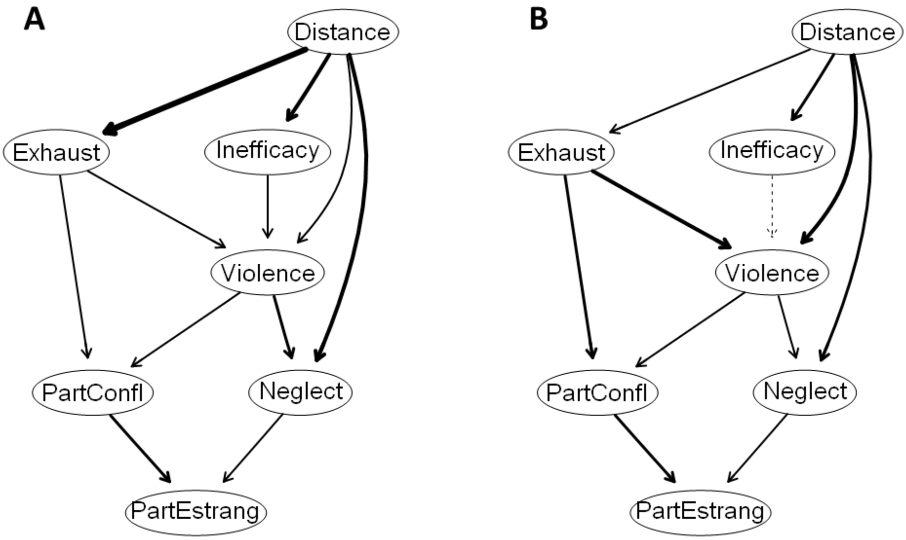


Figure 8. Directed Acyclic Graphs (DAGs)

Note. Panel A: Presence of edges: Edge thickness indicates the importance of that edge to the overall network structure, with greater thickness signifying that an edge is more crucial to model fit. Thickness reflects the change in the Bayesian Information Criterion of the model when that edge is removed. For this graph, solid lines represent that the present of an edge improves model fit (a dashed line would represent an edge whose presence worsens model fit). Panel B: Direction of edges: edge thickness indicates directional probability, or in what percentage of the fitted networks the edge went in that direction. Edge thickness is drawn proportionately, such that a thicker arrow indicates a higher directional probability. For this graph, a solid line represents that an edge was present in its current direction in at least 51% of the 10,000 bootstrapped networks, while a dotted line represents an edge present in its current direction in less than 51%. For both Panels A and B, exact edge weights can be found in Table S2 in Appendix B. Exhaust = Emotional exhaustion; Distance = Emotional distance; Inefficacy = Loss of parental accomplishment and efficacy; Neglect = Neglectful behaviors toward children; Violence = Violent behaviors toward children; PartEstrang = Partner estrangement; PartConfl = Conflicts with partner.

In Figure 8B, edges signify directional probabilities: an edge is thicker if it points from one node to another in a greater proportion of the bootstrapped networks. The thickest edges connect emotional distance to violence (.82; i.e., this edge was pointing in that direction in 8,200 of the 10,000 bootstrapped networks and in the other direction in 1,800 of the 10,000 bootstrapped networks) and emotional exhaustion to violence (.81). The thinnest edges connect inefficacy to violence (.51) and emotional distance to exhaustion (.53). The exact directional probability for each edge in Figure 8B can be found in Table S2 in Appendix B.

Structurally, emotional distance arises at the top of the DAG, directly influencing the other parental burnout variables (emotional exhaustion and inefficacy), which then directly influence violence toward children, which in turn directly influences partner conflict and neglect toward children, which both directly influence partner estrangement.

Preregistered additional analyses

Recent commentators have argued that differential variability—that is, the phenomenon that variables have drastically different variances—may distort conclusions about node centrality (e.g., Fried, 2015; Terluin et al., 2016). We confirmed that differential variability did not influence the interpretation of the centrality indices (see supplementary materials).

Following previous studies in the field (e.g., Mikolajczak, Raes, et al., 2018), we also tested whether our results differed depending on the gender of the parent or the number of children. Using the Network Comparison Test (van Borkulo et al., 2023), we found no evidence of a difference between the networks' overall connectivity or structure between mothers compared to fathers, or parents of one, two, or three and up children (see supplementary materials). Only one edge differed between parents of one child and parents of two children (see Table S3 in Appendix B): the network for parents with one child does not include an edge between neglect and partner estrangement, whereas the network for parents with two children does.

4.4 Discussion

In this study, we set out to elucidate how parental burnout features interact among themselves and with family-related variables, as well as whether any specific variables are especially central to the way parental burnout operates within a family system. To do so, we computed network analyses by using two distinct computational network approaches: a GGM and a DAG. Perhaps the most striking finding was the observation that GGM and DAG pointed to emotional distance as an especially potent variable at play. In both network models, our results imply that when emotional distance is activated, it drives the activation of the rest of the network. Additionally, both models indicated that the parental burnout variables are mainly associated with partner conflict and partner estrangement through the child maltreatment variables. This suggests that it is not parental burnout per se that is associated with marital issues, but the child maltreatment accompanying parental burnout that, in turn, drives marital conflicts. This is also apparent in the high centrality values of the nodes representing neglect and violence toward children, which confirm that these variables connect a lot of the other variables in the network.

In terms of community detection, it is interesting that loss of parental accomplishment and efficacy forms its own community, while the other two hallmark features of parental burnout, emotional distance and emotional exhaustion, cohere together in a community along with neglect and violence. Perhaps inefficacy is on its own because it relates to the skills-based parenting evaluation of loss of parental accomplishment and efficacy, which is qualitatively different from the more affective components of emotional exhaustion and emotional distance. Inefficacy demonstrating a high bridge expected influence value suggests that it bridges the different communities together and therefore is a possible target deserving a careful audit during prevention and treatment. However, seeing as it is the only community made up of one node, it is only logical for inefficacy to have a high bridge expected influence, since any of its connections to other nodes count as inter-community connections. It is also of note that emotional distance and emotional exhaustion are categorized into the same community as neglect and violence toward children. This implies that these four nodes are more strongly connected to each other than they are to nodes outside their community. This means that to reduce neglect and violence toward children, it would be important to target emotional distance and emotional exhaustion in treatment efforts.

Even though the three parental burnout scales are categorized into separate communities in this study, previous research supports that these three scales together adequately assess parental burnout, and factor analyses support the use of three subscales (Roskam et al., 2017). Although the current study uses the PBI (Roskam et al., 2017) which was adapted from the job-related Maslach Burnout inventory (Maslach et al., 2010), a new scale has since been developed using an inductive approach. This scale, the Parental Burnout Assessment (PBA; Roskam et al., 2018), encompasses four subscales: emotional exhaustion in one's parental role, emotional distancing from one's children, feelings of being fed up, and contrast with a previous period. This implies that parents lose pleasure and fulfillment as parents, instead of losing their parenting accomplishment and efficacy, thus bringing an emotional component to this dimension. A critical next step would thus be to conduct a network analysis using the PBA to test whether a similar community structure as found in this study which uses the three PBI subscales also appears with the four PBA scales.

Overall, these results suggest that emotional distance plays a key role in propagating activations to the other parental burnout nodes, as well as child maltreatment and partner conflict ones. This nuances the common conception of emotional exhaustion as the core dimension of burnout (Halbesleben & Bowler, 2007); it has been conceptualized as such partly because it is the common variable that ties disparate definitions of burnout together (e.g., Demerouti et al., 2001; Maslach et al., 2010; Pines & Aronson, 1988) and because several studies suggested that it was the first step of the burnout process (Lee & Ashforth, 1993; Leiter, 1993; Leiter & Maslach, 1988). Emotional exhaustion also appears influential for conceptions of burnout within the parenting sphere: a recently developed theoretical framework for parental burnout posits that parental burnout arises when there is a chronic lack of resources to meet parenting needs, leading to a depletion of energy and so emotional exhaustion (Mikolajczak & Roskam, 2018). Yet in our results, emotional distance was more central in the GGM and more highly placed in the DAG than emotional exhaustion. A potential reason for this could be that this data was collected at one time-point, and in Mikolajczak and Roskam's (2018) framework, emotional exhaustion was specifically posited as a key variable in the *development* of parental burnout. As our study relies on a cross-sectional design, one cannot exclude that emotional distance could be key in *maintaining*—while playing no role in instigating—the network system comprising parental burnout, marital conflicts, and the associated child maltreatment (see also the discussion below about the distinction between probabilistic dependence and temporal antecedence). There is however no other evidence regarding this hypothesis (either for or against), as so far, most research in parental burnout has viewed and analyzed parental burnout as a unitary construct, utilizing only the total parental burnout score. We thus call on future studies to use not only the global score but to also distinguish between the subscales. This will allow researchers to disentangle the role each of the subscales plays in the onset and maintenance of parental burnout, as well as to identify if the subscales differentially impact related variables, such as child maltreatment or partner conflict.

Distinguishing variables that differentially affect onset and maintenance of parental burnout also would be pivotal in identifying targets ripe for prevention and intervention. Emotional exhaustion and the associated chronic lack of resources, if implicated in the onset of parental burnout, would be important to target for prophylactic efforts. Emotional distance, if central in the maintenance of parental burnout as this study suggests, would be crucial in the development of treatments. Indeed, network theory suggests that "turning off" a highly central node, such as by down-regulating a variable through treatment, should decrease activation throughout the network (McNally, 2016). If this holds true for the models investigated in this study, a treatment that specifically decreases the amount of emotional distance a parent feels toward their child(ren) should also lead to downstream improvement not only in the other parental burnout symptoms but also in child-maltreatment and partner-conflict variables. An example of attempting to specifically reduce emotional distance could include mental imagery exercises such as imagery rescripting or positive imagery re-training (Holmes & Mathews, 2010), where parents would concretely visualize enjoying and valuing moments spent with children (for more clinical recommendations about mental imagery, see also Blackwell, 2019). Similarly, because the network perspective assumes that nodes having high bridge centrality may be more likely to propagate activation through the entire network when re-activated after treatment (e.g., Cramer et al., 2010), future research should examine whether ineffective parenting acts as a potential harbinger of relapse deserving careful audit during follow-up. Note that current treatment efforts for parental burnout already aim to do so via work with parents to improve child-rearing practices (Brianda, Roskam, Gross, et al., 2020). Before investigating treatment options in more depth, however, future research needs to further investigate the distinct roles of the different parental burnout subscales and confirm this distinction between onset and maintenance variables, using longitudinal data.

Indeed, one of our study's limitations is that it analyses cross-sectional data. The only insight into the direction of associations is from the DAG, which uses probabilistic methods to provide clues about the direction of probabilistic dependence between the variables. However, when considering DAGs, one must not confuse the direction of probabilistic dependence between variables with temporal antecedence (for a discussion, see Pearl, 2009). Indeed, DAGs encode the conditional probability distribution of the variables. In this way, DAGs can be decomposed as a product of the conditional distribution of each node given its parent nodes in the graph, thus rendering DAGs capable of indicating whether the presence of node A probabilistically implies node B more than vice versa (Heeren et al., 2020). However, it does not imply the temporal precedence of node B. Moreover, the DAGs, by definition, assume that the relationships between variables are acyclic (i.e., no feedback loop). On the other hand, it has been suggested that one may easily gauge the extent of the directional flip-flopping in the DAG by looking at the thickness of the edge (e.g., Heeren et al., 2020; McNally, Heeren, et al., 2017). Indeed, the thickness of the edge represents the probability that the edge points in the direction depicted but does not imply that the probability of the edge pointing in the other direction is zero (Heeren et al., 2020). Here, several edges were thin. For instance, the edge from inefficacy to violence pointed in its depicted direction in only 51% of the bootstrapped networks, thus indicating that it pointed in the other direction in 49% of the bootstrapped networks. Likewise, the edge from emotional distance to exhaustion pointed in its current direction in only 53% of the bootstrapped networks, thus indicating that it pointed in the other direction in 47% of the bootstrapped networks. The direction of prediction between these variables may thus go both ways. To confirm and further investigate the direction of association between variables, the next step is to move to intensive time-series data which will allow a direct investigation into the temporal relationship between variables, while also assessing possible feedback loops (Bringmann et al., 2013). This would also allow a micro-level assessment of how parental burnout variables interact with each other and with child-maltreatment and partner-conflict variables, thus allowing insight into the mechanisms of parental burnout.

We also did not measure any potentially relevant moderators, such as personality traits, attachment patterns, or psychopathology symptoms, all of which could affect how parental burnout impacts the children or the spouse. However, the number of parameters grows quickly with each additional node (Epskamp, Waldorp, et al., 2018), and so for this first network analysis on parental burnout, we chose to include the minimum number of nodes in the network to ensure that we had adequate power (for information on estimating power in network analyses, see the supplementary materials). We hope that future research, however, will investigate how these types of moderating variables could impact the way parental burnout interacts with behavior toward children and toward the spouse.

Some might find it a limitation that this study does not directly compare parents with severe parental burnout to those unaffected; however, we do not view this as a limitation as using an unselected sample allows for the generalizability of the results and protects against a restricted range issue. More crucially, however, there currently are no accepted clinical cut-offs for parental burnout (Roskam et al., 2017, 2018). And although all scales are skewed toward lower scores (since our sample is unselected), all variables nonetheless show a wide variability in scores. The only variables whose scores do not span the entire scale are neglect and abuse, which is to be expected since maximum scores on these scales represent frequent and regular abuse toward children (although there are also no clinical cut-offs for which scores on these scales "officially" cross the line into child abuse). Thus, a connection between (for example) emotional distance and neglect represents that higher scores of emotional distance are associated with an increased frequency of neglectful behaviors, but this increased frequency could be from "never" to "once a month."

Finally, another limitation is that parents self-reported all variables. Parents may thus have felt uncomfortable answering some of the items truthfully, particularly items relating to child maltreatment. On the other hand, the only validated method to measure parental burnout is currently with self-reported questionnaires. In any case, it is extremely challenging to collect datasets large enough for network analyses (e.g., minimum hundreds of participants). that include third-party assessments on child maltreatment—or, indeed, on parental burnout. Future research should nonetheless attempt to investigate other forms of data collection, such as through third party observations from other family members or teachers, or through structured interviews with psychologists or other trained professionals.

Conclusion

Emotional distance emerges as one of the most central variables in the network, driving the activation of the other parental burnout variables, as well as variables measuring child maltreatment and parental conflict. Neglect and violence towards are also highly central and connected to most of the network. In contrast, partner estrangement and partner conflicts are only associated with parental burnout variables through child maltreatment variables, suggesting that problems with the partner arise from child maltreatment. These results, stemming from cross-sectional data, suggest that emotional distance could play a crucial role in maintaining parental burnout and the associated child maltreatment, but further research is needed to confirm and further disentangle the mechanisms of parental burnout.

Chapter 5.

How to measure parental burnout fluctuations: Developing longitudinal items for parental burnout & its family context

Parental burnout is a growing subject of research, but thus far this research has not examined whether the features of parental burnout fluctuate over time. Moreover, parenting and parental burnout are inextricable from their family context. Therefore, a critical next step involves examining how parental burnout features temporally unfold and interact with the ever-changing family environment. To do so, we developed an 11-item experience sampling methodology (ESM) tool to measure self-reported parental burnout features (specifically emotional exhaustion, emotional distance, and feeling fed up), as well as partner relationship, children's behavior, behavior toward children, social support, and perceived resources. We conducted two two-week periods of ESM data collection (one with French-language ESM items; n = 9; one with English-language ESM items; n=23) and one eight-week data collection with the French-language ESM items (n=50). We collected the ESM data using formr, an open-source platform, and we provide open access to all materials (including a formr template, allowing free use of the assessment tool), analysis code, and data: https://osf.io/s2yv5/. Participants' responses indicated sufficient within-person variability (assessed via intraclass correlation) and support for convergent and discriminant validity (assessed by correlating aggregated ESM responses with retrospective questionnaire scores on parental burnout, depression, anxiety, and stress). Lastly, we found that the three parental burnout ESM items had high between-subject reliability and moderate within-subject reliability. Participating parents found the ESM survey easy to answer and not burdensome. Finally, we discuss how assessing parental burnout over time can help usher parental burnout research and treatment forward.

Reference

Blanchard, M. A., Revol, J., Hoebeke, Y., Roskam, I., Mikolajczak, M., & Heeren, A. (2023). On the temporal nature of parental burnout: Development of an experience sampling methodology (ESM) tool to assess parental burnout and its related ever-changing family context. *Collabra: Psychology, 9*(1), 74614. https://doi.org/10.1525/collabra.74614

5.1 Introduction

When parents face challenging parenting demands without sufficient parenting-related resources, they become stressed and exhausted—when this stress reaches an extreme state, this is called parental burnout. Parental burnout involves four main features (Roskam et al., 2018): emotional exhaustion relating to the parental role, emotional distance from the child(ren), feeling fed up with the parental role, and a sense of contrast with the parent they used to be. Following the current theoretical framework, parental burnout arises when parents chronically lack the resources to balance out their parental stressors (Mikolajczak & Roskam, 2018). These resources and stressors are specific to each parent's situation, and many of them vary over time, such as perceived partner support, leisure time, family organization, financial resources, and social support; all of these can change within weeks, days, or even hours (Mikolajczak, Raes, et al., 2018; Mikolajczak & Roskam, 2018).

Parents experiencing burnout within their parental role is a growing societal concern, as parental burnout is linked with suicidal ideation, addictions, child neglect and abuse, and marital conflict (Mikolajczak, Brianda, et al., 2018). Although research into parental burnout is still in its infancy, there are already initial findings of a temporal link between parental burnout and subsequent parental neglect and violence (Mikolajczak et al., 2019). These findings highlight the pressing need to better understand how parental burnout develops and how it impacts parenting behaviors, including neglect and violence, so as to develop effective prevention and treatment tools. Parental burnout is particularly concerning in the context of the COVID pandemic, which has changed experiences of daily life across the globe and has perturbed the equilibrium of many parents (Griffith, 2022).

The current method to measure parental burnout, the Parental Burnout Assessment (PBA; Roskam et al., 2018), asks parents to estimate the frequency with which they experience 23 specific feelings or behaviors on a scale ranging from "never" or "a few times a year or less" to "a few times a week" or "every day." The PBA has been adapted into and psychometrically validated for many languages, including Arabic (Gannagé et al., 2020), Portuguese (Matias et al., 2020), Chinese (Cheng et al., 2020), Farsi (Mousavi et al., 2020), Japanese (Furutani et al., 2020), Polish (Szczygieł et al., 2020), Romanian (Stănculescu et al., 2020), and Turkish (Arikan et al., 2020). As can be seen from the frequency responses for the PBA items, the PBA assesses the overall level of parental burnout at a specific timepoint but cannot capture day-to-day fluctuations.

Examining temporal fluctuations is an important next step. For one thing, there has been a growing move across psychology domains to consider the temporal dynamics of psychological constructs by repeatedly measuring the construct in everyday life (e.g., experience sampling methodology; ESM; Myin-Germeys et al., 2018; Trull & Ebner-Priemer, 2020). For most psychological constructs that are relatively stable across time, their features still fluctuate. This is for instance the case for depression. Although it is relatively constant across time, its actual symptoms, and particularly their interrelations, fluctuate on (at least) a weekly basis (Bringmann et al., 2015). Examining temporal fluctuations via ESM can grant insight into how a psychological construct develops or persists. For instance, researchers can use intensive longitudinal data to determine if specific variables influence one another in a feedback loop (Borsboom & Cramer, 2013). ESM data can also be used to grant clinical insight into a person's individual trajectory (although the clinical validity of such measures is still under investigation; Wichers et al., 2017). One such method is inspired by research on critical transitions in ecosystems (e.g., Hirota et al., 2011) and can be used to identify when someone is about to tip into a disordered state (e.g., developing or relapsing psychopathology) or return to an ordered one (e.g., recover; Wichers et al., 2016, 2019). And, last but not least, one can also capitalize on ESM data to generate idiographic (i.e., specific to a single person) network models that illustrate how an individual's behaviors and cognitions influence one another over time. Idiographic networks can thus yield information about potential clinical pathways and central variables for a single person (Bastiaansen et al., 2020) and open up radically new vistas for personalized treatment options (Fisher et al., 2017).

Examining the temporal dynamics of parental burnout is therefore a promising and timely endeavor. On the one hand, there is already evidence that emotional exhaustion, at least, fluctuates on a daily basis (Gillis & Roskam, 2020). It is therefore likely that emotional distance and feeling fed up (which were not measured daily in Gillis & Roskam, 2020) also vary from day to day; the fourth feature of parental burnout (i.e., a sense of contrast with the previous parental self) is by definition likely more stable across time. On the other hand, parental burnout does not exist in isolation but always within a specific family context, since parenting is a subjective experience that can fluctuate with shifts in the parents' or children's behavior or with the ever-changing situational demands of the environment (e.g., Aldao, 2013). A critical next step in parental burnout research would thus be to examine how the fluctuations of parental burnout features interact with contextual factors, such as the parent's relationship with their partner, the children's behavior, or the resources available to parents. Examining how these different variables interact over time will shed light on how parental burnout develops and evolves (for further discussion, see Blanchard et al., 2021). In addition, generating intensive time-series data for individuals allows for personalized clinical applications, such as estimating idiographic models that can be used for quantitative individualized case conceptualization (Fisher & Bosley, 2020).

Despite the aforementioned promising perspective, there is currently no available tool to measure the features of parental burnout and related situational variables on a repeated and frequent basis. Developing ESM items markedly differs from creating a traditional (single time-point) questionnaire. When developing ESM items, it is crucial that the phrasing and content of items mirror how people actually think about and describe their own moment-to-moment experiences (Myin-Germeys et al., 2018; Varese et al., 2019; Vasconcelos e Sa et al., 2019). Concretely, this means using terminology the targeted population would themselves use and employing unambiguous language (Myin-Germeys et al., 2018; Varese et al., 2019). Since participants will be answering items frequently, the items should thus focus on experiences common to daily life and avoid extreme states (Myin-Germeys et al., 2018; Varese et al., 2019). Because statistical analyses of ESM data depend on sufficient variance in participants' responses over time (Vasconcelos e Sa et al., 2019), adequate item development also has statistical importance.

Beyond the items' content per se, all ESM surveys should also be extremely brief, taking a maximum of two to three minutes to complete, since participants will be taking time, over weeks or months, from their daily lives to answer these questions (Myin-Germeys et al., 2018; Varese et al., 2019). The questions should not feature too many repetitive items, since the participants are already answering the same items repeatedly and often find answering similar items on top of that to be frustrating and confusing; some researchers therefore suggest employing single-item measures for straightforward unidimensional constructs (e.g., Robins et al., 2001; Wanous et al., 1997).

These varied requirements for developing optimal ESM items have rendered the assessment of their validity and reliability remarkably difficult (for an accessible overview of test reliability and validity concepts, see Jhangiani et al., 2019). For example, test-retest reliability assumes that a test is reliable if responses are stable across time, but this goes against the purpose of ESM, which is to measure fluctuations; changes in responses across time are expected (Eisele et al., 2021; Varese et al., 2019). If using single-item measures as is common for ESM studies (Eisele et al., 2021), internal consistency reliability (i.e., whether items measuring the same content yield similar responses) cannot be estimated (Wanous et al., 1997). Traditional measures of validity, such as criterion validity, are also complicated, since ESM items might not strongly correlate with one-time questionnaires measuring the same construct (Varese et al., 2019). One-time questionnaires typically measure a construct from a static perspective, covering long periods of time (and therefore potentially including recall bias; Varese et al., 2019); measuring that same construct from a dynamic perspective and then aggregating the responses does not necessarily yield similar results (e.g., since state measurements could vary depending on the context; Hektner et al., 2007). This is reflected by how uncertainty still abounds regarding the relationship between state and trait measures, which can be divergent (as is the case with state versus trait anxiety; for discussion, see Balsamo et al., 2013; Heeren et al., 2018).

That said, there are still ways to assess an ESM item's reliability and validity. Multilevel reliability (which assesses both between and within-person reliability) can be examined if multiple items form a scale assessing the same overall construct (Cranford et al., 2006; P. Shrout & Lane, 2012). At a very basic level, researchers can evaluate the content validity of ESM reports by examining whether internal experiences make sense together—that is, similar states should positively correlate, whereas dissimilar states should negatively correlate (Eisele et al., 2021; Hektner et al., 2007). Researchers can also examine convergent and discriminant validity by correlating ESM measures with retrospective measures of the same construct and different constructs (Eisele et al., 2021). Since ESM items aim at capturing daily fluctuations, it is also necessary to ensure that item responses have sufficient within-subject variability. To do so, one can examine the intra-individual standard deviation and intraclass correlation coefficients (ICCs) of each item (Trull & Ebner-Priemer, 2020). Nonetheless, since assessing the reliability and validity of ESM measures is arduous, it is vital that the item-development phase yields items that are self-explanatory and unambiguous (i.e., have high face validity) while covering all aspects of the construct (i.e., have high content validity). The recommended method to develop ESM items with high face and content validity is to pilot potential items with members of the targeted community, collect qualitative feedback from these community members, and review item phrasings and content with experts (Myin-Germeys et al., 2018; Varese et al., 2019; Vasconcelos e Sa et al., 2019).

Following the above guidelines, we accordingly had three major goals. First, we sought to develop a brief set of ESM items (in French and English) to assess parental burnout and relevant contextual variables, adapted to daily measurement with high face and content validity. To meet the first goal, we carried out an intense item development process, which included adapting items from existing one-time questionnaires (which measure their construct over a window of anywhere from the past two weeks to the past few years) into comparable measures of state constructs, as well developing new items from scratch; we developed the items in French and then back-translated them into English. Since we did not find many resources on how to develop ideal ESM items, we also wanted to share all aspects and materials of this item development process. Second, we wanted to ensure that parents were comfortable using this tool (i.e., minimal burden and applicable to their daily experience). To assess this second goal, we asked the parents who participated in the different ESM data collections for qualitative feedback regarding their experience answering the ESM items and any suggestions they had to render future studies easier for participants. Third, we wanted to examine how the ESM items performed when administered to multiple samples over different lengths of time. To assess this third goal, we examined whether items yielded sufficient within-person variance (i.e., low ICCs), whether the parental burnout items had good reliability, and whether the ESM items had good convergent/divergent validity with retrospective questionnaires. If we met these three goals, we wanted to share the resulting ESM items, administration methods, and materials to further stimulate research on the dynamics of parental burnout and the family context and ease direct comparisons between different studies.

5.2 Method

We report our rationale for the technical details of this ESM questionnaire (e.g., why we chose single-item measures, a daily interval for ESM assessment, and a continuous response scale) in Appendix C.

Item development

Development procedure

We first decided on the desired categories for items. We naturally included the features of parental burnout, but we also included relevant contextual family variables that would likely vary daily and interact with the parental burnout variables. Since a previous (cross-sectional) network study has identified that partner conflict and estrangement, as well as child neglect and abuse, interact with parental burnout (Blanchard et al., 2021), we wanted to include variables that broadly represent partner relationships and behaviors toward children. However, in the present study, we sought to investigate how parental burnout variables interacted more generally with the family context. In addition, ESM items should not assess extreme states but instead focus on behaviors common in daily life to yield sufficient temporal variation (Myin-Germeys et al., 2018; Varese et al., 2019; Vasconcelos e Sa et al., 2019). Since the theoretical framework of parental burnout also emphasizes the role of parental resources, we also included a category for daily resources and support from their social circle.

We then sought (French-language or translations of) questionnaires that assessed these categories (Beck et al., 1998; Brianda et al., 2019; Fombonne et al., 1988; Mikolajczak, Brianda, et al., 2018; Mikolajczak & Roskam, 2018; Roskam et al., 2018; Spitzer et al., 2006; Zacchilli et al., 2009) and adapted relevant items (typically with high factor loadings) for daily use. We included both negatively and positively worded items, following the recommendation of Vasconcelos e Sa and colleagues (Vasconcelos e Sa et al., 2019), to ensure sufficient variability. We also aimed for items that were broadly applicable to all parents and children at least some days (e.g., not questions that were only relevant for parents of infants). In total, we adapted and created 70 items. The first, second, and last author of this paper intensively workshopped the phrasing and content of these items so that these accurately assessed the relevant categories and suited a daily assessment. We then sought preliminary feedback from parents ($n = 7$) and adapted the questions based on their suggestions. The complete list of item revisions and permutations can be found on the Open Science Framework (OSF): https://osf.io/46z8x/.

We then conducted a pilot phase of data collection with a small group of parents ($n = 5$; see Table 3 for more information), so that parents could experience and give feedback on answering items daily. In this way, we could gauge which items yielded insufficient variation in participant responses, and parents reported to what extent they judged items as clear, unambiguous, and easy to answer. Based on participant feedback about how repetitious the items were, as well as the reasoning explained above regarding single-item measures, we chose one item per category based on response variability, parent feedback, and expert opinion. We then had another round of workshopping and discussion regarding item phrasing. We then conducted a back-translation process to arrive at an English-language version of these ESM items. More details on this process are available in Appendix C.

Table 3. Information on the Different Data Collections

Sample	N	Mothers, *n* (%)	Single parents, *n* (%)[a]	Dates collected	# days	Briefing type	Median survey time	Compliance rate
Initial Pilot	5	3 (60%)	1 (20%)	Oct.-Nov. 2020	7	Text	--	--
French 2-week	9	8 (89%)	1 (11%)	Dec. 2020	14	Text	1.33 min	95%
English 2-week	23	22 (96%)	0	May-June 2021	14	Recorded video	1.39 min	87%
French 8-week	50	46 (92%)	4 (8%)	April-June 2021	56	Video call	1.32 min	94%

Note. [a] Single parent is defined here as someone parenting their children without a second co-parent (even if they are in a relationship with someone).

Item testing

Software: formr

A primary barrier to implementing ESM in research and clinical practice with up-to-date technology (e.g., electronic software, time-stamped entries) is the expense of the software (Varese et al., 2019). Indeed, software and applications designed explicitly for ESM purposes are often either expensive or suboptimal (e.g., restricted to android devices and thereby limiting who can participate). Therefore, we opted to use *formr*, a free and open-source software (Arslan et al., 2020); this also means we can share all materials. Since *formr* operates via webpages, it can be used on any device and does not require participants to download any software or application. In addition, many aspects are also customizable using R; we discuss more details of *formr* and how to use it for ESM studies in Appendix C. We share *formr* templates for our ESM studies (French-language version: https://osf.io/c492m/, and English-language version: https://osf.io/p95zs/). Templates include all steps of collecting the ESM data: the enrollment survey wherein participants choose their notification time and type of notification (text message or email), the demographic questionnaire, the ESM items, and the figures apprising participants of their progression through the study.

Testing procedure: Data collection (English & French versions)

We conducted two preliminary phases of data collection (one for the French version of the items, one for the English version) to examine the variance in participants' responses to ESM items over two weeks, as well as to examine the three parental burnout items as a parental burnout ESM scale. As part of another project, we also conducted a longer data collection (spanning 8 weeks) with 50 parents. Details on the different data collections can be found in Table 3 and demographic information in Table 4, while additional information about the different data collections are reported in Appendix C. All responses were collected using *formr* (Arslan et al., 2018). All parents provided written informed consent to share their anonymized data, and the project received the approval of the Biomedical Institutional Review Board of UCLouvain (approval date: 11 May 2020; protocol title: PBNET). We recruited parents that had at least one child living at home; we purposefully did not restrict our sample to parents with only young children (or only older children), since we aimed to assess the questionnaire's suitability for parents with children of varying ages. For the preliminary data collections, we recruited parents from among the authors' acquaintances, and via emails to relevant groups; for the eight-week data collection, we recruited parents using flyers, social media posts and emails to relevant groups. Of critical interest, we also sought in-depth qualitative feedback from the parents about the items and their experience as participants.

Table 4. Demographic Information

Demographic Variable	Mean	SD	Min	Max
Initial Pilot				
Age of parents	42.20	5.64	35	50
Number of children (living under same roof)	1.6	0.55	1	2
French-language Two-week Data Collection				
Age of parents	44.67	5.55	36	52
Number of children (living under same roof)	1.78	0.97	1	4
Parental Burnout Assessment (total score)	15.44	15.03	3	49
English-language Two-week Data Collection				
Age of parents	39.3	7.95	27	60
Number of children (living under same roof)	2.39	1.34	1	5
Age of children in years (living under same roof)	10.50	5.65	0.2[a]	21.8
Parental Burnout Assessment (total score)	22.90	16.20	2	61
DASS (total score)	16.70	12.80	0	48
French-language Eight-week Data Collection				
Age of parents	37.56	5.28	30	50
Number of children (living under same roof)	1.98	2.96	1	4
Age of children in years (living under same roof)	7.94	5.53	0.01[b]	21.70
Parental Burnout Assessment (total score)	37.84	25.81	8	131
Generalized Anxiety Disorder-7 Questionnaire	5.58	5.35	1	20
Beck Depression Inventory	11.20	10.22	0	52

Note. DASS = Depression Anxiety Stress Scales.

[a] This refers to a child that was 3 months old at the start of the data collection

[b] This refers to a child that was a few days old at the start of the data collection

French-language two-week sample

First, we administered the French-language versions of the 11 items (see Appendix C) for 9 French-speaking parents. On the first day, participants filled out a demographic questionnaire and the Parental Burnout Assessment (PBA; Roskam et al., 2018). Based on participant feedback from the pilot (see Appendix C), participants could also select in this initial questionnaire whether they received the notifications containing the survey link via text message or email. They were also invited to choose their favored time, between 6 PM to 9 PM, to receive the notification (i.e., they were asked to choose a time after most of their interactions with their children were over). They received one reminder notification two hours later, and the questionnaire expired after five hours. The phrasing of the items was specifically tailored to the participants: if they had one child, the item included the words “my child”; if they had multiple children, the question referred to “my children.” As French is a gendered language, the same principle applied for the gender markers of the participant and their partner, as well as if the participant was a single parent (questions referring to “partner” were then not shown). This tailoring of item phrasing was performed automatically through *formr* (for details, see the *formr* template: https://osf.io/c492m/, and specifically the survey entitled ‘ESM_PB’). The ESM items were shown in randomized order. Figure 9A illustrates how the ESM questionnaire appears on participants' phone screen. Since this data collection occurred during a semi-lockdown period in Belgium (due to COVID-19), we added a question to assess how much time parents had spent with their children that day. However, this could also be generally useful to assess if children spend a few days away from their parents (e.g., because they are with grandparents or on a fieldtrip, because the parents are away for a business trip).

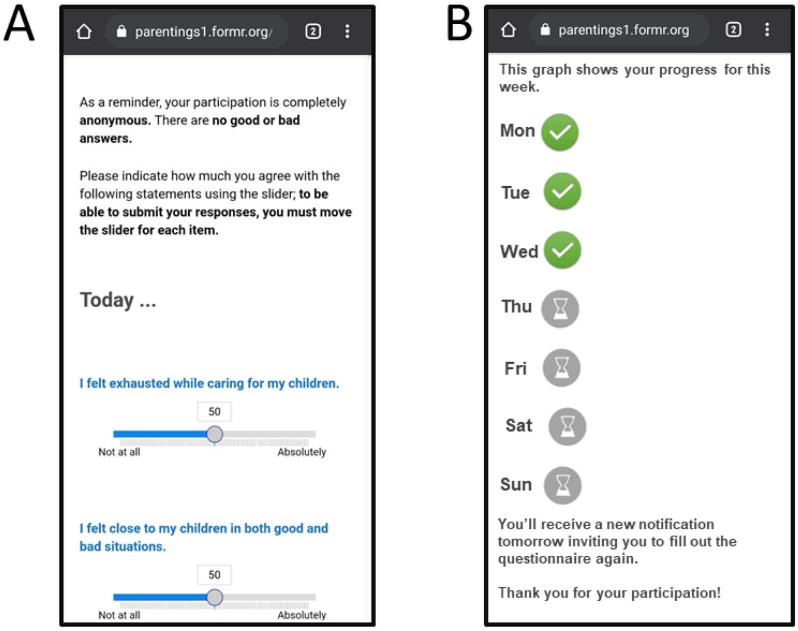


Figure 9. ESM Items and Daily Feedback (Participant View)

Note. A) Participant view of the ESM survey, accessed by a personalized webpage sent to them by an email or text message notification. An example version of this survey can also be viewed here: https://exempleparent.formr.org/. B) Completion feedback screen participant sees about how many ESM surveys they have answered so far during the week and how many they have left.

After completing the daily survey, the participants saw a specialized feedback message (specific to the day of the week and the percentage of time-points completed, to maintain participants' motivation and investment in participating in the study). Next, they saw a feedback graph showing them how many surveys they had filled out so far that week (see Figure 9B). On the fifteenth day, participants answered a follow-up feedback questionnaire, assessing their understanding of the items and soliciting any qualitative feedback about item phrasing, and their overall experience of the study.

English-language two-week sample

Next, we conducted a two-week data collection period with 23 parents using the English-language ESM items. The procedure was the same as for the French-speaking parents, except parents also filled out the Depression Anxiety Stress Scales (DASS; Lovibond & Lovibond, 1995). These parents had a high level of English (all but one were either native English speakers or used English regularly at their work) and lived in Belgium.

French-language eight-week sample

Since the previous data collections only spanned two weeks, we also included data collection from another project (https://osf.io/pshdn/) which spanned eight weeks. Before parents started this study, a researcher conducted an in-depth introductory briefing session with them to go over each item one-by-one and answer any questions; to this end, we created a *formr* survey with just the ESM questions, to introduce participants during the introductory briefing session to how the software operates, as well as for interested readers to take a look: https://exempleparent.formr.org/. This type of introduction and training meeting is known to enhance participant compliance (Trull & Ebner-Priemer, 2020).

We present preliminary results from this dataset (including ICCs, multilevel reliability for the three parental burnout items, and parent feedback) to examine the long-term use of these ESM items. Compared to the previous data collections, French-speaking parents additionally filled out the Generalized Anxiety Disorder-7 questionnaire (GAD; Spitzer et al., 2006) and the Beck Depression Inventory (BDI; Beck et al., 1998).

Statistical analyses

In addition to summarizing the qualitative feedback from parents, we also used quantitative methods to assess whether the final items were adequate for ESM use. First, we assessed whether the item responses yielded sufficient within-subject variability. To do so, we computed the ICC for each variable using intercept-only multilevel models, as is common within ESM studies (Gabriel et al., 2019). The ICC represents the proportion of the total variance due to between-person variance (Lüdtke & Trautwein, 2007; Snijders & Bosker, 1999); for example, an ICC of .40 indicates that 40% of the variance is due to between-person variation and 60% to within-person variation. Therefore, low ICC values indicate that an item possesses substantial within-subject variability; ICCs below .5 are usually accepted within ESM research, although some ESM researchers find typical ICC values to be between .2 and .4 (Eisele et al., 2021). We also report the intraindividual means and standard deviations, following the recommendation of Trull and Ebner-Priemer (2020).

Next, for samples in which we collected retrospective questionnaires about both parental burnout and distinct constructs (e.g., anxiety, depression, or stress), we examined the correlation table of all ESM items (aggregated within persons) with the questionnaire scores. Although within-person measures cannot be expected to match perfectly with between-person measures, we can still expect that the three parental burnout ESM responses should be more closely correlated with retrospective parental burnout measures than with retrospective measures of depression, anxiety, or stress, in line with expected convergent and discriminant validity (Eisele et al., 2021).

Lastly, although we could not assess the multilevel reliability of all variables, since they were all single-item measures, three items did represent three features of parental burnout. We therefore computed the multilevel reliability of a 'parental burnout scale' which included the Exhaustion, Emotional Distance (reverse-coded), and Fed Up variables, following the method of Cranford et al. (2006) and Shrout & Lane (2012). We used the *mlr* function in the R package psych version 2.0.9 (Revelle, 2020) to partition variance into between-individual (denoted R_{KR}) and within-individual reliability (denoted R_C). Low or moderate within-individual reliability is acceptable, since we expect substantial within-person variability, but high between-individual reliability is required for an ESM scale to be reliable.

We carried out all analyses in R 4.0.3 (R Core Team, 2021) and generated all plots with the *ggplot2* package, version 3.3.2 (Wickham, 2016). All anonymized data and R code can be found at https://osf.io/b35ec/.

5.3 Results

Finalized ESM items

The English translation of the finalized ESM items appears in Table 5.

Table 5. Parental Burnout and Family Context: ESM Items

Category	Item (English version)
Emotional Exhaustion	I felt exhausted while caring for my children.
Emotional Distance	I felt close to my children in both good and bad situations.
Feeling Fed Up	I felt overwhelmed caring for my children.
Partner Support	I received help from my partner with caring for my children.
Partner Conflict	I had some misunderstandings, tension, or arguments with my partner.
Kids' Behavior [a]	My children were difficult to manage.
Behavior towards Kids (+) [a]	I shared positive moments with my children.
Behavior towards Kids (-) [a]	I got angry with my children.
Resources	I lacked the means (for example, time, energy, material resources) to take care of my children.
Social Support	I received help from friends or family (other than my partner) with caring for my children.
Time with Kids	Around how many hours did you spend near your children today (outside of sleeping hours)?

Note. These ESM items were presented in random order on the same page, with "Today" at the top. ESM = experience sampling methodology.

[a] For these three item labels, we changed the exact label in subsequent publications that used this ESM tool, on the suggestion of a reviewer to make the label closer to the actual item content. As such, in Blanchard et al. (2023) and likely in subsequent publications, these items are labeled as "Difficult to Manage (Kids)" with the shorter version (e.g., for figures) as "DiffKids," "Positive Moments (Kids)" with the shorter version "PosMoKids," and "Angry (Kids)" with the shorter version "AngKids."

The finalized French ESM items, as well as figures showing the mean daily response with standard deviations, can be found in Appendix C.

French-language two-week data collection

The median time participants took to complete the ESM survey was 1.33 minutes. Overall, the rate of participant compliance was 95%. Each item appears to have sufficient within-person variation for analyses: there is a substantial intraindividual standard deviation for each item, and the ICCs of almost all items indicate that most item responses exhibit more within-person variation than between-person variation (i.e., ICCs < .50; see Table 6). Partner conflict was the only item with a somewhat larger ICC (.60), but it still contains a large intraindividual standard deviation. Using variance partitioning methods to estimate the multilevel reliability of the parental burnout scale (made up of the three parental burnout features: Exhaustion, Emotional Distance, and Feeling Fed Up), we found that the between-individual reliability was relatively high ($R_{KR} = 0.85$) while the within-individual reliability was, as expected, moderate ($R_C = 0.53$).

Table 6. Intra-individual Means, Standard Deviations, and ICCs for Each ESM Item

	French 2-Wk Sample			English 2-Wk Sample			French 8-Wk Sample		
ESM item	**iiMean**	**iiSD**	**ICC**	**iiMean**	**iiSD**	**ICC**	**iiMean**	**iiSD**	**ICC**
Emotional Exhaustion	20.41	21.27	0.26	33.31	19.01	0.51	29.2	24.31	0.23
Emotional Distance	87.57	11.02	0.32	85.02	9.09	0.62	76.11	15.74	0.38
Feeling Fed Up	14.2	17.34	0.15	24.53	17.76	0.40	22.52	21.89	0.16
Partner Support	52.79	24.25	0.40	71.96	18.5	0.57	62.51	25.15	0.42
Partner Conflict	30.49	22.14	0.60	21.78	19.1	0.31	18.97	22.87	0.19
Kids' Behavior	20.19	19.72	0.28	23.01	17.97	0.24	26.18	22.54	0.19
Behavior toward Kids (+)	82.35	14.98	0.37	85.41	11.98	0.36	76.82	16.38	0.38
Behavior toward Kids (-)	22.7	21.79	0.27	24.05	20.22	0.25	22.92	23.37	0.18
Resources	22.62	23.95	0.12	24.93	19.1	0.33	32.44	24.84	0.26
Social Support	27.75	27.91	0.28	13.76	13.72	0.46	25.86	32.02	0.14
Time Spent with Kids	6.09	2.60	0.47	7.43	2.78	0.39	8.01	3.62	0.24

Note. Wk = Week; ICC = intraclass correlations; ESM = experience sampling methodology; iiMean = intraindividual Mean; iiSD = intraindividual standard deviation.

English-language two-week data collection

The median time participants took to complete the daily ESM survey was 1.39 minutes. The compliance rate was 87% (although *formr* did have server issues during this period of data collection, leading many participants to miss one questionnaire). Although most ICCs were below .50, a few were higher (specifically, emotional exhaustion at .51, emotional distance at .62, partner support at .57; see Table 6). However, when we removed participants who showed no variation in responses (n = 2), and afterward removed participants who showed very little variation (n = 7), these ICCs did decrease (see Section 8 of the appendix)[1]. When examining the multilevel reliability of the three items making up the parental burnout scale, the between-individual reliability was high ($R_{KR} = 0.94$), while again the within-individual reliability was

[1] Note that these participants all had high levels of English and were of similar nationalities as the participants with greater response variation (e.g., Belgian, American, Canadian, etc.).

moderate ($R_C = 0.52$). Parental burnout ESM items correlated significantly with a retrospective questionnaire on parental burnout but did not do so with retrospective measures of depression, anxiety, or stress (see Figure 10).

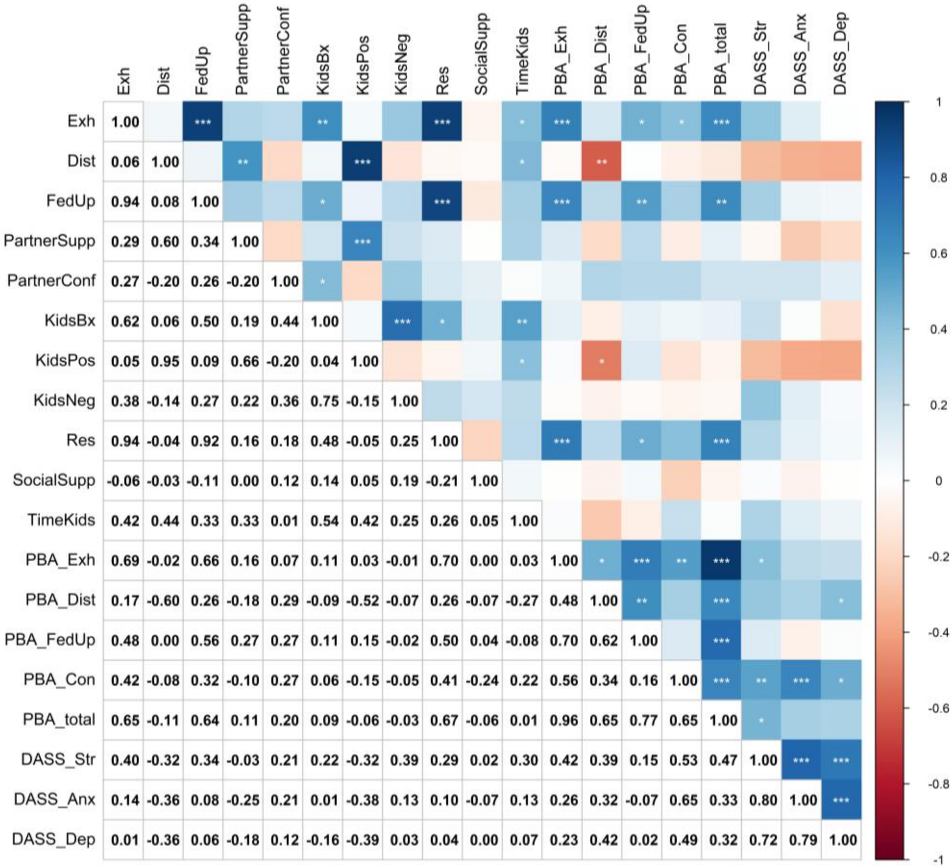


Figure 10. Correlation Plot of All Aggregated ESM Variables from English-language Two-week Data Collection

Note. The correlation table includes data averaged within person instead of all the data because of the temporal dependence between time-points. ESM = experience sampling methodology; Exh = emotional exhaustion; Dist = emotional distance; FedUp = feeling fed up; PartnerSupp = partner support; PartnerConf = partner conflict; KidsBx = children's behavior; KidsPos = positive behavior toward children; KidsNeg = negative behavior toward children; Res = resources; SocialSupp = social support; TimeKids = time spent with kids; PBA = Parental Burnout Assessment; DASS = Depression, Anxiety Stress Scales; Str = stress subscale; Anx = anxiety subscale; Dep = depression subscale. *p < .05. **p < .01. ***p < .001.

French-language eight-week data collection

Participants took a median of 1.32 minutes to answer the ESM items, and the overall compliance rate was 94%. All items showed sufficient intraindividual variability (all ICCs were below 0.5, and except for Partner Support all were below 0.4; see Table 6). Again, the between-individual reliability for the three parental burnout ESM items was high ($R_{KR} = 0.96$), whereas the within-individual reliability was moderate ($R_C = 0.58$). Parental burnout ESM items were more strongly correlated with the parallel retrospective subscale (e.g., the Exhaustion subscale with the PBA Emotional Exhaustion subscale) and the total PBA score than with the BDI or GAD (and in fact, only the Emotional Distance ESM item was significantly correlated with the GAD score; see Figure 11).

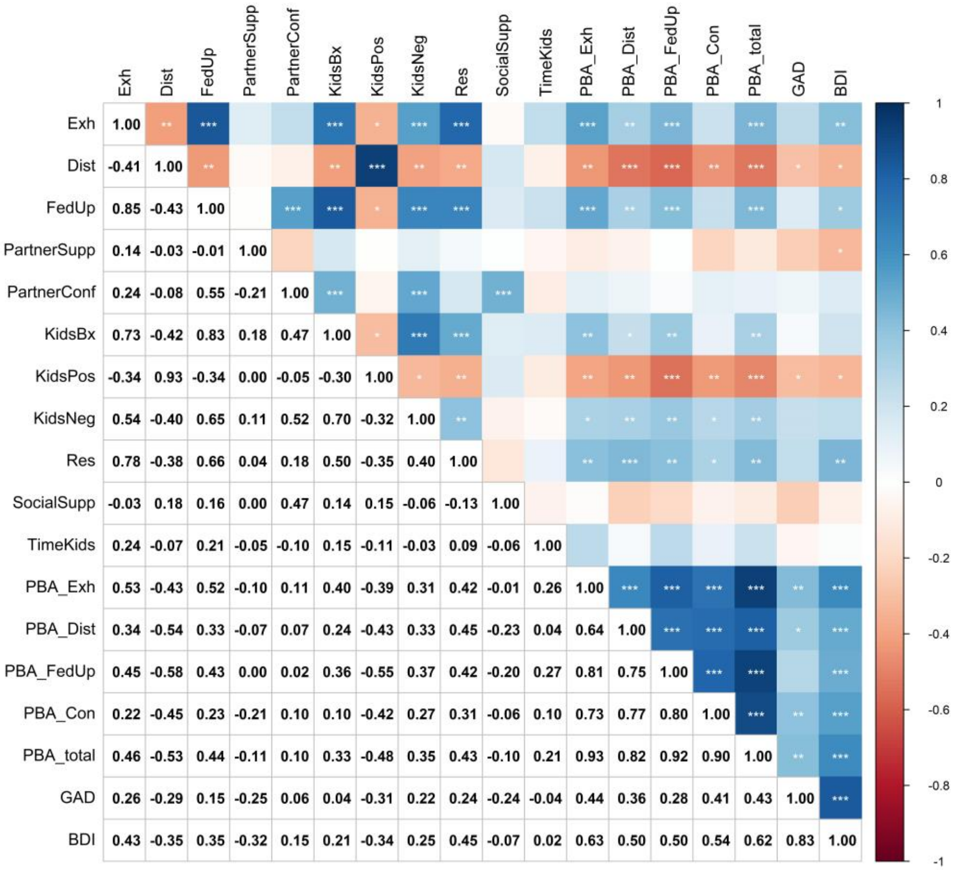


Figure 11. Correlation Plot of All Aggregated ESM Variables from French-language Eight-week Data Collection

Note. The correlation table includes data averaged within person instead of all the data because of the temporal dependence between time-points. ESM = experience sampling methodology; Exh = emotional exhaustion; Dist = emotional distance; FedUp = feeling fed up; PartnerSupp = partner support; PartnerConf = partner conflict; KidsBx = children's behavior; KidsPos = positive behavior toward children; KidsNeg = negative behavior toward children; Res = resources; SocialSupp = social support; TimeKids = time spent with kids; PBA = Parental Burnout Assessment; GAD = Generalized Anxiety Disorder questionnaire; BDI = Beck Depression Inventory.

**p* < .05. ***p* < .01. ****p* < .001.

Participant feedback

French-language two-week data collection

The participants overall thought the general interface and slider were easy to use, and parents generally reported appreciating participating in the ESM study (further parent feedback is reported in Appendix C). Parents rated whether they would recommend the studies to others on a scale from 1 (*not at all*) to 7 (*absolutely*); their mean response was 6.29.

English-language two-week data collection

After the two weeks of ESM questions, participants rated whether they thought the daily questions were clearly phrased on a scale from 1 (*not at all*) to 5 (*completely*) with a mean response of 4.73 (SD = .45) and whether they found the questions applicable to their daily life with a mean response of 4.36 (SD = 0.85). On whether they found the study to be a burden, participants answered with a mean response of 1.23 (SD = 0.61).

French-language eight-week data collection

After eight weeks of answering daily ESM questions, parents rated to what extent they found the study to be a burden, from 1 (Not at all) to 5 (Extremely); their mean response was 1.53 (SD = 0.71).

5.4 Discussion

Thus far, research on parental burnout has not examined whether parents experience fluctuations in the features of parental burnout. However, the theoretical framework conceptualizing parental burnout emphasizes the crucial role of the parent's time-varying family context—particularly, their parenting resources and stressors (Mikolajczak & Roskam, 2018). Therefore, a critical next step in parental burnout research should involve examining how the specific features of parental burnout fluctuate from day to day in interaction with the ever-changing family context. Investigating daily fluctuations involves ESM methods, with parents answering the same questions repeatedly in their daily lives; one can then use the resulting intensive longitudinal data to address research questions about the onset, maintenance, and interactions of parental burnout features within their family context. However, there is currently no existing tool to measure the features of parental burnout and related family variables on a repeated and frequent basis. This paper therefore had three complementary goals: first, to develop such an ESM tool; second, to ensure parents found it easy to use and not burdensome; and third, to assess the performance of the ESM tool.

For our first goal, we set out to develop ESM items in line with the theoretical framework of parental burnout: items assessing the main features of parental burnout and parenting behavior, but also contextual variables that are likely to change over time, such as partner relationship, child behavior, perceived social support, and perceived general resources. We solicited qualitative feedback regarding item content and phrasing from a total of 21 parents, as well as from parental burnout experts. This procedure winnowed an initial pool of 70 items to a final pool of 11 items, each with unambiguous phrasing that assessed the entire contents of its intended construct through a statement applicable to most parents daily.

Our second goal was to ensure this tool was easy to use and not burdensome for parents. Since analyses using ESM data require many time-points (and perform better with no or minimal missing data), and since ESM methods intrude into participants' daily lives by their very nature, this aspect is vital. Further, the target sample for these ESM items are parents, all of whom are busy and some of whom might be extremely exhausted or burned out. Therefore, these ESM items must present a minimal burden so that parents can comfortably answer the items over long periods of time. Across all samples, the ESM items took parents a median time of less than 2 minutes to complete, below the recommended three minutes for ESM questionnaires (Eisele et al., 2021; Varese et al., 2019). All three studies also had high compliance rate, with all parents answering over 85% of ESM prompts—in fact, the only sample with a compliance rate lower than 90%, the English-language sample, experienced issues with the *formr* server that caused most parents to miss at least one ESM prompt. In addition, parents reported in feedback questions that the ESM questionnaire was not a burden, that the questions were clearly phrased and applicable to their daily life, and that they would recommend the study to other potential participants.

Our third goal was to assess the performance of these ESM items across three different samples, two languages, and two different lengths of time. Across these different samples, we examined whether the ESM items yielded sufficient within-person variability, which is necessary to investigate how variables fluctuate in time, the purpose of ESM (Varese et al., 2019). In the eight-week data collection, we found that the ESM items do yield sufficient within-person variability, as all ICCs were below 0.5. The French-language two-week data collection had one ESM variable with an ICC over 0.5 (e.g., showing slightly more between-subject than within-subject variation), and the English-language two-week data collection had a few more. However, these data collections spanned a shorter amount of time (and so allowed less varied experiences than over eight weeks). In addition, multiple of these items had substantial intra-individual standard deviations, and a few high ICCs seemed related to a few parents in particular answering exactly the same over two weeks. This might have had to do with only the French-language data collection having a face-to-face (video-call) briefing to explain the study to participants, test the items out together, and answer all questions. Therefore, we would recommend for future ESM studies with these items to include a face-to-face briefing (as recommended for ESM studies generally; Trull & Ebner-Priemer, 2020) and to include sufficient time-points to allow parents to experience varied aspects of parenting and family life. We conclude that overall, these items do show sufficient variation for use in ESM studies; this finding in itself has theoretical importance, since it demonstrates that parental burnout features fluctuate over time in an unselected sample of parents. This aligns with previous findings about emotional exhaustion (Gillis & Roskam, 2020), but is the first to extend these results and find that emotional distance and feeling fed up also fluctuate daily.

For the two samples where parents answered retrospective questionnaires about both parental burnout and other constructs (e.g., depression, anxiety, and stress), we also examined the correlations of the aggregated ESM variables with these questionnaires. The three parental burnout ESM items (Emotional Exhaustion, Emotional Distance, and Feeling Fed Up) were more strongly correlated with the parental burnout subscale than with other constructs, providing initial convergent and discriminant validity. In fact, in our English-language data collection, these parental burnout ESM items were significantly correlated with the PBA but not with the DASS subscales (see Figure 10). In the eight-week French-language data collection, meanwhile, there were significant correlations between the parental burnout ESM items and the BDI as well as with the PBA, but the correlations with the corresponding PBA subscale were always stronger (see Figure 11). In addition, the BDI includes items that are not specific to depression (e.g., weight loss, insomnia, irritability), which is not the case for the DASS (Lovibond & Lovibond, 1995). This might have been the reason for this difference between datasets. In addition, when examining the correlation plot of all aggregated ESM variables, the correlations are all in the expected directions based on item content and valence; this is a promising preliminary finding regarding content validity (Hektner et al., 2007). The last part of assessing our ESM items involved examining reliability. Although we developed each item as a single-item measure for its construct (making it difficult to assess reliability overall), the three items measuring different features of parental burnout (Exhaustion, Emotional Distance when reverse-coded, and Feeling Fed Up) could also be summed into an ESM scale measuring overall fluctuations in parental burnout. When combined into one parental burnout scale, these three items had good-to-excellent multilevel between-subjects reliability.

Overall, we therefore provide an ESM tool with good initial estimates of reliability and validity to encourage research into the dynamics of parental burnout and its family context and set the scene for future transfer into ideographical clinical research. We also share the materials we used to collect data (e.g., formr templates), in the hope that they might help other researchers interested in the dynamics of parenting, parental burnout, and the family context. Establishing shared ESM items allows for consistency in the definition and operationalization of items and concepts, as well as response scales, which is necessary to be able to compare ESM results and generalize findings (Mofsen et al., 2019; Myin-Germeys et al., 2018; Trull & Ebner-Priemer, 2020).

Limitations

There is the possibility that participants could show reactivity in response to seeing the ESM questions, based on comments from one parent saying she appreciated participating in the study as it allowed her to evaluate her emotional state when interacting with her children. However, since we value limiting the burden on parents, we do not think it would be feasible to implement potential solutions to reactivity (e.g., variable sampling frequencies; Nehrkorn-Bailey et al., 2018). In addition, since we are especially interested in how these items influence one another (and not necessarily in what specific level each construct is at each day), we do not necessarily view reactivity as a crucial issue. And, in any case, previous research involving parent daily diaries about parental warmth and conflict found that parents showed minimal reactivity over two months, and any small changes were mostly positive, in line with literature on self-monitoring (Reynolds et al., 2015). We are more concerned that the items might induce social pressure for parents or feelings of guilt (Meeussen & Van Laar, 2018). We thus formulated ESM items so that parents would not have to highly endorse sentences that went against social parenting norms (e.g., the item assessing emotional distance assesses its opposite, emotional closeness). A further limitation relates to the generalizability of samples: all samples include mostly mothers and relatively few fathers, and were limited to parents living in Belgium. Although the purpose of this project was to detail the creation of ESM items and initially test them, future research should also confirm that these items are appropriate for more diverse and heterogeneous populations.

Future directions

Being able to explore the temporal dynamics of parental burnout opens new horizons for parental burnout research. ESM research can clarify the basic understanding of how a psychological construct develops and evolves within its context (Myin-Germeys et al., 2018). This is especially relevant for parental burnout, since its context—the parent's family situation—is a key component within the theoretical framework of parental burnout (Mikolajczak & Roskam, 2018). There might also be differences in the fluctuations of parents with parental burnout and those without, and understanding these differences would provide insight into how parental burnout operates. For example, it is unclear at this stage whether these fluctuations vary depending on the level of burnout; it is possible that when parents are burned out, they experience fewer fluctuations of the parental burnout features. In addition, since previous research has shown associations between parental burnout on the one hand, and marital conflict and neglect and violence toward children on the other (Mikolajczak, Brianda, et al., 2018), investigating how parental burnout features interact with behavior toward the children and the partner can help clarify the exact pathways between parental burnout and family dysfunction. Overall, ESM research encourages dynamic questions investigating change, stability, catalysts of change, and temporal interactions; many analysis types exist to investigate these questions, including multilevel models, temporal networks, and dynamic systems modeling (Jordan et al., 2020; Myin-Germeys et al., 2018; Nehrkorn-Bailey et al., 2018).

Integrating ESM methods within clinical research has also been recommended: adding in-the-moment assessments of a patient's state as an outcome measure for clinical trials can improve measurement precision and lower recall bias (Mofsen et al., 2019). Possibilities also arise to directly integrate ESM methods with clinical work. One example involves generating an idiographic network for a patient, which would allow both the clinician and the parent to visualize how that parent's symptoms interact over time (Epskamp & Fried, 2018; for a discussion, see Blanchard & Heeren, 2020). Another example involves detecting early warning signals heralding a transition into (or out of) a state of parental burnout, using methods from research on critical transitions in ecosystems (Scheffer et al., 2009). For instance, previous research has demonstrated such warning signals for depression (van de Leemput et al., 2014; Wichers et al., 2016) and bipolar disorder (Bayani et al., 2017), and more transdiagnostic approaches have been proposed as well (Wichers et al., 2019). Although these individualized applications are not yet ready for clinical implementation, requiring further methodological and analytic refinements (Bastiaansen et al., 2020; Wichers et al., 2017), they set the scene for radically new ways to inform clinical work and future development in parental burnout research.

Conclusion

Parental burnout is a growing area of research, but this research has not yet considered the possibility of day-to-day fluctuations in its features. Only one study so far has examined the daily fluctuations of a feature of parental burnout: parental exhaustion. We provide initial evidence that not only emotional exhaustion, but also emotional distance and feeling fed up—three of the main features of parental burnout—fluctuate daily. This ushers new possibilities for research examining how parental burnout develops and persists within its specific family system. To allow consistency in future research on the temporal dynamics of parental burnout, we detail our methods in creating an ESM tool with 11 items measuring parental burnout and related family variables (parent behavior, child behavior, partner interactions, social support, and resources).

Chapter 6.

Current and optimal practices for temporal network analyses: A scoping review

---

Network analyses have become increasingly common within the field of psychology, and temporal network analyses in particular are quickly gaining traction, with many of the initial articles earning substantial interest. However, substantial heterogeneity exists within the study designs and methodology, rendering it difficult to form a comprehensive view of its application in psychology research. Since the field is quickly growing and since there have been many study-to-study variations in terms of choices made by researchers when collecting, processing, and analyzing data, we saw the need to audit this field and formulate a comprehensive view of current temporal network analyses. To systematically chart researchers' practices when conducting temporal network analyses, we reviewed articles conducting temporal network analyses on psychological variables (published until March 2021) in the framework of a scoping review. We identified 43 articles and present the detailed results of how researchers are currently conducting temporal network analyses. A commonality across results concerns the wide variety of data collection and analytical practices, along with a lack of consistency between articles about what is reported. We use these results, along with relevant literature from the fields of ecological momentary assessment and network analysis, to formulate recommendations on what type of data is suited for temporal network analyses as well as optimal methods to preprocess and analyze data. As the field is new, we also discuss key future steps to help usher the field's progress forward and offer a reporting checklist to help researchers navigate conducting and reporting temporal network analyses.

Reference

Blanchard, M. A., Contreras, A., Kalkan, R. B., & Heeren, A. (2022). Auditing the research practices and statistical analyses of the group-level temporal network approach to psychological constructs: A systematic scoping review. *Behavior Research Methods, 55*, 767-787. https://doi.org/10.3758/s13428-022-01839-y

6.1 Introduction

The study of complexity has gained traction within multiple fields over the last few decades (Barabási, 2012), and psychology has recently embraced approaches exploring complexity as well. One such example is the network approach, which graphically models psychological phenomena as dynamic systems of interacting variables: a network (for an overview of the methodological approach, see Borsboom et al., 2021). Most of the research applying a network approach within psychology has specifically focused on psychopathology, viewing mental disorders as arising from the interactions between symptoms instead of from one common cause (Borsboom & Cramer, 2013). This view of psychopathology as a system of symptoms that interact and exacerbate one another is termed the network theory of psychopathology, and it comes with several crucial assumptions: that a central, or well-connected, symptom plays a crucial role in maintaining the entire psychopathology network, for example, or that a network with more connections between symptoms indicates more severe pathology (Borsboom, 2017). These assumptions are still under investigation (for reviews, see Blanchard & Heeren, 2022; Bringmann et al., 2022; McNally, 2021), but researchers have embraced the crucial tenet that psychopathology should be investigated by examining the complex interactions between symptoms—even outside the context of network theory. Indeed, the network approach within psychology has also included variables other than symptoms (Jones et al., 2017) and constructs other than psychopathology (e.g., personality: Cramer et al., 2012; intelligence: Van Der Maas et al., 2017; attitudes: Dalege et al., 2016). However, most psychology research involving network analyses thus far has utilized cross-sectional networks, and therefore mainly investigate the interactions between variables at one timepoint. Most psychological constructs, however, evolve over time, and examining relationships between variables from one timepoint to the next is the next horizon in the network approach to psychology (Blanchard & Heeren, 2022; Bringmann et al., 2022).

Temporal network analyses are a specific type of network analysis estimated using intensive longitudinal data (e.g., data collected repeatedly over time), allowing researchers to visualize the temporal connections between psychological variables. These networks graphically model how variables interact and predict one another from one timepoint to the next (more so than just the past information from one variable alone, and so providing information about Granger causality; Jordan et al., 2020). Temporal networks thus have enormous potential to inform our theoretical insight into psychological constructs by investigating dynamic questions, such as how variables evolve or interact over time (Jordan et al., 2020; Wichers et al., 2017). For example, in a recent study on social media use and well-being, time spent on Facebook at one timepoint predicted negative affect, decreased self-esteem, feeling insecure and social comparisons at the next timepoint. In contrast, none predicted Facebook use, suggesting that the relationship between Facebook use and these variables is one-way and not bidirectional (Faelens et al., 2021). In another example examining eating disorder symptoms in at-risk populations, cognitions about fatness and fear of weight gain predicted many other symptoms, highlighting the central role of these variables in the development of eating disorders (Levinson et al., 2020). Both of these examples also illustrate the usefulness of temporal network analyses for clinical applications, by suggesting key variables or symptom inter-relations to target in interventions. Although most of the temporal network articles so far have focused on clinical populations or research questions, others have investigated questions involving general emotion dynamics (Bringmann et al., 2013; Elovainio et al., 2020; Martín-Brufau et al., 2020; Meng et al., 2020) or even personality (Lazarus et al., 2020; Pavani et al., 2017).

Temporal networks offer great potential to psychology, but they also require demanding data collection and novel analysis methods. Indeed, temporal networks are estimated on data with many timepoints such as from ecological momentary assessments (EMA)[1]: this data is typically collected in participants' daily lives, with participants answering short questions multiple times a day over several days or weeks (Myin-Germeys et al., 2018). Dynamic networks can also incorporate intensive longitudinal data that is not self-reported, such as from activity trackers (Nehrkorn-Bailey et al., 2018) or passive digital phenotyping (Lydon-Staley et al., 2019). The intensive longitudinal data is typically analyzed using multilevel vector autoregressive (mlVAR) models, which regress a variable at time *t* on that same variable at time *t-1* within a multilevel framework (to account for timepoints within subjects). However, other analyses to generate group-level temporal networks also exist. Examples include Bayesian multivariate multilevel models and group iterative multiple model estimation (GIMME), which builds a group network structure from subject-specific networks (for a clear overview of these different analysis methods, see Jordan et al., 2020; for an in-depth explanation, see Epskamp et al., 2018). For most temporal network models, one can estimate up to three networks, each visualizing a different type of relationship between nodes: contemporaneous (i.e., relations between nodes within one timepoint, controlling for temporal relationships; hypothesized to capture processes that occur faster than the timepoints are measured; Epskamp et al., 2018), temporal (i.e., relations between nodes from one timepoint to the next), and between-subjects (i.e., relations between means of different subjects). After generating networks, many researchers examine specific questions, such as which are the strongest edges (e.g., connections) between nodes, or which nodes are the most central (e.g., most strongly connected to other nodes, and therefore assumed to influence the entire network).

However, as the data collection and analysis methods are novel and constantly evolving, there is substantial heterogeneity in network study designs. This is also heightened by temporal network analyses being relatively new to psychology, with the methodology to compute temporal networks only recently introduced (Bringmann et al., 2013; Epskamp, Waldorp, et al., 2018). Some of these varied practices also pertain to difficulties relevant to other fields (e.g., which assumptions do researchers check), while others are specific to temporal network analysis (e.g., varying methods to estimate and visualize temporal networks). Different research

[1] Although researchers use a variety of terms for intensive longitudinal data collection methods, including experience sampling methodology (ESM) and ambulatory assessment (AA), as well as daily diaries for repeated daily measures, we relied upon the term ecological momentary assessment (EMA) throughout the article, as this term is considered to have a broad methodology including many longitudinally intensive assessment traditions (Trull & Ebner-Priemer, 2009).

questions may also lead to different research practices: researchers could estimate just one network and examine specific edges or nodes, or they could compare networks between two groups (sometimes visually, sometimes with statistical analyses), sometimes focusing on differences in edges or nodes, sometimes in more global network topology such as connectivity. These varied approaches also lead to heterogeneous results that can be difficult to compare or interpret, and the field overall lacks a comprehensive view.

Network analyses have become increasingly common in psychology, and temporal network analyses, in particular, are quickly gaining traction, with many of the initial articles earning substantial interest from psychology researchers. Since the field is so quickly growing, we saw the need to formulate a comprehensive view of current temporal network analyses. More specifically, we wanted to form a clearer understanding of what choices researchers made when collecting, processing, and analyzing data, with the goal of helping future researchers make informed decisions when conducting their own temporal network analyses. To be clear, temporal networks are a new and complex methodology, and the purpose of this review is to identify areas where the field as a whole can improve and to support coherence and consistency in how methods and analyses are reported. To systematically chart researchers' practices when conducting temporal network analyses, we reviewed articles conducting temporal network analyses on psychological variables (published until March 2021) in the framework of a scoping review. Our goal with this systematic audit of the field is to determine the common practices involved in conducting temporal network analyses, and thus identify trends and gaps (Paré et al., 2015). We also synthesize this data to provide guidelines and guidance to help future researchers using temporal network analyses in psychological science.

6.2 Methods

A scoping review maps the key concepts or practices in emerging fields, and identifies gaps in a research field (Peters et al., 2015). This is more suited to our goal of auditing the specific data collection and analysis practices for psychological studies conducting temporal network analyses than a systematic review, which strives to "answer a clinically meaningful question or provide evidence to inform practice" (Munn et al., 2018; for examples of scoping reviews auditing data collection and analysis practices, see Kjellberg et al., 2016; Zarin et al., 2017). We followed the PRISMA extension guidelines for scoping reviews (PRISMA-ScR; Tricco et al., 2018). Before beginning this scoping review, we preregistered our protocol, including our objectives, inclusion criteria, and methods on the Open Science Framework (OSF): https://osf.io/jwqmk/. Note that we slightly updated this registration (i.e., changes to the title and a few of the variables to be extracted). A complete list of the changes can be found here: https://osf.io/pgkr2/.

Search strategy

We searched in two psychology databases: Scopus and PsycInfo. The initial search took place in January 2020, with two updated searches in March 2020 and March 2021. The exact search strings used can be found in Figure 12. Although most of this research has been conducted within clinical psychology, we did not restrict the scoping review to a specific domain of psychological sciences to be as inclusive as possible and maximize generalizability from the audit. Figure 12 shows the exact search phrase for each database.

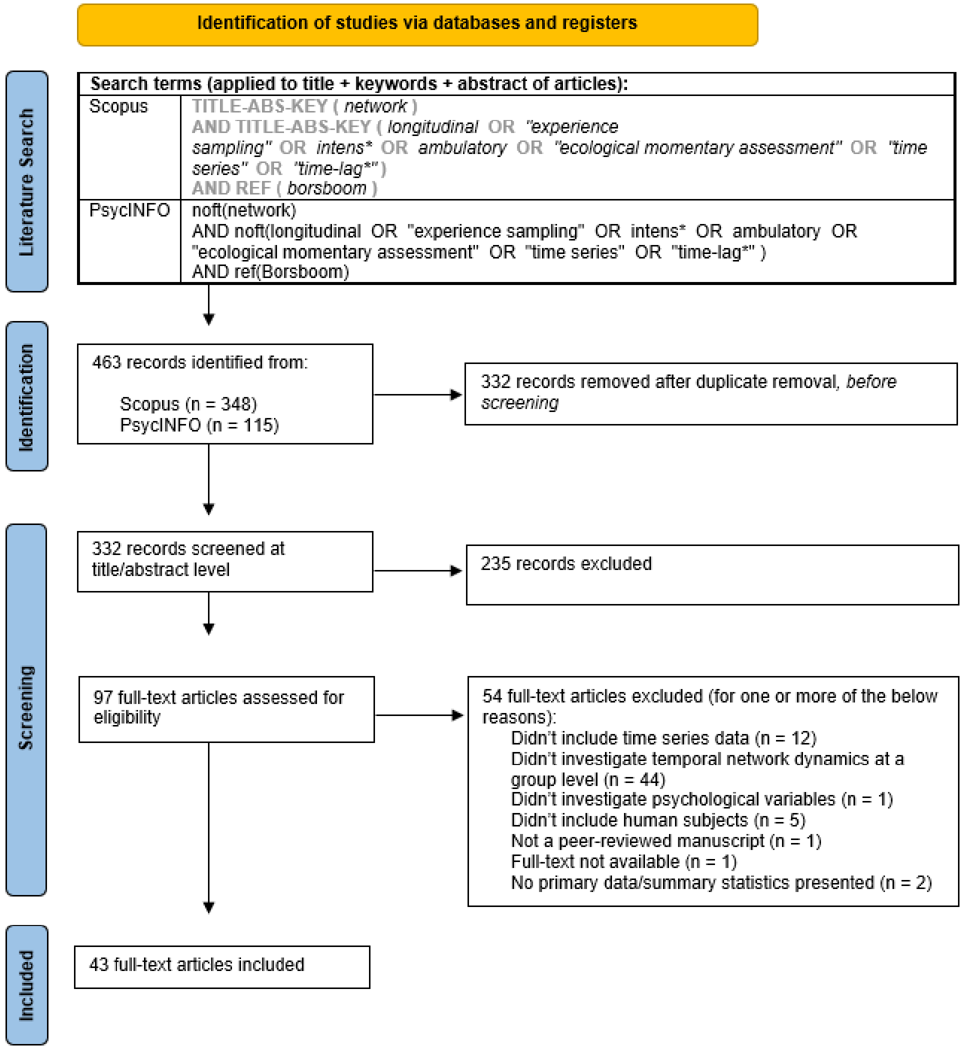


Figure 12. Literature Search Process

Note: We included a search term for their reference section of all articles: articles had to reference D. Borsboom, since he has authored the main initial conceptual papers linking network analysis with psychological research; we assume that any article discussing the psychological implications of temporal network analyses would cite at least one of Borsboom's initial articles. We included this search term to exclude articles with similar keywords that were not relevant to our search, such as research involving brain imaging, social networks, ecology, and so on. Figure template from (Page et al., 2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/.

Inclusion criteria

We limited our search to English-language studies that included data collected from human subjects and pertained to psychological variables. We limited our search to peer-reviewed articles to ensure that the methods and analyses reviewed in this scoping review were approved by experts. Eligible articles also had to include time-series data and investigate temporal network dynamics at a group level. Although there are also a number of articles examining idiographic (i.e., personalized) networks, we decided to limit our review to nomothetic (e.g., group-level networks), since the data collection, methods and analyses vary substantially.

Study selection process

Two reviewers independently inspected all titles and abstracts for eligibility according to the inclusion criteria. In a second phase, the criteria were reviewed in the full texts of the articles. In the case of inconsistencies, the two reviewers discussed with a third reviewer to resolve any remaining disagreements. We used CADIMA (Kohl et al., 2018) for the study selection process. Due to a technical issue with CADIMA, we only have an inter-rater agreement for the third wave of the study selection process (assessing 105 abstracts), with a kappa value of 0.61, which is considered "good." This translates to disagreements on 18 abstracts, all of which were discussed and resolved with a third reviewer. Most inconsistencies were between "unclear" and "no" for criteria, reflecting that information about the inclusion criteria was not always apparent in the abstract.

Data extraction (i.e., charting the data)

Data extraction included variables about metadata, sample characteristics, data collection, data preprocessing, estimation of data analysis, visualization of data analysis, robustness[2], and open science practices. We preregistered the exact variables for data extraction along with our protocol (https://osf.io/58swq/).[3] Three reviewers jointly developed the data extraction form to determine which exact variables to extract, and they piloted it before use on two randomly selected temporal network articles.

Two reviewers independently read all articles that met inclusion criteria and extracted the relevant information using a Google form (the contents of which can be found here:

[2] By robustness, we refer to "the stability of statistical inference under the variations of the accepted distribution models" (Shevlyakov & Vilchevski, 2002, p. *vii*); in other words, whether the model estimation yields comparable results with slight variations of the data (e.g., bootstrapped samples, case-dropping).

[3] The original preregistered list of variables for data extraction can be found here: https://osf.io/v73nb/. We updated our preregistration in March 2021 when adding a year onto our search end date, and we also added five additional questions to our data extraction form, which we found pertinent after reviewing the first batch of articles. These changes are detailed in blue in our updated preregistration (https://osf.io/pgkr2/), and have to do with authors computing multiple networks in their paper, the specific model they use, and how the authors handle missing data.

https://osf.io/mpj89/, and the link to the original form is in the preregistered protocol). The mean interrater agreement for categorical variables, calculated using the same procedure as (LópezNicolás et al., 2021) was a Cronbach's kappa of k = .67 (SD=.25), signifying substantial agreement. The range of agreement across variables was quite wide, however, with kappa's ranging from .08 to 1, reflecting how difficult it was to determine specific information for many of the variables investigated, in many cases because the articles themselves did not clearly describe all details of the data collection or analysis process. Only a few variables had a kappa below .5 (see Table S5), and we present more information regarding the Cronbach kappa's for the categorical variables in the supplementary materials. All disagreements between the two reviewers were extensively reviewed and discussed with a third reviewer.

Synthesis of results

The full dataset is available on OSF as a .csv file (https://osf.io/jsqrc/). For convenience, we also share a web-based interface that can be used to filter or group any variable of the full dataset in an easier fashion (see https://airtable.com/shrcC5MOBhHwEbwD8). In this review, we also summarize key information relating to the data collection, preprocessing, data analysis, open science, and robustness practices. We present graphs for key results that were created using R, adapted from the R code of López-Nicolás et al. (2021).

6.3 Results

The PRISMA flow diagram in Figure 12 illustrates the study selection process for this scoping review. For each extracted variable, readers can see the full list of possible answers for categorical variables (e.g., answered via multiple choice option by the two extractors) and the exact wording for free-response variable in the data extraction form: https://osf.io/mpj89/.

Study & sample characteristics

Eighty-eight percent of articles (n=38) were either part of a larger project or a reanalysis of already preexisting data. The sample size per network ranged immensely (see Table 7). Regarding the type of sample, 51% of articles had a clinical or health-related sample (n=22), 33% of articles had a non-clinical (i.e., healthy) sample (n=14), and 16% of articles included both clinical and non-clinical participants (n=7).

Table 7. Descriptive Information about Network Variables

Variable	*M*	Median	*SD*	Min	Max
Number of participants	119.6	66	210.7	5	1354
Total possible number of timepoints	60	61.2	21.3	7	140
Mean number of completed timepoints	47.9	47.5	22.8	13.8	114.8
Number of nodes	9	7	4.9	4	30

Data collection

Timescale

A crucial point for any study assessing variables over time is the timescale, and the articles varied enormously in the frequency of assessment, from ten times a day to once a year. However, the vast majority of articles (81%, n = 35) used a daily or multiple-times-a-day frequency of assessment, over a period of time ranging from five days to 56 days. We report the exact assessment schedule of each study in Table S6 in Appendix D.

Number of timepoints & missing timepoints

Another key aspect of data collection is the number of timepoints, which also varied widely (see Table 7). Sixty-seven percent of articles (n=29) also reported the mean number of completed timepoints (i.e., not including the timepoints that participants missed answering or that were excluded during analyses; see Appendix D). Finally, the scoping review also revealed that temporal network research takes a variety of actions in response to participants missing timepoints: 39.5% of articles (n=17) do not exclude participants based on missing timepoints, 18.6% of articles (n=8) do not specify any action, while 41.9% of articles (n=18) do exclude participants if they miss answering a certain percentage of timepoints. More specifically, 23.3% of articles (n=10) excluded participants if they did not complete at least 30% of timepoints, 4.7% (n=2) excluded participants not completing at least 40% of timepoints, 9.3% of articles (n=4) excluded participants without at least 50% of valid responses while 4.6% of articles (n=2) exclude participants not completing more than 50% of total timepoints.

Data collection method & design

We also examined details regarding how the researchers collected data in the reviewed articles: 77% of articles (n=33) relied on an electronic device, 9% (n=4) used a combination of electronic device and pen and pencil, and 14% (n=6) did not specify. The most frequent data collection designs (categorized following the descriptions provided by Myin-Germeys et al., 2018) were a fixed sampling design, with questionnaires completed at equal time intervals (33% of articles, n=14) and a pseudo-random sampling schedule, with questionnaires completed at random times within a specific time interval (44% of articles, n=19). Furthermore, to increase participants' compliance with the questionnaires, 49% of articles (n=21) compensated participants for their time. More information about the exact devices and platforms, as well as data collection designs and compensation details, can be found in the Appendix D or in the full dataset.

Data preprocessing

In 98% of studies, authors used R to conduct analyses[4]. Most studies reported that the variables they included in their networks were ordinal (Likert scale) variables (79% of studies) and that nodes were composed by using only one item (93% of studies). The number of nodes varied widely (see Table 7). For more information about the statistical software and packages the authors report using, as well as the type of variables assessed and the methods of defining nodes, see the Appendix D.

Assumptions

A key point before starting analyses is examining if there are violations of relevant assumptions, but only a minority of articles assessed whether their data was normal and stationarity (see Figure 13). Table 8 lists the main methods used to assess and correct normality and stationarity, while more detailed explanations and references to specific articles can be found in the Appendix D.

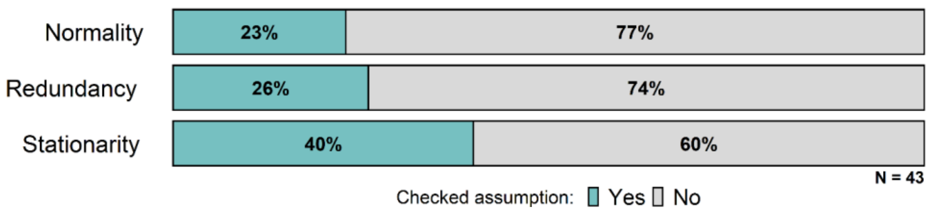


Figure 13. Assumptions

Another step authors could verify before beginning analyses is whether nodes overlapped in content. Only 11 articles report checking for potential redundancy in node content (see Figure 13); the methods used are listed in Table 8, while more information can be found in Appendix D.

Missing data

Regarding missing data, only five articles (12%) report any specific actions they took regarding missing data; the specific methods are listed in Table 8 (with more information and references in Appendix D). Since most of the models rely on mlVAR, most models therefore use the automatic action within the mlVAR package to use listwise deletion to deal with missing data (Epskamp et al., 2019); two articles explicitly mentioned this.

[4] Several authors reported using a combination of software types, but all but one used R for at least one portion of their analyses. For many other variables as well, authors sometimes utilized multiple of the answer options so that the total would amount to over 100%.

Table 8. Specific Methods Used to Assess Data or Network During Temporal Network Analyses

Action	Specific Methods
Checking normality (n=9)	• Visually inspecting histograms (n = 3) • Kolmogorov-Smirnov test (n = 3) • Shapiro-Wilk test (n = 2)
Correcting for normality (n=3)	• Log transformation (n=2) • Transformation via the normal quantile distribution (n=1)
Checking stationarity (n=17)	• Kwiatkowski-Phillips-Schmidt-Shin (KPSS) unit root test (n=11) • Dickey-Fuller test (n = 1) • Visual inspection of participants' responses over time (n = 2) • Add time variable (and compare model fit; n=2; or see whether variables depend on this time variable; n=2) • Theoretically assume violation (e.g., participants in treatment, n=1)
Correcting for stationarity (whether or not checked; n=18)	• Detrending the data (n = 10) • Including a linear trend (n = 8)
Checking redundancy (n=12)	• Choose nodes on a theoretical basis so they don't overlap (n=1) • Use model robust against high overlap in variables (e.g., PCMCI; n=1) • Look for multicollinearity or high correlations between variables (n=6) • Factor analysis (n=2) • Use the goldbricker function to identify which nodes likely measure the same underlying construct (n = 2)
Missing data (n=8)	• Taking specific action (n = 5): Multiple imputation; imputation; estimating data with Kalman filter; ensuring equal time intervals between timepoints (using cubic spline interpolation) • Justify that model is robust against missing data (n=3): because of Bayesian estimation, full-information maximum likelihood, or model itself
Network comparison (using formal statistical methods; n = 15)	• Using permutation tests (n = 10) • Using t-tests (n=3) • Using a Bayesian network model with 95% credible interval (n=1) • Correlating networks indices (to compare similarity of samples; n=1)
Robustness (n=6)	• Bootstrap confidence intervals (n=3; difficult with multilevel models) • Credible interval (uncertainty estimate with Bayesian estimation; n=1) • Case-dropping methods to assess stability (n=2)
Sensitivity (n=14)	Focus: • Variable transformation (e.g., standardization, to correct stationarity or normality violations; n=8) • Model choice or model specifics (n=4) • Node construction or selection (n=2) • Inclusion of specific participants (n=4)

Note. More details (including the full list of articles) for each method can be found in the supplementary materials or in the full dataset (see https://airtable.com/shrcC5MOBhHwEbwD8). The number of articles utilizing a specific method does not always sum up to the number of articles taking a specific action, since 1) not all authors specified how they conducted a specific action (e.g., some authors stated their data was not normal but did not explain how they assessed this) and 2) some authors did multiple types of actions (e.g., many authors did multiple sensitivity analyses). PCMCI = Peter and Clark Momentary Conditional Independence.

Data analysis: Estimation

Models

Within all articles, researchers estimated a group-level temporal network; 51% of articles (n=22) also included a group-level contemporaneous network, and 36% of articles (n=17) also included a between-subjects network. Almost all articles reported using a vector autoregressive (VAR) model to estimate the temporal network; the only other types of models reported were a GIMME model and a Peter and Clark Momentary Conditional Independence (PCMCI) model (see Figure 14A). The vast majority of VAR models were multilevel and estimated through sequential univariate estimations, although 2 articles used multivariate estimation (see Figure 14B). Only 2 articles did not use multilevel models (Lazarus et al., 2020; McCuish et al., 2021), but instead pooled data and then regularized it, with one additional article (de Vos et al., 2017) comparing a multilevel VAR model to a sparse VAR model (see Figure 14B).

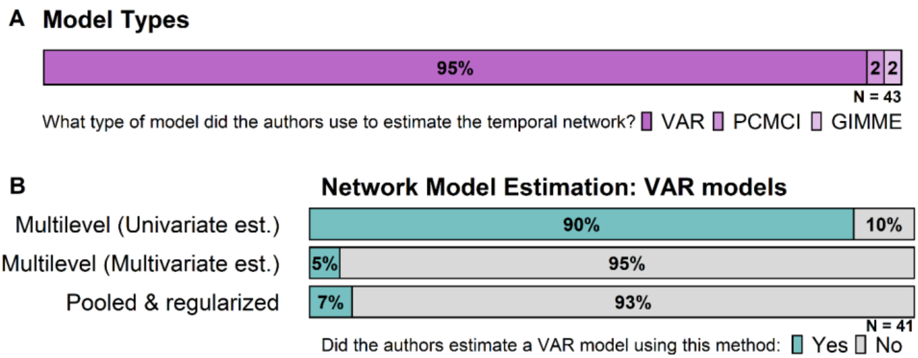


Figure 14. Models

Note. All percentages reported in the figure are rounded. VAR = Vector autoregressive; PCMCI = Peter and Clark Momentary Conditional Independence; GIMME = group iterative multiple model estimation; est. = estimation.

All models used a lag of 1 (i.e., from time t-1 to time t), although three articles did run alternate models with other lags (t-2 or t-3) and chose the model with the best fit, which was always t-1 (Bringmann et al., 2013, 2015; Kaiser & Laireiter, 2019).

Network comparison & centrality

More than half of the articles estimated at least two networks and compared them in some manner, including visually (see Figure 15A); around half of these compared the topology of networks, and almost a fourth compared centrality indices (see Figure 15B). Table 8 lists the specific methods used, and more information can be found in Appendix D.

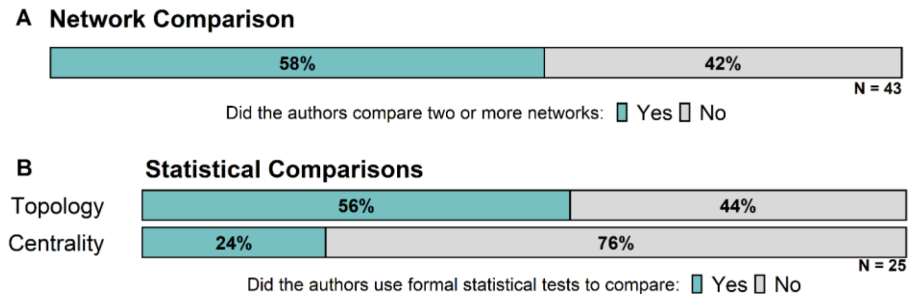


Figure 15. Network Comparisons

Open science practices

Only one article (Groen et al., 2020) preregistered their hypotheses and/or analyses for their network analyses, despite 61% of articles including specific hypotheses (see Figure 16A). Few articles openly (e.g., include a link within the article itself) share their code and fewer the data (see Figure 16B). None of the studies were replications of previous studies (i.e., investigating the same construct), but two studies included validation samples within the same article to replicate their own results (Bringmann et al., 2013; Huckins et al., 2020).

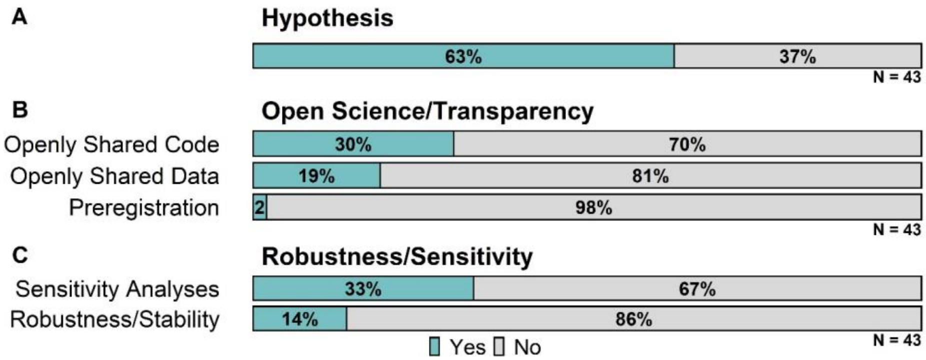


Figure 16. Open Science and Robustness Practices

Robustness and sensitivity

Only a few articles assessed the robustness of their edge or centrality estimates (see Figure 16C), using either Bayesian estimation or case-dropping bootstrap methods to do so (see Table 8, with more details in Table S8 of Appendix D).

One-third of the articles included sensitivity analyses around a specific analytical choice to observe its impact on the results (i.e., running the analysis with and without that choice; see Figure 16B). These articles performed sensitivity analyses on the impact of variable transformation, model or node specifics, and so on (see Table 8).

6.4 Discussion

Key findings

A striking observation throughout our audit of the methodology and analysis of temporal networks is that there is substantial heterogeneity in almost all aspects, from whether and how articles report node development and selection and preprocessing choices to how articles deal with missing data and assumption violations. This heterogeneity is reflected in the relatively modest interrater reliability ratings, both for article selection and for data extraction: there is no standard way of reporting information (about data collection or analyses), nor standard set of information to report. This is not unusual for a novel and growing field. It echoes our original reason to conduct a scoping review on this field: to understand what the current main practices in data collection and analyses for temporal network analyses are, so as to help interested researchers figure out what options they have when performing their own temporal network analyses.

Recommendations

We thus present Figure 17, which illustrates the main considerations involved when planning and conducting a temporal network analysis. The recommendations discussed are guided both by the literature, with relevant references cited, as well as by the results of this scoping review. We also discuss these recommendations in greater depth in the following sections. We hope readers interested in conducting temporal network analyses can use this information to help them make and justify decisions about collecting, preprocessing, and analyzing data. We also provide these considerations in the form of a reporting checklist, to help authors ensure they are reporting all relevant details in their manuscript and/or supplementary materials; this form can also be used as a preregistration checklist (https://osf.io/e6wp3/).

Formulating research questions

Before starting a temporal network analysis, researchers should ensure that they have thoroughly developed their research question(s), and that these research questions are suited to temporal network analysis. Temporal network analyses are ideal for investigating how dynamic variables interact with one another over time; researchers can also more specifically investigate which edges or nodes are especially important, or how the temporal networks of two different groups compare. It is crucial, though, that researchers ground the justifications

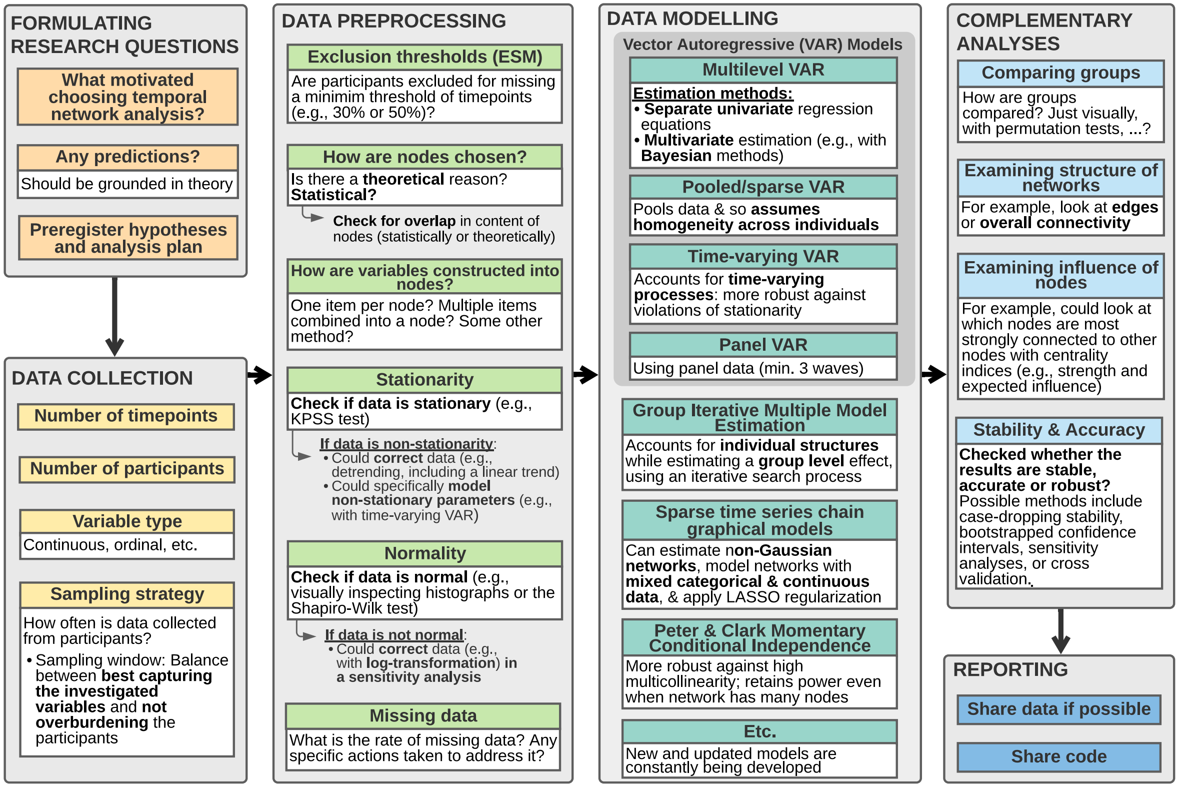


Figure 17. Temporal Network Analyses: Considerations

Note. The considerations in this figure were formulated using the extracted data from this scoping review as well as relevant literature (citations in the discussion section); more detailed explanations for all aspects of this figure can be found in the discussion section. This information is also presented as a reporting or preregistration checklist here: https://osf.io/e6wp3/. KPSS = Kwiatkowski-Phillips-Schmidt-Shin (KPSS) unit root test; LASSO = least absolute shrinkage and selection operator.

and predictions of their network analysis in theory. At the most basic level, this means that researchers should use temporal network analysis techniques because investigating temporal dynamics in a multivariate fashion will shed light on heretofore unanswered questions in their topic of interest.

This also means researchers should use theory to ground their choice of nodes. For example, Faelens et al. (2021) sought to understand the relationship between using social network sites (SNS) and well-being. Previous cross-sectional research had identified several indicators that played a role in the relationship between well-being and SNS use, including social comparison, self-esteem, and repetitive negative thinking. Moving to a temporal network perspective allowed these researchers to investigate the direction of relationships between all of these variables. Theory should also justify hypotheses. For example, Groen et al. (2020) investigated the temporal dynamics between depressive and anxiety disorders, and made hypotheses based on network theory that overlapping mental states would function as bridges connecting the two disorders (e.g., activating symptoms in the other disorder, leading to comorbidity).

Despite the great efforts of research to include theoretically relevant variables in temporal models, the vast majority so temporal network analyses thus far are data-driven: authors investigate a potential network structure by matching a statistical (network) model to data. Although some of the reviewed articles include specific hypotheses, none are derived from a testable and falsifiable theoretical network model. This would involve developing a theoretical network model of the psychological construct of interest: identifying the nodes in the network, as well the dynamic associations between nodes (e.g., how changes in one node might propagate throughout the network), and thereby developing testable (and importantly, falsifiable) hypotheses (Bringmann et al., 2022). Bringmann and colleagues (2022) note that there are few of these theoretical network models; examples include a network theory of how attitudes change (Dalege et al., 2016) and a computational model of panic disorder (Robinaugh et al., 2019). Testable hypotheses derived from theoretical network models can be examined through fitting statistical network models to data (potentially through "confirmatory" network models which test model structure and fit; Epskamp, Rhemtulla, et al., 2017), although also through other designs (e.g., experimental). Network analysis is still a new field, and using data-driven methods to explore plausible network structures is also a worthwhile pursuit, but hopefully there will be more theory-driven and confirmatory approaches to complement the current exploratory nature of temporal network analyses.

Preregistration

We urge researchers to preregister their planned preprocessing and analytical steps, as well as their data collection plans (or knowledge of preexisting data if conducting a reanalysis; Weston et al., 2019). This is especially important since so many of the preprocessing and analytical choices authors can make are arbitrary (e.g., excluding participants for missing a specific threshold of timepoints, constructing nodes, correcting assumption violations), and de

Vos et al. (2017) firmly demonstrate that arbitrary analysis choices (e.g., definitions of density/connectivity, data transformation) can influence results.

We also encourage researchers to preregister any hypotheses they develop. Over half of reviewed articles did include predictions of some type, but only one (Groen et al., 2020) preregistered its hypotheses and analysis plan. We note that many studies are exploratory and designed to generate hypotheses for future research; in this case, the authors should clearly state this in the manuscript itself (and in the potential preregistration). If researchers do have hypotheses, though, we would recommend that they preregister them.

Data collection

Number of timepoints

When collecting data, it is necessary to have enough timepoints per person. The median number of timepoints in the reviewed articles was 60. Within both the EMA literature (Palmier-Claus et al., 2011) and for most statistical models (Jordan et al., 2020), the absolute minimum number of timepoints suggested is 20, although this is typically the minimum number of timepoints the model needs to run (e.g., Epskamp et al., 2019). We recommend including many more timepoints, perhaps around 60 (the mean in this scoping review and another large EMA meta-analyses; e.g., Vachon et al., 2019). In general, the more timepoints, the better.

However, alternatives exist if a dataset has very few timepoints (e.g., less than 20 timepoints per person), such as using network models made for panel data (i.e., with at least 3-4 waves; Epskamp, 2020). Other alternatives include pooling data (such as by using pooled VAR to generate networks; Lazarus et al., 2020; McCuish et al., 2021); however, de Vos et al. (2017) caution that this method should only be used if there is little heterogeneity among participants' responses, since it assumes each participant is a replication of all others.

Number of participants

The median number of participants in reviewed articles is 66 per network. In general, including more participants indicates greater power, but actual power analyses are complicated with temporal network models. It is possible to conduct simulations and vary the number of subjects, nodes, and timepoints (this is integrated within the mlVAR package, for example), but the researcher should know *a priori* the network structure (Jordan et al., 2020). However, if researchers would like to explore simulating different numbers of participants and timepoints while taking into account temporal dependence (for analyses such as multilevel regression models, if not explicitly for temporal network analyses), Lafit et al. (2021) have recently developed a Shiny app and tutorial for doing so.

In general, though, researchers should know that power for longitudinal data used in multilevel models takes into account both number of timepoints and number of participants (Jordan et al., 2020) and is also dependent on the effect size, type of model, and so on

(Bringmann et al., 2022). In addition, number of nodes is relevant to this calculation: a network with more nodes includes more parameters to be estimated, thus requiring more timepoints and ideally more participants.

Variable type

Continuous variables allow more variability in participant responses. Aalbers and colleagues (2019), for example, use a scale of 0-100 to prevent a restricted range issue (i.e., a variable with minimal variance), which can distort conclusions made from network models (e.g., Terluin et al., 2016). We therefore recommend a scale with a large range, such as from 0-100, since other scale types (e.g., Likert) could yield limited variability in participant responses since they include fewer response options.

The most common variable type, used in 79% of reviewed articles, were ordinal variables (specifically, Likert scale). From a statistical standpoint, it is optimal if all variables have the same scale. Binary variables can be challenging to implement with current network models (although Klippel et al., 2018, conducted sensitivity analyses using multilevel logistic regressions).

Sampling strategy

Reviewed articles included sampling strategies ranging from 10 times per day over a few days to once a month for four years. In general, the sampling window should be selected to best capture the investigated variables while also not overburdening participants or leading to missing timepoints; finding the correct balance often takes piloting (Varese et al., 2019). The frequency of sampling (e.g., many times a day, twice a day) should be chosen to best capture the temporal variations of the variable of interest (Trull & Ebner-Priemer, 2020). For example, if a variable fluctuates intensely throughout the day, such as with emotional states, it would likely be better to include many measurements throughout the day. If, on the contrary, a variable fluctuates at a slower rate, such as sleep quality, it might make more sense to measure daily (Dejonckheere & Erbas, 2021). In addition, researchers should keep in mind that they might need to specify in their R code if they are measuring participants multiple times a day for some packages (e.g., mlVAR).

It is typically simpler for most of the models used in the reviewed articles (e.g., models such as multilevel VAR, that model time in a discrete fashion instead of continuously) to have all variables assessed on the same timescale (e.g., all assessed daily). However, this might not be possible if the variables of interest occur at different timescales (e.g., if investigating both sleep and emotions) or if the variables only occur occasionally (e.g., panic attacks); in these cases, continuous time dynamic models could be employed, as they grant greater flexibility with timescales (for an overview, see Bringmann et al., 2022).

Data preprocessing

Exclusion thresholds based on participant response rate

The EMA literature recommends that each participant have responses for at least 20 timepoints (Palmier-Claus et al., 2011). Twenty-one percent of reviewed articles excluded participants if they did not answer a minimum threshold of EMA items, most commonly 30% of items or less than 20 timepoints.

We recommend a flexible approach. For instance, if researchers collect 200 timepoints, a participant who only answered 30% of prompts still answered 60 timepoints. It could therefore make more sense to establish a minimum number of timepoints, such as only including in analyses participants who answered at least 20 timepoints (often considered the absolute minimum number of timepoints needed for analyses such as VAR models to run) or some other number of timepoints. It is also worth noting that participants with higher amounts of missing data, or even attrition, in longitudinal research may be those with higher symptom severity (e.g., Abdullah et al., 2021; Lamers et al., 2012), and so it can bias the results to exclude them.[1]

Node selection

Researchers should carefully consider and justify their choice of nodes – are they chosen for theoretical reasons? Statistical ones? Typically, researchers should choose which variables they include as nodes based on theoretical background, as discussed in the 'Formulating research questions' section with the example of Faelens et al. (2021). However, researchers might decide not to include specific nodes for statistical reasons, such as if they show no temporal fluctuations. If possible (e.g., if the researchers are collecting the data themselves), all variables pertinent to the research question should be included in the network without extraneous variables (Bringmann et al., 2022), and these variables should be assessed in a manner suitable to frequent assessment.

In general, an increase in the number of nodes in a network goes along with an increase in model complexity. In reviewed articles, the number of nodes ranges from four to 30. Researchers should check for overlap in the content of nodes, either statistically or theoretically. If high overlap is unavoidable for theoretical reasons (e.g., studying a construct which has highly overlapping features), researchers could use a model that allows many nodes and highly correlated variables (such as the PCMCI model; Runge et al., 2019) or a model that explicitly takes into account the overlap between nodes (e.g., a latent network model; Epskamp, Rhemtulla, et al., 2017).

Variable construction into nodes

[1] We are thankful to one of the reviewers of this study for pointing our attention to this issue.

Researchers should make sure to report how they constructed nodes from their variables. For example, did each item form a node? Were multiple items combined into a node, and how? Was some other method used?

Stationarity

Since stationarity is one of the main assumptions of most models with EMA data, researchers should check if their data is stationary (e.g., if the mean and variance of the EMA data remain unchanged; Jordan et al., 2020). One way to do so is to use the KPSS unit root test (Kwiatkowski et al., 1992), as 11 reviewed articles did. If data is non-stationary, researchers could correct this by detrending the data (e.g., Hoffart & Johnson, 2020) or including a linear trend (e.g., Bringmann et al., 2015). However, one may wonder about the influence of these types of corrections on the interpretability of the results (Jordan et al., 2020). Therefore, researchers could instead directly use a model that allows non-stationary data by specifically modeling non-linear time-varying parameters, such as time-varying VAR (Haslbeck et al., 2020). Perhaps researchers are investigating processes that change over time (e.g., during treatment, or a transition into or recovery from a mental disorder); in this case, researchers could use models that explicitly look for abrupt or gradual changes, although these models typically require many more timepoints (for an overview, see Bringmann et al., 2022).

Normality

Another crucial assumption for most temporal network models (e.g., VAR models; Epskamp, Waldorp, et al., 2018) is (multivariate) normality, and so researchers should check whether the assumption of normality is violated. They can do so by visually inspecting histograms of the residuals or with the Shapiro-Wilk test (Shapiro & Wilk, 1965).[2]

If data is non-normal, it is currently unknown exactly how this will bias temporal network results if using an estimation technique that assumes normality (Epskamp, 2020), especially since violations of normality can occur for many different reasons (Epskamp, Waldorp, et al., 2018). For example, if data is non-normal because it was measured on a different scale (e.g., Likert-type ordinal scale) but the underlying process is normal, it can be effective to transform the data back to normal (Epskamp, Waldorp, et al., 2018). For instance, among reviewed articles, this was carried out by de Vos et al. (2017), who transform all items using the normal quantile transformation, since these items had Likert-scale responses with skewed distributions; they also report differences in network density compared with non-transformed data (among other sensitivity analyses). However, data that is non-normal due to the process itself being non-normal (e.g., if the residuals are skewed) pose more of a problem (Epskamp, Waldorp, et al., 2018).

One possible action when faced with non-normal data is to transform the data, although how transforming time-series data affects temporal networks is, to the best of our knowledge, currently unknown. Among the reviewed articles, in addition to de Vos et al. (2017), only two

[2] Although the Komogorov-Smirnov test was also used by three reviewed articles, the Shapiro-Wilkes test is favored as less biased and more powerful (Razali & Wah, 2011; Steinskog et al., 2007)

additional articles performed a transformation for non-normal data. First, Aalbers et al. (2019) report the distributions of their variables, as well as the normality of within-person means; upon the request of a reviewer, they then log transformed items that had a skew or kurtosis outside the acceptable range of -2 to 2, if that transformed reduced the skew and kurtosis to within that acceptable range. They state that the overall pattern of results was unaltered by transforming the data (and only report the networks generated from the original non-corrected data). Second, Faelens and colleagues (2021) report that some variables had a non-normal distribution (as did their within-person means). They therefore include in their supplementary analyses a sensitivity analysis where they generate networks from log-transformed variables that had unacceptable ranges of skew and kurtosis. Faelens et al. (2021) state that the transformed and non-transformed networks show similar relationships between variables; and upon visually comparing the networks, the edges that are additional or missing are typically the thinnest edges. Possibilities for transforming non-normal data other than log-transforming exist, such as the non-paranormal transformation sometimes used in cross-sectional network studies to yield multivariate normal data (Epskamp, Waldorp, et al., 2018), which has been shown through simulation studies to be helpful when generating cross-sectional networks from continuous skewed data (Isvoranu & Epskamp, 2021). However, it is unknown how this transformation, applied to all variables at once, might affect the temporal dependencies present in time series data, or how it might impact the resulting temporal network estimation.

Since there is so little currently known about the impacts of violations of normality when generating temporal networks or transforming time-series data, the most we can recommend is that researchers check and clearly state whether the assumption of normality is upheld or violated. If it is violated and authors choose to transform their data in some manner, we recommend they do so as a sensitivity analysis and compare results with and without data transformation (which could briefly be described in the main manuscript but more fully included in the supplementary materials, for example), at least until more thorough investigations into non-normality, data transformation, and temporal networks are carried out. We also recommend that especially since the impact of normality violations are not yet fully understood, researchers could also check the robustness of their network results (e.g., by using case-dropping bootstrap techniques; Epskamp, 2020).

Unequal time observations

It is recommended to insert missing values when there are unequal time intervals between observations (or potentially impute missing data) to create a dataset with (roughly) equidistant timepoints (Hamaker et al., 2018). This is because some models (e.g., VAR models) assume equidistant data, or roughly equal time intervals between any two consecutive assessments (Jordan et al., 2020). One of the reviewed articles (i.e., Kaiser & Laireiter, 2019) responded to this model assumption by using cubic spline interpolation to generate equally spaced time-series. Another possibility, especially if investigating variables that operate on different timescales (e.g., behaviors and affect) is to use models that do not assume equal time intervals between timepoints, such as continuous time dynamic models (for more information, see Bringmann et al., 2022).

Missing data

Models estimated with mlVAR automatically use listwise deletion with missing data (Jordan et al., 2020), but other possibilities include imputing missing data (e.g., McCuish et al., 2021) or estimating missing data using a Kalman filter (e.g., Levinson et al., 2018). Two articles specifically used multiple imputation, which involves generating multiple imputed datasets, performing the required analyses with them, and then averaging the results: de Vos et al. (2017) specify they impute 10 datasets and averaged the subsequent results, and Levinson et al. (2020) use multiple imputation but do not further specify their methods. Some estimation methods are also more robust against missing data, such as full-information maximum likelihood (only when data is missing at random, however; Cham et al., 2017) or Bayesian estimation (e.g., Groen et al., 2020).

Ji et al. (2018) recommend against listwise deletion for missing data in intensive longitudinal assessment, as it is likely to lead to biased results, and instead suggest using multiple imputation (including partial multiple imputation supported by full-information maximum likelihood) and averaging results using Rubin (1996)'s pooled average. Ji et al. (2018) suggest at least 5 replications/imputations (their dataset has 30% missingness and around 100 participants) but discuss using more imputations for datasets with a higher percentage of missingness, smaller sample size, or complex pattern of missing data. We would therefore suggest that especially if authors have substantial amounts of missing data, they use a method robust against missing data, such as Bayesian estimation or multiple imputation (with sufficient replications).

Lag

All articles in this review use lag-1 (from time point *t* to the most recent time point previous, *t*-1), representing that they assume each timepoint is independent from others except for only the previous timepoint; other lags (such as lag-2) are more complex and require more repeated measurements (Jordan et al., 2020). However, some researchers argue that although defaulting to lag one (due to convenience or possibly for theoretical reasons) can be acceptable, ignoring higher-order lagged relationships that are present in the data can bias results (Jacobson et al., 2019). For example, if data is collected daily but variables are primarily predicting one another on a weekly basis, a lag of one (e.g., daily) will miss predictions at a higher-level lag (e.g., lag of 7 or weekly). However, a few possibilities do exist for researchers to investigate the optimal lag for their data: if a researcher is using mlVAR models, they can directly compare how different lags affect model fit using the *mlVARcompare* function and the *compareToLags* argument of the *mlVAR* R package (Epskamp et al., 2019). If a researcher is using another model or wants a more general method, they could use the Differential Time-Varying Effect Model tool (DTVEM package in R; Jacobson et al., 2019), which identifies the optimal time lag(s) in a dataset. We recommend that if researchers do investigate the effect of different lags, they report the results of the different resulting networks as sensitivity analyses.

Data modeling

Vector Autoregressive Models (VAR)

In general, VAR models regress a variable at time *t* on that same variable at time *t-1*. VAR models can also be extended to a multilevel framework where individuals are allowed to vary, called multilevel VAR (mlVAR); these are the most common type amongst reviewed articles (estimated by 95%). mlVAR models can be easily estimated using separate univariate regression equations (as generally implemented in the mlVAR package); the vast majority of authors used this estimation method (e.g., Contreras et al., 2019; Curtiss et al., 2019; Kuranova et al., 2021). Researchers generally choose to estimate mlVAR through separate univariate regression equations because they are convenient and fast to estimate in R (while still able to estimate correlated random effects), especially with fewer than 8 nodes (Epskamp, Waldorp, et al., 2018). However, this estimation method has weaknesses, including that it cannot directly estimate correlations between error terms or be combined with regularizations techniques to generate sparse networks (de Vos et al., 2017; Epskamp, Waldorp, et al., 2018).

Another possibility is to use multivariate estimation through Bayesian methods. Although this is not currently available through a package in open-source software (but can be done using the DSEM package in MPlus), it can model dependencies between error terms and is also more robust against missing data. Two reviewed articles used this method: Groen et al. (2020) and van Roekel et al. (2019). Groen et al. (2020) specify that they used Bayesian estimation because it can directly estimate a hierarchical model and subject-specific parameters (instead of indirectly as when a multilevel VAR model is estimated with separate univariate regression equations), can generate credible intervals (easily allowing an assessment of uncertainty or stability of parameters), and can handle missing data and unequal time observations well.

There is also the possibility to estimate a VAR model that is not multilevel, by pooling the data. This model assumes homogeneity across individuals and only estimates fixed effects (so overall is not very realistic for most data). However, this method can be used with regularization techniques and is also able to handle more nodes since there are less estimated parameters (with no random effects). Both McCuish et al. (2021) and de Vos et al. (2017) state they chose to use a pooled VAR for its regularization possibilities, since they wanted a sparse network with minimal spurious edges. However, de Vos and colleagues (2017) directly compare a pooled VAR model with a univariate estimation of a multilevel VAR model and conclude that since their data includes substantial individual differences (as is commonly the case), the multilevel model better fits their data.

There are other types of recently developed VAR approaches that are currently less common but offer great promise, although no reviewed articles utilized these approaches. For example, time-varying VAR models specifically account for time-varying processes and therefore are more robust against violations of stationarity (Haslbeck et al., 2020). For networks based on panel data or data that has less than 20 timepoints, panelVAR is a good option (Epskamp, 2020).

Other (non-VAR) models

Although VAR models are the most common among reviewed articles, accounting for 95% of estimated models, there are also other models suited to estimate temporal networks. For example, group iterative multiple model estimation (GIMME) takes individual structures into account while estimating a group level effect (similar to multilevel modeling), but does so using an iterative search process that tries to identify associations occurring consistently throughout individuals (Lane & Gates, 2017). Among reviewed articles, only Ellison et al. (2020) estimated a GIMME model, and they specify that they chose to do so because they wanted to base their models on within-person processes and build on within-person commonalities to form group networks, instead of assuming homogeneity across persons by directly modeling with a group or hierarchical approach. Another type of model, used only by Huckins et al. (2020) among reviewed articles, is the Peter and Clark Momentary Conditional Independence (PCMCI) algorithm that is more robust against high multicollinearity and missing data while also retaining high power even for networks with many nodes (Runge et al., 2019). Although not used by any reviewed articles, sparse time series chain graphical models are another interesting alternative, as they can estimate non-Gaussian networks, model network with mixed categorical and continuous data, and can also be used with LASSO regularization (Abegaz & Wit, 2013; Jordan et al., 2020). A more in-depth overview of most of the discussed models, as well as the software/packages to estimate them, are given by both Epskamp et al. (2018) and Jordan et al. (2020). In addition, all of the above models include discrete intervals (e.g., lags), but there is also the possibility to use continuous time dynamic models; these can be especially useful if including variables that operate at different timescales (for an overview, see Bringmann et al., 2022).

Complementary analyses

Reviewed articles did not just estimate temporal network models, but also typically examined specific parts of the networks or compared networks. More specifically, 58% of articles compared (visually or statistically) at least two temporal networks; to statistically compare network topology or centrality indices, most articles used permutation tests. In other articles, researchers examine the structure of networks, such as looking at the edges as well as the overall network connectivity. It is essential for those examining network connectivity to decide beforehand (and ideally, preregister) a specific definition or index of connectivity. In many articles, researchers also examine how central, or connected, specific nodes are. Strength is the most commonly used centrality index within reviewed articles (in 87% of articles), and it is also recommended as the centrality index that makes the most theoretical sense to use within psychology networks (Bringmann et al., 2019). We therefore recommend using (in- and out-) strength to assess centrality if edges all have the same signs (positive or negative), or (in- and out-) expected influence if there are both negative and positive edges (Robinaugh et al., 2016). It is important to note, however, that statisticians caution researchers regarding interpreting centrality: although examining which nodes are especially connected to other nodes can help generate hypotheses for future research, a node having high centrality does not necessarily mean that it plays a causal role in the network or that it is linked with clinical outcomes (Bringmann et al., 2019).

Assessing robustness and accuracy

Especially since temporal networks are a new (but growing) field, it is critical to assess whether the network results are stable and accurate; this is a concern that is currently being addressed for cross-sectional networks (e.g., Epskamp, Borsboom, et al., 2018). Few of the reviewed articles performed procedures to assess the robustness of models. However, one key method across analysis types to assess the accuracy of an estimated parameter (e.g., edge estimates) is to use bootstrapped confidence intervals, and these are very computationally intensive for multilevel models (and might not even be possible on a standard computer; Bringmann et al., 2013). It is possible to compute bootstrapped confidence intervals on pooled VAR models (e.g., with the *bootnet* package in R, as done by Lazarus et al., 2020; McCuish et al., 2021), although multilevel models are preferred to pooled models unless there is little heterogeneity among subjects (de Vos et al., 2017). It is more feasible to assess the stability of temporal networks through methods such as case-dropping (e.g., can assess if network parameters such as centrality indices or edge estimates are similar after dropping a randomly-selected percentage of participants), which can be conducted on any type of model. Jongeneel et al. (2020) employ this approach, dropping a randomly-selected 80% of participants. Epskamp (2020) also suggests that researchers can assess stability by dropping blocks of data (so as to take temporal dependencies into account, instead of dropping participants/individual rows as done to assess stability in cross-sectional networks; Epskamp et al., 2017). Another possibility is to use Bayesian estimation, as authors can then examine the credible intervals of parameters and thereby evaluate the uncertainty of estimates, as done by Groen et al. (2020). Assessing stability of temporal network analyses is important on its own merits, but it is even more crucial when there are violations of assumptions, since how these violations affect the resulting network is not yet clear (Epskamp, 2020)

Another crucial point is to examine the replicability of temporal networks. Among reviewed articles, 5% of articles split their sample to cross-validate their network models, an encouraging start to examining the replicability of temporal networks. However, that number is still quite low, and there were no studies replicating other temporal networks, which Guloksuz et al. (2017) point out is a crucial step to assess the stability and reproducibility of network models.

Another method to examine the robustness of the analyses is to conduct a sensitivity analysis, following 33% of reviewed articles. Sensitivity analyses involve running analyses with and without a specific change (e.g., preprocessing choices like the method used to compute variables or transformations to correct for an assumption violation, or model choices) to examine its impact on the results. Sensitivity analyses can be particularly useful when researchers have doubts about the impact of a specific arbitrary analysis choice. Since the field of temporal network analyses is so new and the impact of many analysis decisions on the results is not yet fully understood, it could also be interesting for researchers to perform multiverse analyses (Steegen et al., 2016) to examine the impact of many arbitrary analysis choices at once; a limited version was done by (de Vos et al., 2017).

Reporting

We urge researchers to share their code and anonymized data if at all possible on a platform such as the Open Science Framework (OSF) or GitHub. Sharing analysis code allows others to fully understand and replicate all preprocessing and analytical choices. Sharing anonymized data might not always be possible, but it can greatly contribute to the field when possible. For one thing, temporal network analyses rely on intensive longitudinal data, and this data is extremely difficult to collect; so sharing data allows others to investigate the same dataset as well, perhaps with updated methods as the field continues to advance (Bringmann et al., 2022) or to assess the impact of arbitrary analytical choices (e.g., with a multiverse analysis; Steegen et al., 2016).

It is also important that researchers report all relevant details of their data collection, preprocessing, and analysis, which was not always the case in the reviewed articles. A lack of clarity in reporting data collection and methodology is not uncommon in EMA research (Vachon et al., 2019), but since this field involves not only EMA research but also novel analytical methods, it is especially important to clearly report all aspects of data collection and analysis methodology so that the field can grow. We hope that researchers will find Figure 17 helpful in reviewing all essential information that should be reported in their manuscript (or supplementary materials), and we have also provided this information in a fillable checklist (https://osf.io/e6wp3/). For researchers conducting EMA data collection, they can also look at the reporting checklist by Trull & Ebner-Priemer (2020).

Limitations

This scoping review has some limitations. Since we wanted to survey a wide variety of methodology and analysis practices, we chose to limit the scope of our review to group-level temporal networks, which had to include time-series data. We therefore did not include articles focusing solely on idiographic temporal networks or panelVAR networks (based on cross-lagged panel data), although we do believe audit of these specific network types through systematic scoping reviews would be useful in the future. Our inclusion criteria did lead to some articles that were ambiguous on whether or not they should be included in the review. For example, Wichers (2014) focused on illustrating the temporal network approach, without the goal of providing an empirical contribution to the literature (other examples are discussed in Appendix D). Another potential limitation is that we use the findings of this review to formulate recommendations, although the field is still young; just because a number of papers use a specific practice does not mean that it is optimal. However, this is why we developed our recommendation section in light of relevant literature on EMA and temporal networks, as well as the results of the scoping review.

Future of the field

Through this scoping review, we also identify a few key areas of development for the field. First: assessing the robustness of temporal network analyses. Although there are some available methods to examine the overall stability of findings, such as through participant-dropping bootstrapping (e.g., removing a certain percentage of randomly-chosen participants and seeing how much this affects results; see Jongeneel et al., 2020), it is difficult to estimate bootstrap confidence intervals around edge/centrality estimates for multilevel temporal models (Bringmann et al., 2013). Although it is possible to pool data and not use multilevel models, this is only recommended when there is little heterogeneity among participants, since it is assumed that all participants are replications of one another (de Vos et al., 2017). One possibility to examine the robustness of point estimates is when using Bayesian estimation methods, since you can then build credible intervals around the estimated parameter; however, easily (e.g., with already built packages) using Bayesian methods to estimate temporal network requires using MPlus, a proprietary software. Being able to examine the robustness of network analyses has been deemed crucially important to the future of the field (Guloksuz et al., 2017), and we hope methods to easily assess the robustness of temporal networks will be forthcoming as well.

As we discussed in the 'Normality' section above, very little is currently known about how violations of normality affect temporal networks as well as how transformations of non-normal time-series impact results. We therefore reiterate the message from Epskamp, Waldorp, and colleagues (2018): We urge future researchers to investigate normality, assess how its violations and transformations impact temporal networks, and potentially develop new methods and models to address these challenges.

There have also been a variety of new models developed that address some of the shortcomings of current temporal networks. For example, time-varying VAR allows researchers to explicitly model non-stationarity data (Haslbeck et al., 2020), instead of correcting the data or ignoring the assumption violation, both of which have unknown impacts on the results (Jordan et al., 2020). However, this review has not identified much use of these models. We hope that future research will continue to explore and use new models as the field grows.

Section 3.

Patterns across time

Chapter 7.

A temporal network analysis: Parents from the general population

Many parents have days where they encounter emotional exhaustion, emotional distance from their children, and feeling fed up with being a parent. Some parents experience these characteristics to a severe extent—a clinical phenomenon termed parental burnout. Parental burnout arises when parents chronically endure severe stress without sufficient resources to cope, which may lead to detrimental consequences not only for the parent, but also for their partner (e.g., marital conflict) and children (i.e., neglect and violence). However, uncertainty remains regarding how these features interact and trigger one another over time (potentially becoming increasingly severe), nor how the daily variations of the family context influence these features. Therefore, in this study, we recruited 50 parents (with main analyses focusing on 43 mothers with a coparent, and sensitivity analyses with the full sample) from the general population to rate the core features of parental burnout and the family context daily over 56 days. We used multilevel vector autoregressive models to generate network models. Results suggest that exhaustion contributes to parental burnout: it self-predicts and is closely associated with feeling fed up and finding children difficult to manage. Distance, by contrast, is mainly negatively connected to sharing positive moments with children. Contextual variables also interact with parental burnout features, illustrating the relevance of examining parenting within the family system context. If future research confirms a central role of exhaustion in parental burnout development, prevention efforts can focus on decreasing parental exhaustion.

Reference

Blanchard, M. A., Hoebeke, Y., & Heeren, A. (2023). Parental burnout features and the family context: A temporal network approach in mothers. *Journal of Family Psychology, 37*(3), 398-407. https://doi.org/10.1037/fam0001070

7.1 Introduction

Parents have a challenging role, without question: they are responsible for caring for their children and raising them into adults. While parenting can be rewarding and joyful on some days, it can be difficult and frustrating on others. And for some parents, parenting brings overwhelming amounts of stress and exhaustion (e.g., pressure to be a perfect parent, dysfunctional family dynamics) without sufficient resources (e.g., supportive partner or extended family, emotional regulation skills) to cope (Mikolajczak & Roskam, 2018). If this imbalance persists for too long, the parent can experience parental burnout, involving four features: emotional exhaustion, emotional distance from the children, feeling fed up with parenting, and a sense of contrast with the previous parental self (Roskam et al., 2018). Research into parental burnout has only recently begun in earnest (Roskam et al., 2017), but it is already widespread across many languages and cultures (e.g., Arikan et al., 2020; Furutani et al., 2020; Mousavi et al., 2020).

Indeed, parental burnout is reported in countries around the world, with the highest prevalence rates rising to 8% of parents (Roskam et al., 2021). It has been linked with negative consequences for the parent (e.g., suicidal ideation and addiction), the couple (e.g., marital conflict), and the child (e.g., neglect and abuse) in cross-sectional research (Hansotte et al., 2021; Mikolajczak, Brianda, et al., 2018), as well as in longitudinal (Mikolajczak et al., 2019) and intervention (Brianda, Roskam, Gross, et al., 2020) research. Parenting stress and pressure have only increased during the COVID-19 pandemic (Griffith, 2022), with more parents (especially those with fewer resources) feeling exhausted and burned out (Aguiar et al., 2021; Kerr, Fanning, et al., 2021). This growing priority to help parents in difficulty, as well as the association of parental burnout with severe consequences for the children, highlights the pressing need to understand precisely how parental burnout develops and persists. Only with specific knowledge on the development of parental burnout can practitioners effectively prevent and treat parents with parental burnout.

Because parental burnout research is still in its early days, however, critical gaps remain in the literature. For example, most research has only investigated parental burnout as a cohesive and unitary phenomenon, with all four features (i.e., exhaustion, distance, feeling fed up, and contrast) summed into one whole. However, the few studies that have investigated these features separately (e.g., Blanchard, Roskam, et al., 2021; Hansotte et al., 2021; Kalkan et al., 2022) have all found that specific parental burnout features have distinct associations with family-related variables. For instance, emotional distance is most strongly associated with neglect toward children. In addition, the literature on parental burnout (Mikolajczak & Roskam, 2018) and on burnout more generally (Lee & Ashforth, 1993; Leiter, 1993) posits that exhaustion is the first step toward developing burnout. This possibility promotes investigating of the four features of parental burnout separately, to examine whether certain features are implicated in the instigation or maintenance of parental burnout.

Another crucial area for growth in the parental burnout literature involves the temporal unfolding of parental burnout. Until now, most research has utilized a cross-sectional approach and only investigated parental burnout at one timepoint. There are a growing number of longitudinal studies, but most did not investigate the evolution of the different features of parental burnout (e.g., Mikolajczak et al., 2019; Yang et al., 2021). Among the few studies that did (e.g., Aguiar et al., 2021; Roskam & Mikolajczak, 2021), they demonstrate that the features evolve in different patterns—and suggest that emotional exhaustion is the first step to developing parental burnout. Yet these studies all investigated parental burnout at month-long intervals, as required to understand the long-term evolution of parental burnout. However, the experience of parenting is something that varies from day to day, according to the ebb and flow of daily interactions with the children (e.g., Rodriguez & Silvia, 2022), partner or extended family (e.g., Gillis & Roskam, 2019), and wider family context (e.g., Malinen et al., 2017). It thus is also relevant to examine how the experience of parenting (including specifically feeling exhausted, fed up, and distant from one's children) can fluctuate from day to day, and how these fluctuations interact with the family context. These characteristics—feeling exhausted, distant, and fed up—are sometimes experienced to an extreme extent, as when a parent is in a severe state of burnout; but all parents will sometimes feel exhausted, distant, or fed up to some degree. To allow us to examine how parental burnout might develop, as well as how these characteristics are experienced by most parents and influenced by the family context, we decided to focus on the experiences of parents in the general population. However, our data collection ended up including almost all mothers. We therefore focus the analyses in this manuscript on this sample of mothers (but include sensitivity analyses with the full sample in the supplementary materials).

Our goal was therefore to examine the fluctuating experiences of the parental burnout features and interactions with the children, partner, and wider family environment (e.g., social support and parenting resources). We wanted to examine how these variables interact with each other and with the family context, to better understand their daily dynamics. To model the dynamic interactions between many variables, we used temporal network analyses, which are especially suited to visualizing dynamic multivariate relationships. Specifically, we generated three networks: 1) a temporal network to examine which variables predicted others from one day to the next; 2) a contemporaneous network to inspect how variables interrelated within the same day; and 3) a between-subjects network to observe the mean-level relationships between variables.

7.2 Methods

Sample size

Since there is currently no possibility to estimate a priori power analyses for temporal networks (especially without previous temporal network analyses on the same variables), we preregistered recruiting a minimum of 40 participants with 80% compliance, based on studies with a similar number of timepoints and nodes (Curtiss et al., 2019; de Vos et al., 2017; Lutz et al., 2018). We recruited parents that had at least one child living at home.

Participants

We recruited fifty French-speaking parents in Belgium through Facebook parenting pages and other online spaces. Of these, three were single parents[1], who were not included in the primary analyses (since these networks include nodes for partner support and conflict) but are included in a sensitivity analysis in the supplementary materials (with all 50 parents; see Figure S24 in Appendix E). As only four parents were men (not enough to make strong conclusions on the experience of fathers), we focus our analyses in this study on the 43 mothers with coparents; this is a deviation from our preregistration. However, we include a sensitivity analysis with all 47 parents with a coparent in Appendix E (see Figure S21). All parents with a partner were in a male/female relationship.

The total sample included in the main results therefore includes 43 mothers (see Table 9 for further demographic information). Participants' net family income was average-to-high for Belgium (see supplementary materials; Statbel, 2021). The mean number of surveys answered per person was 52.28, with a compliance rate of 93%.

Table 9. Demographic Information

Demographic Variable	Mean	SD	Min	Max
Age of parents	37.30	4.08	30	50
Number of children (living under same roof)	2.02	0.80	1	4
Age of children in years (living under same roof)	7.95	5.39	0.01	21.70
Parental Burnout Assessment (total score)	38.81	26.13	8	131
Balance between Risks and Resources (BR2)	53.60	52.43	-100	159
Generalized Anxiety Disorder-7 Questionnaire	7.98	5.60	1	20
Beck Depression Inventory	12.28	10.560	0	52

Baseline measures

Parental Burnout Assessment (PBA)

We assessed parental burnout using the PBA (Roskam et al., 2018), which measures the four features of parental burnout. Parents answered questions about emotional exhaustion (9 items; e.g., *When I get up in the morning and have to face another day with my child(ren), I feel exhausted before I've even started*), emotional distance toward their child(ren) (3 items; e.g., *I'm no longer able to show my child(ren) how much I love them*), feeling fed up (5 items; e.g., *I can't take being a parent anymore*), and a sense of contrast with their previous parental self (6 items; e.g., *I'm ashamed of the parent I've become*). Participants rated each item using a 7-point Likert scale ranging from 0 (*never*) to 6 (*every day*), and relevant items were reverse-scored. Scores could range from 0 to 138. Within the present sample, internal reliability was good for both

[1] "Single parents" here refers to a parent who is not parenting with a partner (e.g., someone with whom the participant shares childcare duties daily, typically but not necessarily living together).

the global scale (Cronbach's α = .97) and the individual subscales (Exhaustion: α = .94; Distance: α = .84; Feeling Fed Up: α = .89; Contrast: α = .91).

Balance of Risks and Resources (BR2) questionnaire

We examined participants' parenting risks and resources using the BR2, including perfectionist personality traits, stress management capabilities, parenting practices, co-parenting, etc. (Mikolajczak & Roskam, 2018). Each of the 39 items took a bipolar form, with a risk factor on the left side (e.g., *I find it difficult to reconcile my family life and my professional life*) and the corresponding resource on the right side (e.g., *I can easily reconcile my family life and my professional life*). Parents answered each item from -5 to 5, with a negative number indicating the risk factor statement more closely mirrored their experience, and a positive number indicating the resource statement more closely reflected their experience. A rating closer to |5| signifies a stronger endorsement of the (negative or positive) statement, while 0 indicates that the parent possessed neither risk factor nor resource factor. Scores could range from -195 to 195, with a negative score indicating parents have more risks than resources (and vice versa for a positive score).

Beck Depression Inventory-II (BDI-II)

We assessed depression symptoms using the BDI-II (Beck et al., 1996). Participants answered 21 questions, each time choosing the one statement out of four that most described how they felt during the past two weeks regarding specific symptoms (e.g., *I do not feel sad* to *I am so sad and unhappy that I can't stand it*) on a four-point scale from 0 to 3. Scores could range from 0 to 63. Within this sample, internal reliability was high, with a Cronbach's α of .93.

Generalized Anxiety Disorder scale (GAD-7).

We examined trait anxiety using the GAD-7 (Spitzer et al., 2006). Participants indicated their anxiety over the past two weeks on seven items (e.g., *worrying too much about different things)* using a four-point scale from 0 (*Not at all*) to 3 (*Nearly every day)*. Scores could range from 0 to 21. Internal reliability was high within this sample, with a Cronbach's α of .93.

Daily diary survey

The daily dairy survey consisted of 11 items focused on the parenting experience and family context. Ten items were assessed with a slider scale from 0 (not at all) to 100 (absolutely). These items focused on parental burnout (specifically emotional exhaustion, emotional distance, and feeling fed up[2]), partner relationship (partner support and conflict), children-focused relationship (finding children difficult to manage, sharing positive moments, and getting angry), resources, and social support. Parents answered the last item, measuring hours

[2] We do not include a daily item for "a sense of contrast with the previous parental self," since although this is an important component of parental burnout, its definition implies a stability over time (e.g., likely changing only over months or longer) that cannot be captured with daily surveys.

spent with kids, by entering a number between 0 and 24. The exact items can be found in Table 10 and we have previously described their development procedure and psychometric properties (Blanchard et al., 2023).[3] We reverse-scored parents' responses for emotional distance and resources.

Table 10. Parental Burnout and Family Context: Daily Diary Items

Category	Item (English version)
Emotional Exhaustion	I felt exhausted while caring for my children.
Emotional Distance	I felt close to my children in both good and bad situations. *(Rev)*
Feeling Fed Up	I felt overwhelmed caring for my children.
Partner Support	I received help from my partner with caring for my children.
Partner Conflict	I had some misunderstandings, tension, or arguments with my partner.
Difficult to Manage (Kids)	My children were difficult to manage.
Positive Moments (Kids)	I shared positive moments with my children.
Angry (Kids)	I got angry with my children.
Resources	I lacked the means (for example, time, energy, material resources) to take care of my children. *(Rev)*
Social Support	I received help from friends or family (other than my partner) with caring for my children.
Time with Kids	Around how many hours did you spend near your children today (outside of sleeping hours)?

Note. These daily diary items were presented in random order on the same page, with "Today" at the top. Rev = reverse-scored.

Daily diary procedure

First, a researcher conducted an introductory briefing session with each parent individually over a video call, with the experimenter explaining the overall study and demonstrating the daily diary items and software to the participant. We used *formr*, an open-source software, to collect data (Arslan et al., 2020); for further information, see Blanchard, Revol, et al. (2022). Participants received the link to complete the demographic questionnaires, and then, the next day, the first of the 56 daily surveys started. Parents received a notification (email or text message, as they preferred) with that day's survey each evening, between 6 PM and 9 PM (we asked parents to choose a time after most of their interactions with their children were over for the day). Participants received compensation: Each participant received 10€, and

[3] These items were conceptualized, created and piloted within the context of the experience sampling methodology (ESM) literature, which entails asking questions from one to many times a day over a long period of time. During this development period, described in detail in Blanchard, Revol, et al. (2022), we decided that a frequency of once per day would be most suitable for the parents (as many parents described spending mostly the evening with their children) as well as the items themselves (which we hypothesized would vary daily). Since these items were only assessed once per day, we describe this data collection as a "daily diary," despite the items themselves being created in the context of ESM literature.

those who completed at least 80% of surveys received an additional 25€. Parents provided written informed consent to share their anonymized data, and the project received approval from the local Biomedical Institutional Review Board (approval date: May 11, 2020; protocol title: PBNET). Each day, daily diary questions appeared in a random order (except 'Time with Kids', which always appeared last). We decided on a daily sampling scheme since parenting is typically more active at some times of the day over others (e.g., on weekdays: in the early morning before school but mostly the afternoon/evening after school). In addition, previous research suggests that parenting exhaustion varies from day to day (Gillis & Roskam, 2019); we assume that emotional distance, feeling fed up, and other parenting-related variables would also vary daily, and so a daily sampling scheme would be suitable to investigate these variables. The eight weeks of data collection ran from April through June of 2021. Further procedure information is reported in the supplementary materials.

Data analysis

We performed all analyses using R Statistical Software (v4.1.0; R Core Team, 2021) and used packages *mlVAR* (Epskamp et al., 2019) to estimate network models and *qgraph* (Epskamp et al., 2012) to visualize them.

Assumptions: Normality and stationarity

We checked for violations of normality using the Kolmogorov-Smirnov test, correcting for multiple testing with Bonferroni, following Aalbers et al. (2019). We also examined residual plots. We found that no variables were normally distributed. We followed our preregistration and log-transformed any variables with skew or kurtosis values outside the acceptable range of -2 to 2 (Aalbers et al., 2019). However, this transformation did not render variables closer to a normal distribution. We therefore attempted a different transformation using the R package *LambertW*, which uses an automatic procedure to optimally transform heavy-tailed and skewed data (Goerg, 2015). Although Kolmogorov-Smirnov tests suggested this transformed data was still not normally distributed, the skew and kurtosis values were at least within the range of -2 to 2. We thus used this transformation for the remainder of analyses (a deviation to our preregistration of only using a log transformation for non-normal data). As preregistered, we conducted a sensitivity analysis to examine whether transforming data to adhere to a normal distribution would change the pattern of results; it did (specifically for the temporal network, which was sparser with transformed data; see the supplementary materials for more details). As there is little information on how transforming non-normal intensive longitudinal data could impact the interpretation of temporal network analyses (Blanchard et al., 2023), we report network analyses based on the raw results in the manuscript. However, we report and discuss a model estimated from the transformed data in the supplementary materials (see Figure S22 in Appendix E).

To check for violations of stationarity, we used the Kwiatkowski-Phillips-Schmidt-Shin unit root test (KPSS; Kwiatkowski et al., 1992) to verify that the variance of all variables remained stable over time, as recommended by Jordan et al. (2020). We conducted the KPSS test for each variable of each participant, correcting for multiple testing with Bonferroni. All variables appeared stationary.

Network analyses

We modeled the parenting and family-related variables as networks using a multilevel vector autoregressive (VAR) approach. VAR models regress a variable at time *t* on itself and on all other variables at time *t-1*: they thus estimate how well each variable predicts all other variables at the next timepoint (Epskamp, Waldorp, et al., 2018). To account for the dependency of timepoints within subjects, we estimated the VAR model using a multilevel framework. This resulted in a *temporal network*, which visualizes the associations between variables from one timepoint to the next using arrows, while controlling for all other associations. The participant means are then used to generate a *between-subjects network*. This network shows the associations between variables on average across participants, collapsing across time (and controlling for all other variables). The between-subjects network is most comparable to partial-correlation cross-sectional networks (showing mean-level associations remaining after controlling for all other associations). Next, the contemporaneous (e.g., within the same timepoint) associations between all variables are estimated by regressing the residuals of the multilevel VAR model on all other residuals from that same timepoint. The resulting *contemporaneous network* visualizes how variables are related within the same timepoint, after controlling for all other contemporaneous associations and temporal associations; it can thus be thought of as a partial-correlation network. Epskamp et al. (2018) propose that the contemporaneous network likely captures processes that occur more quickly than the lag interval in the data (e.g., daily for this study).

Model specifications. We estimated the multilevel VAR model through sequential estimation of univariate multilevel regression models (Bringmann et al., 2013; Epskamp, Waldorp, et al., 2018). We allowed random effects in the model to be correlated. When visualizing the contemporaneous and between-subject networks, we used the "and" rule, requiring edges included in the networks to have both coefficients (from node X to node Y and vice-versa) be significant.

Additional analyses

We calculated strength centrality[4] (i.e., how connected a node is with other nodes) for all nodes in all networks (McNally, 2016). We report both the specific methods and the centrality

[4] We do not calculate expected influence centrality, although it considers positive and negative edges, since if a node had multiple positive edges, one negative edge would decrease the strength of the expected influence estimate. Although this could make sense for a network with all symptoms (i.e., all variables that presumably have the same valence), it makes less sense for a network representing the family system, with many variables representing contextual factors, and also variables that cannot be straightforwardly classified as "negative" or "positive." For example, "I got angry with my children" could be a justified and needed response if one's children decide to knowingly engage in dangerous behaviors while playing, but less so in other situations. For a further discussion on centrality and its complications for psychology networks, see the section on Centrality in the Discussion.

indices in Appendix E (see Figure S19). We also estimated the stability of these centrality indices using case-dropping, following the method and code of Jongeneel et al. (2020). The exact methods and results are reported in Appendix E (see Figure S20), but overall, the centrality indices were stable for the contemporaneous and temporal networks, and not stable for the between-subjects network.

Transparency and openness

We report how we determined our sample size, all data exclusions as well as the reason for exclusion, all manipulations (e.g., none), and all measures (either in the manuscript or in the supplementary materials). This study was exploratory, but we preregistered our study design, data collection procedure, and analysis plan following the preregistration template for experience sampling methodology (ESM; i.e., intensive longitudinal data collected in participants' everyday lives) research from Kirtley et al. (2021): https://osf.io/dz7nv. We specify in the manuscript where we deviated from this preregistration. We also share the R code, materials, and anonymized data on the Open Science Framework (https://osf.io/pshdn/), as well as Appendix E (also available at https://osf.io/g5t7h/).

7.3 Results

Descriptive statistics

For each variable, intra-individual means, standard deviations, and intraclass coefficient correlations, can be found in Table S9 in Appendix E. For all baseline measures (e.g., depression, anxiety, parental burnout), the vast majority of the sample was below clinical cut-offs, confirming that this sample was indeed from the general population.

Network analyses

Temporal network

Figure 18 represents how one variable at one timepoint predicts another (or itself) at the next time point, accounting for all other associations (edge values can be found in Table S10 in Appendix E). Most variables temporally predicted themselves (i.e., autocorrelated), except for distance, feeling fed up, and anger toward children. Of note, therefore, emotional exhaustion was the only parental burnout feature that self-predicts from one timepoint to the next. No bidirectional associations or cycles were present in the graph. The strongest edge pointed from positive moments with children negatively predicting distance. Among variables relating to the family context, partner conflict at one timepoint predicted greater social support at the next, which in turn predicted less partner support. Increased partner support, for its part, predicted decreased emotional distance.

Contemporaneous network

For this network showing partial correlations within the same day (with edge values in Table S11 in Appendix E), the strongest edge still negatively connected emotional distance with sharing positive moments with children. Emotional exhaustion and feeling fed up were strongly connected to one another, as well as to perceived resources and finding children difficult to manage. Finding children difficult to manage, for its part, showed a strong connection to getting angry with children, which was strongly connected with feeling fed up. Within one day, partner support and partner conflict were only weakly connected to each other, and isolated from other nodes.

Between-subjects network

For this network visualizing the mean-level partial associations between variables (with edge values in Table S12 in Appendix E), the thickest edge still negatively connected emotional distance and sharing positive moments with children. Mothers who reported higher levels of exhaustion also tended to report higher levels of feeling fed up, greater resources, and less partner conflict. Feeling fed up was negatively associated with partner support and positively associated with finding children difficult to manage. Social support, for its part, was positively associated with partner conflict and negatively associated with getting angry with children.

Sensitivity analyses

We conducted four different sensitivity network analyses: a model with all 47 mothers (including the four fathers), a model with Gaussianized-transformed data, a model with raw data including the variable 'Time with Kids,' and a model including all 50 parents (including single parents, and therefore without the partner-related variables). The specific changes between those networks and the one reported in this manuscript are detailed in the supplementary materials, but overall, the networks remain relatively consistent: most thick edges remain in all networks. The major exception is with the Gaussianized-transformed temporal network, which is sparser than the temporal network with raw data. One interesting small difference is an edge from feeling fed up toward exhaustion appeared in three of the sensitivity analyses: the network with all parents with co-parents (n=47; Figure S21 in Appendix E), network with all parents with coparents, including the variable 'time with kids' (n=47, Figure S23 in Appendix E), and the network with all participants (including single parents) and no coparenting-related nodes (n=50; Figure S24 in Appendix E). With this additional edge, one cycle appears in the network: exhaustion at one timepoint predicted finding children difficult to manage at the next timepoint, which in turn predicted feeling fed up, which predicted exhaustion.

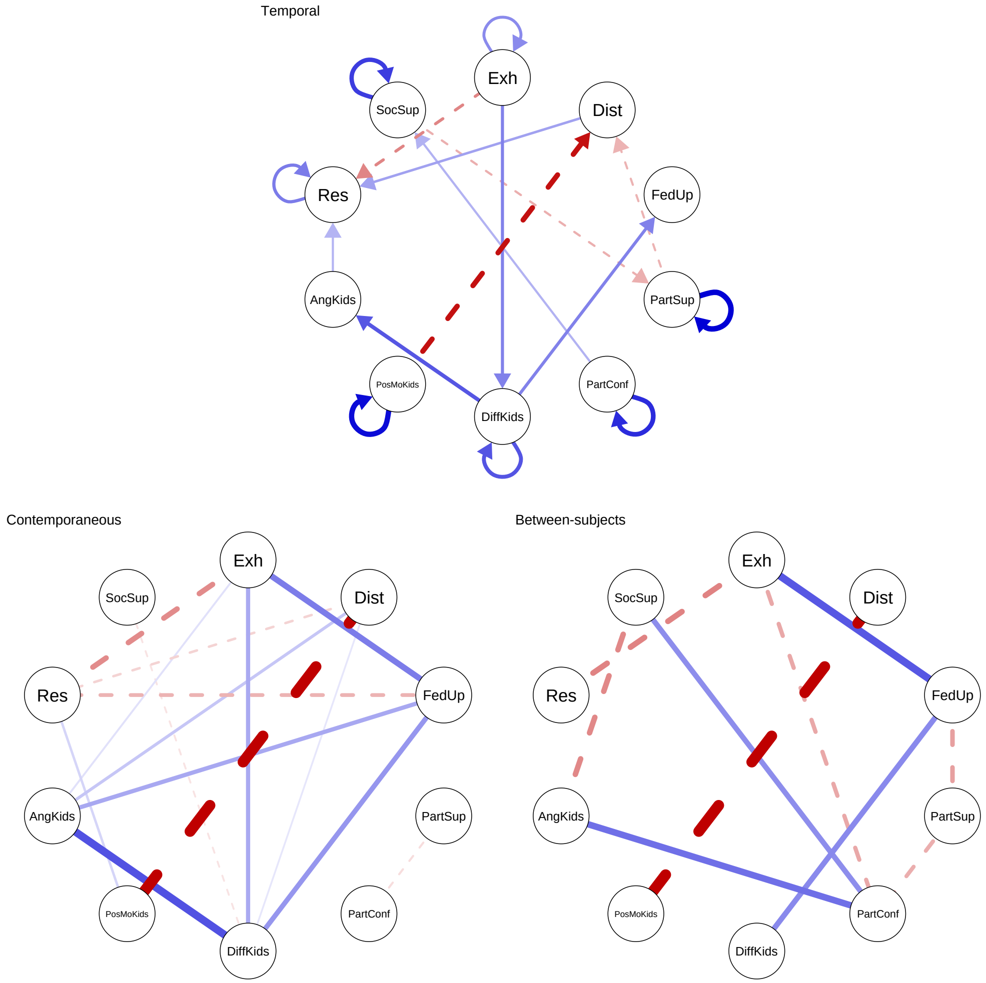


Figure 18. Network analyses: Parents from the general population

Note. Solid blue edges represent positive associations, while dashed red lines represent negative associations. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PartnerConf = partner conflict; PartnerSupp = partner support; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support. Edge values for all networks (including p-values, standard errors, etc.) can be found in Appendix E (see Table S10, Table S11, and Table S12).

7.4 Discussion

This study examined how parents' daily experiences of feeling exhausted, distant, and fed up interacted with each other and with their family context. Exhaustion predicted finding children difficult to manage in the temporal network, which in turn predicted feeling fed up. In the contemporaneous network (i.e., within the same day), exhaustion, feeling fed up, and finding children difficult to manage were all related to one another, while in the between-subjects network, exhausted mothers were more likely to also feel fed up, and feeling fed up was associated in turn with finding children difficult to manage. These three variables therefore seem closely connected. Interestingly, exhaustion was the only parental burnout feature to self-predict itself in the temporal network. Taken together, this suggests exhaustion as a kickstart toward activating the key features of parental burnout, which dovetails with prior theoretical (Mikolajczak & Roskam, 2018) and empirical works (Le Vigouroux et al., 2022; Roskam & Mikolajczak, 2021).

Conversely, emotional distance was not connected with either of the other two parental burnout features in any network. Instead, it was strongly connected in all three networks with sharing positive moments with children, and in the temporal network, this edge points from sharing positive moments toward distance. This is consistent theoretically, as sharing positive moments with children builds closeness. Previous empirical research supports this as well; for example, a previous daily diary study found that on days where parents report more warmth (i.e., less distance) toward their adolescent children, these children reported feeling more loved (Coffey et al., 2020). Importantly, research on quality time with children highlights that positive moments between a parent and child do not need to be intensive, long, or preplanned but can be spontaneous and part of everyday activities, as long as both child and parent enjoy the time spent together (Hsin, 2009; Kremer-Sadlik & Paugh, 2007).

Distance being disconnected from the other features of parental burnout also holds with the theory that exhaustion drives the development of parental burnout (Mikolajczak & Roskam, 2018), as well as burnout more generally (Lee & Ashforth, 1993; Leiter, 1993). The present sample consists of mostly mothers from the general population, with overall low levels of parental burnout. It therefore makes sense that exhaustion is especially connected to other variables in the present networks, as exhaustion is the first active component of parental burnout for this population. Previous literature has suggested that emotional distance might play a role in maintaining parental burnout once it has developed (e.g., Blanchard, Roskam, et al., 2021), although this possibility has not yet been investigated empirically. Nonetheless, if emotional distance were crucial in maintaining parental burnout, it would make sense that it would only be very connected to other related variables for parents with high levels of parental burnout. This therefore remains an important question to examine within samples of parents with severe parental burnout: are all three features (emotional exhaustion, emotional distance, and feeling fed up) closely connected in the context of parental burnout? Does emotional distance play a key role in maintaining parental burnout?

When examining the role of the family context, there were no connections between partner variables or social support and any parental burnout variables that emerged in all three networks (and in fact, these nodes were all isolated in the contemporaneous network). However, distinct connections appeared in the temporal (e.g., partner support negatively predicting distance) and between-subjects (e.g., partner support negatively correlating with feeling fed up; partner conflict negatively correlating with exhaustion) networks. These demonstrate that the family context (i.e., interactions with the partner and wider social support) interrelates with a parent feeling exhausted, fed up, and distant. For its part, the node representing resources has strong connections with the parental burnout features, particularly with exhaustion; an edge connects these two nodes in all three networks. Interestingly, all edges connecting resources with other nodes pointed toward resources in the temporal network. For instance, when mothers feel exhausted one day, they feel like they have less resources the next day. On the other hand, if they spend positive moments with their children and if they feel distant toward their children, they feel like they have more resources the next day. This suggests that after mothers are more distant with their children, they feel like they regain resources later on. This is consistent with a previous longitudinal study that examined the differential course of the features of parental burnout (Roskam & Mikolajczak, 2021). The authors suggested that, similarly to theories in the job burnout literature, detachment might "protect parents from negative affect and cognitions about parenting." Overall, these contextual variables interact and predict parental burnout features and child-related behaviors, illustrating how parenting experiences are informed by and interconnected with the family context as a whole.

Taking these results together, exhaustion emerges as a potential kickstart of parental burnout development, as it is highly self-predictive (and so likely accumulates over time) and could lead to a downstream negative cascade through its close connections with feeling fed up and finding children difficult to manage. If future studies confirm that exhaustion drives the initiation of parental burnout, it would be a key target for prophylactic intervention. This coheres with network theory, which posits that targeting a highly central (e.g., very connected within the network) node in an early intervention could lead to a beneficial cascade that "turns off" other nodes (McNally, 2016). Identifying parents at risk of burnout, such as exhausted parents, and preventing them from developing severe parental burnout, is a critical goal, since parental burnout has severe psychological consequences (e.g., neglect and violence toward children; Mikolajczak et al., 2018, 2019). To target exhaustion, one option could involve practitioners focusing on the individual family. For example, such interventions could revolve around decreasing parenting stressors, increasing restorative time for the parents, or lessening the parenting load over the long term (e.g., increasing routines and predictability for the children, or better balancing family demands; Masten, 2018; Mikolajczak & Roskam, 2018). Intervening on emotional exhaustion could also stem from a more system-based perspective, however, by centering on bolstering adequate parenting resources, family resilience, and community support (e.g., Masten, 2018).

Emotional exhaustion can also arise from intense social pressure to be a "good parent" (at least in Euro-American countries, with their intensifying parenting norms; Roskam et al., 2021). Indeed, studies have suggested that high societal standards for parenting can lead to parental exhaustion (Kawamoto et al., 2018; Sorkkila & Aunola, 2020). Parents (and particularly mothers) themselves also describe the pressure they feel to embody the ideal parent, leading them to overinvest and exhaust themselves (Hubert & Aujoulat, 2018). This is not new: Hays discussed the ideal of such child-centered and selfless parenting (often expected and prescribed specifically of the mother) decades ago, terming it "intensive parenting" (Hays, 1996). Parenting experts and researchers pushed for such "intensive parenting" ideals (e.g., Bradley et al., 1997), implying that parents (and particularly mothers) should "constantly delight in their child, never feel as though their child is demanding, accommodate the child's needs, and never want to leave the child" (Liss et al., 2013). As Hays deftly argues, these prescriptions are unrealistically demanding and actively lead to parental guilt, as well as culturally specific and indifferent to diverse parenting compositions and needs (Hays, 1998). Researchers have posited that parents with unattainable parenting goals will use too many resources to (try to) reach these impossible goals, thereby exhausting themselves and being at risk of parental burnout (Le Vigouroux et al., 2022). Prevention efforts could therefore also attempt to alter parents' vision of parenting to something more realistic. For example, for parents (particularly mothers) who prioritize their children's needs above all else and strive to always be positive and warm (as suggested by intensive parenting methods; Hays, 1998), practitioners may help them to better balance the needs of others and to be firm while also being compassionate (Dupont et al., 2022). Another possibility would be to increase parents' resources, such as by improving parents' emotional competencies, which have been shown to buffer against the effects of parental perfectionism (Lin, Szczygieł, et al., 2021). Of course, most helpful would be to shift Euro-American societal parenting norms themselves to be more attainable, but shifting societal norms is slow work implicating political agendas, legal definitions, expert consensus, and broad social narratives.

The present study has limitations. A first limitation is that the residuals were not normally distributed, which is an assumption for the multilevel VAR model. In a recent scoping review about temporal network analyses, less than a quarter of studies examined whether the assumption of normality is violated (Blanchard et al., 2023), and little is known about how non-normal data, or transforming the data, might affect the results or interpretation—although hopefully this will be a target for future statistical and theoretical development. Therefore, we chose to report the raw data in this manuscript and the transformed networks in the supplementary materials as a sensitivity analysis, similarly to Faelens et al. (2021). The results were similar, except for the temporal network generated from transformed data containing fewer edges. We also assessed the stability of the centrality indices to gauge how sensitive our results were to specific subsamples, through case-dropping: the only network that was not stable was the between-subjects network. This is understandable, since individual participants have much more sway on the results of the between-subjects network (as it collapses across all temporal information and uses only mean responses). However, to truly assess whether our results are stable and accurate would require replication in another sample.

Another limitation is the specificity of our sample. Participants were mostly Belgian mothers, and we did not collect information about racial or cultural identification. The sample was a convenience sample (recruited mainly through online parenting pages), and mostly mothers ended up participating. Since we only had four fathers participate, we are not able to confidently generalize the present results to fathers, although it is promising that the sensitivity analysis that includes all parents with coparents, including four fathers (n=47) in Figure S21 is very similar to the results with just the 43 mothers with coparents. However, one main difference is that Figure S21 involves a feedback loop between exhaustion, finding children difficult to manage, and feeling fed up: a thin edge connecting feeling fed up and emotional exhaustion is present in Figure S21 but not in Figure 18 (with only mothers, n=43). We nonetheless believe that emotional exhaustion plays a key role in the parental burnout network for mothers, even without this feedback loop in Figure 18, particularly as emotional exhaustion is still the parental burnout variable with a self-predicting loop. In any case, this study should be repeated with a sample of mostly or all fathers, to examine if the network structure is consistent or not—particularly in light of intensive parenting norms falling especially on mothers, even if through gendered assumptions and not explicitly (Hays, 1998).

Although this study is the first of its kind and grants new information on how parents' daily experiences interact with their family context, there is substantial cultural variation in parenting experiences and parental burnout, specifically relating to individualism (Roskam et al., 2021). Although the results in this manuscript might generalize to other parents in individualistic cultures, there might be important differences with parents in collectivist cultures. For example, parents might have different visions of "ideal" parenting, different childrearing goals, and different expectations of community parenting support (Bornstein, 2012). Future studies should therefore examine daily parenting experiences and the family context in other populations.

Conclusions

Mothers experience different levels of exhaustion, feeling fed up, and distance from their children every day, and these variables continuously interact with each other and with the family context. In line with previous research, exhaustion appears as a potential jumpstart to parental burnout, since it is strongly associated with feeling fed up and finding children's behavior difficult to manage, as well as the only parental burnout feature to self-predict from one day to the next. In this unselected sample, emotional distance is not connected to the other parental burnout features. Sharing positive moments with children does, however, predict feeling less distant. Contextual variables (partner support, partner conflict, social support) also interact with parental burnout variables in various networks (e.g., from one day to the next, on a group level), emphasizing the relevance of viewing parenting and parental burnout within its context of the family system.

Chapter 8.

A temporal network analysis: Parents in burnout

Parental burnout, which includes emotional exhaustion, emotional distance, and feeling fed up with parenting, is a chronic and potentially debilitating stress condition that has become highly prevalent in Western countries. Yet uncertainty remains about parents' daily experience of burnout and how their experiences interact with their family environment. In this study, we surveyed a sample of 40 parents (who felt overwhelmed and exhausted by parenting) daily for eight weeks about their experiences of parental burnout, interactions with their children, and family environment, yielding 1541 total observations; main analyses focus on 36 mothers. We estimated a multilevel vector autoregressive model and created three distinct graphical networks to best reveal the dynamics of these interactions. All models point to interactions with children as a key determinant in the parent's daily experience of parental burnout—particularly predicting emotional distance and feeling fed up. At the temporal level, parental burnout symptoms were characterized by strong self-prediction (autoregression). In addition, the severity of parental burnout was negatively correlated with daily variation of parental burnout features. Together, these results suggest that mothers in a severe state of burnout are less sensitive to their context, with the exception of interactions with their children.

Currently under review, with preprint at:

Blanchard, M. A., Hoebeke, Y., & Heeren, A. (2023). *Parental burnout and decreased sensitivity to context for mothers: A temporal network approach grounded in the family environment*. Manuscript submitted for publication. Preprint available via PsyArXiv at https://doi.org/10.31234/osf.io/hs6tb

8.1 Introduction

Parental burnout is a chronic stress condition (Brianda et al., 2023) characterized by four hallmark features: severe emotional exhaustion, emotional distance from children, feeling fed up as a parent, and a sense of contrast with the former parental self (Roskam et al., 2018). It has severe and potentially debilitating consequences for the parent (e.g., escape and suicidal ideations, sleep problems), but also for the partner (e.g., conflict and disaffection) and the child (e.g., abuse and neglect; Mikolajczak et al., 2018). Researchers currently theorize that parental burnout occurs when parents chronically lack sufficient parenting resources to compensate for the parenting risks and stressors they face (Mikolajczak & Roskam, 2018). However, rates of parental burnout vary across the world, with prevalence rates as high as 8% in some Western countries (Roskam et al., 2021), which suggests that parental burnout could also be related to intensifying cultural norms about "idealized" parenting in these countries (Lin et al., 2023; Liss et al., 2013). Given the widespread prevalence of parental burnout and the serious threat it can pose to parents' mental health, actionable efforts are needed to better understand how parents experience it and how it impacts and echoes through the family system.

The experience of parenting is embedded in the family system: parenting is not an activity that occurs in isolation, but by definition implies interactions and communication with family members (Dallos & Draper, 2015; Kerig, 2019). Therefore, one may wonder how the daily experience of parental burnout affects interactions with a (possible) coparent, with the child(ren), and with the broader and ever-changing family context (Blanchard & Heeren, 2020). All parents experience feelings of exhaustion, distance and being fed up with parenting to some extent, in ways that impact their interactions with other family members (Blanchard, Hoebeke, et al., 2023b). However, previous research examining these interactions focused on parents from the general population (Blanchard et al., 2021; Blanchard, Hoebeke, et al., 2023b). It is likely that parents with severe parental burnout experience parenting differently, however; according to the conceptualization of parental burnout (Roskam et al., 2018), parents with high levels of parental burnout would experience parental burnout symptoms to a severe extent in their daily life. In this project, we therefore expected that parental burnout symptoms would fluctuate less from day to day, as parents would experience these symptoms to a severe extent every day, with less influences from other daily happenings. We investigated this preregistered hypothesis by examining whether within-person standard deviation of daily reports of emotional exhaustion, emotional distance and feeling fed up would be smaller for parents with severe burnout than for parents from the general population (i.e., this latter involving participants from a prior study; Blanchard, Hoebeke, et al., 2023).

No prior study examined how parents with severe burnout interacted with their family system at a micro (e.g., daily) level, even though this fine-grained knowledge would be one key method to understand the processes underlying parental burnout. This would require not only examining the daily experiences of parents with severe parental burnout, but also how these experiences were grounded within the parent's family system. To examine the interactions between these many variables, and to account for the temporal nature of the data, we used temporal network analyses. Specifically, we used a multilevel vector autoregressive (mlVAR) model, which models how each variable is predicted by all other variables (including itself) from the previous timepoint. This type of network approach allows us to examine fine-grained predictions over time while placing the experience of parental burnout in its context—the family.

One previous study has used this same daily diary and temporal network analysis approach (Blanchard, Hoebeke, et al., 2023b) but only in unselected parents, thereby focusing on risk factors; its results support the hypothesis that emotional exhaustion could kickstart the development of parental burnout (Le Vigouroux et al., 2022; Roskam & Mikolajczak, 2021). However, there is still little known about what maintains parental burnout. We have previously proposed that for parents already in a state of burnout, emotional distance might play a maintaining role: the distance parents feel from their children would limit their motivation to seek help and support (Blanchard et al., 2021; Blanchard, Hoebeke, et al., 2023b). If this is the case, then we expect that emotional distance will show strong connections to the other parental burnout features in the network visualization, as well as connections to the child-related variables.

8.2 Methods

Participants

In total, 41 participants started the study, but one participant dropped out after five daily measurements (for information about sample size estimation, see Appendix F: or https://osf.io/4uwaf). Of the remaining 40 participants (36 mothers and four fathers), 30 participants were coparenting with a partner while 10 were single parents (all mothers). To align with a previous project with similar variables but an unselected sample (Blanchard, Hoebeke, et al., 2023b), we decided to only include mothers in analyses in the main manuscript, since the number of fathers was few enough that results might not fully represent or generalize to all parents, but be specific to mothers. As such, the final sample includes 36 mothers (but note that we include a network with all parents as a sensitivity analysis; see Appendix F). To recruit participants, we contacted all psychologists and clinical workers in the francophone part of Belgium that had completed training on parental burnout and asked those who agreed to transmit information about the study to any relevant patients. We also posted about the study in relevant local social media groups (e.g., about parental exhaustion, parental burnout).

Recruitment criteria involved being a parent with at least one child under 17 years living with the parent (although since we had initial difficulties with recruitment, we relaxed this for e.g., parents that had alternate weeks with their child) and feeling exhausted and overwhelmed as a parent. Participants' incomes ranged widely, but at least half of the sample reported average-to-high net monthly incomes for Belgian families (*Statbel: 2020 Average Belgian Income*, 2022). In addition, all but one parents coparenting with a partner (4 fathers and 26 mothers) were in a relationship format of man/woman; one parent was in a relationship format of woman/woman. For further information on participants (i.e., employment status, specific income brackets), see Appendix F.

Sample size

We had preregistered (link: https://osf.io/yfb84) recruiting at least 40 participants, based on ESM studies with similar number of timepoints and number of nodes. However, as we were unsure how difficult recruitment would be (e.g., working with a population with exhaustion and chronic stress on a long study), we had preregistered that if we had less than 30 participants who finished at least 80% of daily surveys, we would extend the data collection until December 2022. In May 2022, we only had 13 participants who had completed the study, and so we extended the study. In the end, our last participant entered the study on 11 January 2023 (we had a few participants meant to begin in December 2022 but who were in the end unable to start until January).

Baseline measures

Parental Burnout Assessment (PBA)

The number of items in each subscale and example items are as follows: emotional exhaustion (9 items; e.g., *When I get up in the morning and have to face another day with my child(ren), I feel exhausted before I've even started*), emotional distance toward their child(ren) (3 items; e.g., *I'm no longer able to show my child(ren) how much I love them*), feeling fed up (5 items; e.g., *I can't take being a parent anymore*), and a sense of contrast with their previous parental self (6 items; e.g., *I'm ashamed of the parent I've become*). Participants rated each item using a 7-point Likert scale ranging from 0 (*never*) to 6 (*every day*), and relevant items were reverse-scored. Scores could range from 0 to 138. Within the present sample, internal reliability was good for both the global scale (Cronbach's α = .97) and the individual subscales (Exhaustion: α = .93; Distance: α = .87; Feeling Fed Up: α = .94; Contrast: α = .93). The cut-offs for parental burnout, as determined in Brianda et al. (2023), are as follows: 52.7 for moderate severity, and 86.3 for severe parental burnout.

Other baseline questionnaires

We assessed a number of other questionnaires at baseline, to have an overall measure of depression, anxiety, and parenting stressors and resources for the sample. Specifically, we assessed depression with the Beck Depression Inventory-II, with possible scores ranging from 0 to 63 (Beck et al., 1996) and anxiety with the Generalized Anxiety Disorder Scale, with possible scores ranging from 0 to 21 (Spitzer et al., 2006). We also measured parenting risks and resources using the Balance of Risks and Resources questionnaire (Mikolajczak & Roskam, 2018), with possible scores ranging from -195 to 195. Further information on these questionnaires (including reliability information) can be found in Appendix F.

Daily diary survey

The 11 daily diary items, assessing daily parenting experiences, interactions with child and coparent (if relevant), and broader context, can be found in Table 11. Their development procedure and psychometric properties are described in depth in previous research (Blanchard et al., 2023). All items were presented on a slider scale of 0 (*not at all*) to 100 (*absolutely*).

Table 11. Parental Burnout and Family Context: Daily Diary Items

Category	Item (English version)
Emotional Exhaustion	**I felt exhausted while caring for my children.**
Emotional Distance	**I felt close to my children in both good and bad situations. *(Rev)***
Feeling Fed Up	**I felt overwhelmed caring for my children.**
Partner Support	I received help from my partner with caring for my children.
Partner Conflict	I had some misunderstandings, tension, or arguments with my partner.
Difficult to Manage (Kids)	**My children were difficult to manage.**
Positive Moments (Kids)	**I shared positive moments with my children.**
Angry (Kids)	**I got angry with my children.**
Resources	**I lacked the means (for example, time, energy, material resources) to take care of my children. *(Rev)***
Social Support	**I received help from friends or family (other than my partner) with caring for my children.**
Time with Kids	Around how many hours did you spend near your children today (outside of sleeping hours)?

Note. These daily diary items were presented in random order on the same page, with "Today" at the top. Item development and preliminary validation procedures are described in Blanchard, Revol, et al. (2023). Rev = reverse scored. Bolded items are presented as nodes in the networks in this article (see Figure 19), and other items are also included as nodes in the sensitivity analyses (see Figure S26 and Figure S27 in Appendix F).

Procedure

Parents attended a virtual one-to-one briefing session with one of the researchers for this study, where the researcher explained the full procedure, went over the exact items, and answered any questions. Participants answered all items (i.e., baseline and demographic questionnaires and daily diary surveys) via the *formr* software (Arslan et al., 2020); further details are described in Blanchard, Hoebeke, et al. (2023). Parents answered daily diary items each evening between 6PM and 9PM (at a time of their choosing) and had five hours to respond; in total, parents received daily diary items for 56 days. Daily diary items were presented in randomized order. Parents were compensated for their participation: all parents received 10€, and parents who answered at least 80% of daily surveys received an additional 40€. All parents provided written informed consent, and the project received approval from the UCLouvain biomedical review board (approval date: May 11, 2020; protocol title: PBNET). In total, data collection spanned from March 2022 to March 2023. Participants took a median time of 1.38 minutes to complete the daily diary.

Data analysis

We performed all analyses using R statistical software (v4.1.0; R Core Team, 2021). Our main analyses used the *mlvar* package to estimate network models (Epskamp et al., 2019) and the *qgraph* package to visualize them (Epskamp et al., 2012), while throughout our R code we made use of tidyverse packages (Wickham et al., 2019).

Assumptions: Normality & stationarity

We assessed whether residuals were normally distributed with the Shapiro & Wilk (1965) test (corrected for multiple testing with Bonferroni), as well as visually examined residuals, both as preregistered. No variables were normally distributed. We attempted an automatic transformation using the Gaussianize function from the R package *LambertW* (Goerg, 2015), but this did not improve (and sometimes worsened) skew and kurtosis values, and so we ultimately did not correct the data for normality. Overall, there is little information on how to deal with non-normality in multilevel VAR models (Blanchard et al., 2023), with some anecdotal suggestions that it is not a crucial assumption with sufficient time-series observations (e.g., McNeish & Hamaker, 2020) or might mainly impact power to detect small edges (Veenman et al., 2022); we discuss our approach further in Appendix F.

To assess stationarity violations, we used the Kwiatkowski-Phillips-Schmidt-Shin unit root test (KPSS; Kwiatkowski et al., 1992) to verify the stability over time for all variables as preregistered and as recommended by Jordan et al. (2020). We corrected for multiple testing with Bonferroni. We found no stationarity violations.

Network analyses

We estimated a multilevel vector autoregressive (VAR) model, which regresses each variable at time t on itself and all other variables at time $t\text{-}1$, thus estimating how each variable predicts all other variables at the next timepoint (Epskamp, Waldorp, et al., 2018). Each variable (i.e., node in the network) was formed from a daily diary item (see Table 11). To account for the dependency of timepoints with each subject, we used a multilevel model: the model therefore includes both fixed (i.e., group-level) effects and random (i.e., individually estimated) effects. From this model, we generated three networks. First, a *temporal network*, which visualizes associations between variables from one timepoint to the next using arrows, while accounting for all other associations. Second, a *contemporaneous network*, which visualizes contemporaneous (i.e., within the same timepoint) associations between variables by regressing the residuals of one variable from the multilevel VAR model on all other residuals from the same timepoint. This network therefore illustrates how variables are related at the same timepoint, after controlling for all other contemporaneous and temporal associations;

thus, it is similar to a partial-correlation network. Epskamp et al. (2018) propose that the contemporaneous network could capture processes occurring more quickly that the lag interval (e.g., daily for this study). Last, participant mean responses are used to generate a *between-subjects network*, which illustrates the associations between variables on average between participants, collapsing across time (and controlling for all other variables).

Model specifications

We estimated multilevel VAR models using sequential univariate regression models (Bringmann et al., 2013; Epskamp, Waldorp, et al., 2018), with random effects allowed to correlate. When visualizing the contemporaneous and between-subject networks, we required edges included in the networks to have significant coefficients in both directions (e.g., from node X to Y and vice-versa), otherwise known as the "and" rule.

Group comparisons

We compared the sample in the present analysis (parents who self-reported being exhausted and overwhelmed by parenting, hereafter termed the "Parental Burnout Sample") with a sample of unselected (i.e., general population) parents from Chapter 7 (i.e., Blanchard et al., 2023). As we had hypothesized that parental burnout symptoms would fluctuate less from day to day for the Parental Burnout Sample, we used one-tailed t-tests.

Additional analyses

Fixed versus random effects

In addition to the networks (which illustrate fixed group-level effects), we followed the example (and R code) of Veenman et al. (2022) and also plotted boxplots of the random effects, with large red dots showing the fixed effects (see Figure 20). This allowed us to examine if the individual estimates clustered around the group-level estimate, or if there was wide variation or differences between these.

Centrality

As there are general concerns with the interpretation of centrality indices (i.e., how connected a node is to all other nodes) for psychology networks (Bringmann et al., 2019), we only report on centrality (i.e., strength centrality) in Appendix F.

Transparency and openness

We preregistered the data collection design and analysis strategy as well as two hypotheses (although the study is mostly exploratory): https://osf.io/yfb84. We share anonymized daily diary data and R code here: https://osf.io/5sgkr/.

8.3 Results

Descriptive statistics

To assess general severity of parental burnout in the sample, we used the cut-off scores for the Parental Burnout Assessment as established by Brianda et al. (2023): 28 participants (70%) scored above the cut-off for moderate severity of parental burnout, while 13 participants (32.5%) reached severe parental burnout.

Table 12. Demographic Information of Sample

Demographic Variable	Mean	Median	SD	Min	Max
Age of parents	37.8	37.5	8.61	15	55
Years of study of parents [a]	10.28	10.0	2.61	0	16
Number of children (living under same roof)	2.10	2	0.90	1	4
Age of children in years (living under same roof)	8.40	7.4	5.51	0.20	20.2
Parental Burnout Assessment (total score)	69.55	64.5	31.81	14	132
Balance between Risks and Resources (BR2)	8.18	6.5	48.89	-91	105
Generalized Anxiety Disorder-7 Questionnaire	10.3	10	5.18	1	21
Beck Depression Inventory	19.98	18	10.17	5	48

Note. [a] Years of study involves how many years of formal education the parent took part in, starting with secondary school (7th grade or age ~12)

Participants answered an average of 52.28 daily surveys, and overall had a compliance rate of 93%. Descriptive information regarding demographics and baseline questionnaires can be found in Table 12, while descriptive information for ESM measures (e.g., intraindividual means, SDs, intraclass correlations) can be seen in Table 13.

Group comparisons: Participant SDs

When comparing the intraindividual standard deviation of daily reports of exhaustion, distance, and feeling fed up between the current "Parental Burnout" sample with an unselected (i.e., general population) sample from a previous study (Blanchard, Hoebeke, et al., 2023b), we found no support for parents in the parental burnout sample having lower standard deviations. In fact, when looking at group means of standard deviations, the parental burnout sample had higher standard deviations (e.g., greater daily fluctuation) than the unselected sample. Exact results are reported in Table S14 and Table S15 in Appendix F. However, we also examined the correlation between PBA total score and intraindividual standard deviation across the two samples, and found significant correlations for all three parental burnout features—specifically: Exhaustion: $r(88) = .22$, $p = .01$; Distance: $r(88) = .43$, $p < .001$; Fed Up: $r(88) = .34$, $p < .001$.

Table 13. Intra-individual Means, Standard Deviations, and ICCs for Each ESM Item

Daily diary item	iiMean	iiSD	ICC
Emotional Exhaustion	41.61	28.44	0.29
Emotional Distance	31.02	21.67	0.37
Feeling Fed Up	30.53	25.91	0.27
Partner Support	52.44	30.23	0.28
Partner Conflict	16.7	25.16	0.08
Kids' Behavior	31.37	25.96	0.23
Behavior toward Kids (+)	70.98	20.25	0.39
Behavior toward Kids (-)	26.13	27.9	0.2
Resources	57.48	27.42	0.35
Social Support	24.55	31.35	0.18
Time Spent with Kids	7.54	3.31	0.53

Note. ICC = intraclass correlations; iiMean = intraindividual Mean; iiSD = intraindividual standard deviation.

Network analyses

Temporal network

The temporal network visualizes how each variable predicts other variables and itself at the next timepoint (e.g., the next day); see Figure 19A. Exact edge values for the temporal network can be found here: https://osf.io/bny23. The temporal network is sparse and almost all edges relate to Positive Moments with Kids: a cycle with Emotional Distance, and a thick positive edge from Positive Moments with Kids toward Resources. The only other edge is from Finding Kids Difficult to Manage toward Feeling Fed Up. Of note, all variables except for Resources and Anger toward Kids have positive autoregressive (i.e., self-predicting) loops.

When looking at how random effects (i.e., individual coefficient estimates) compared to fixed (i.e., group-level) effects (see Figure 20A), most individual estimates cluster together near the group-level estimate, with the important exception of the edge from Positive Moments with Kids toward Resources (which is highly unusual for the group-level estimate to be so far from the mean of random effects; we assume that this could be due to the multilevel structure of the model, which could pull some group-level parameter estimates away from individual values, but this is just a hypothesis and we are unsure exactly why this occurred). Even for all other edges, however, there are almost always at least one or two parents (and sometimes more) where the estimated edge has a very different magnitude or even a different sign. In other words, for at least some parents, the connection between two variables is actually opposite from what is shown in the group-level network in Figure 19A (e.g., instead of Positive Moments with Kids predicting lowered Distance, there are at least two parents where the estimated model shows Positive Moments with Kids predicting greater Distance). The only edges where all random effects have the same sign are the autoregressive loops for Feeling Fed Up and Emotional Exhaustion.

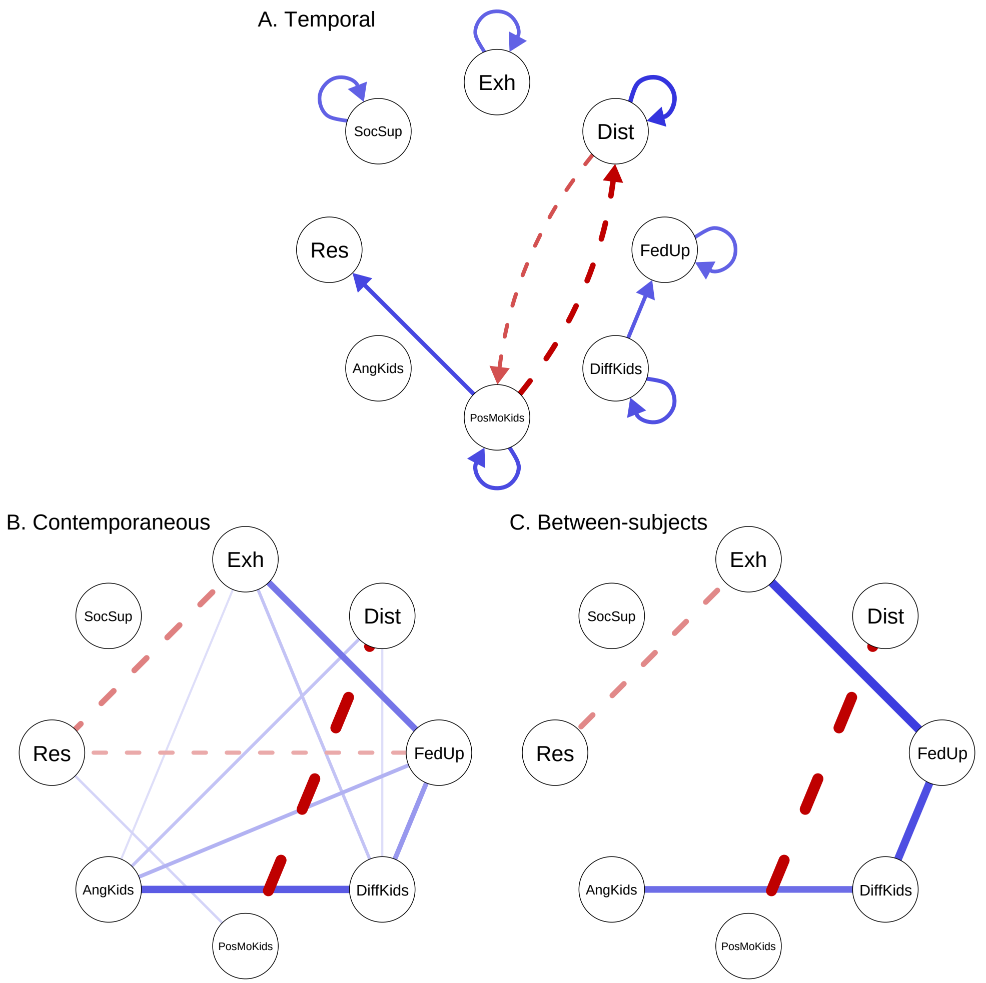


Figure 19. Network analyses for parents in burnout

Note. Solid blue edges represent positive associations, while dashed red lines represent negative associations. (A) The temporal network shows associations between each variable and all other variables (including itself) from one day to the next, controlling for all other associations. (B) The contemporaneous network showed partial associations within the same day. (C) The between-subject network shows mean-level partial correlations between participants. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support. The exact edge values for all networks can be found as supplementary files: temporal (https://osf.io/bny23), contemporaneous (https://osf.io/ypcdj), and between-subjects (https://osf.io/w4bns).

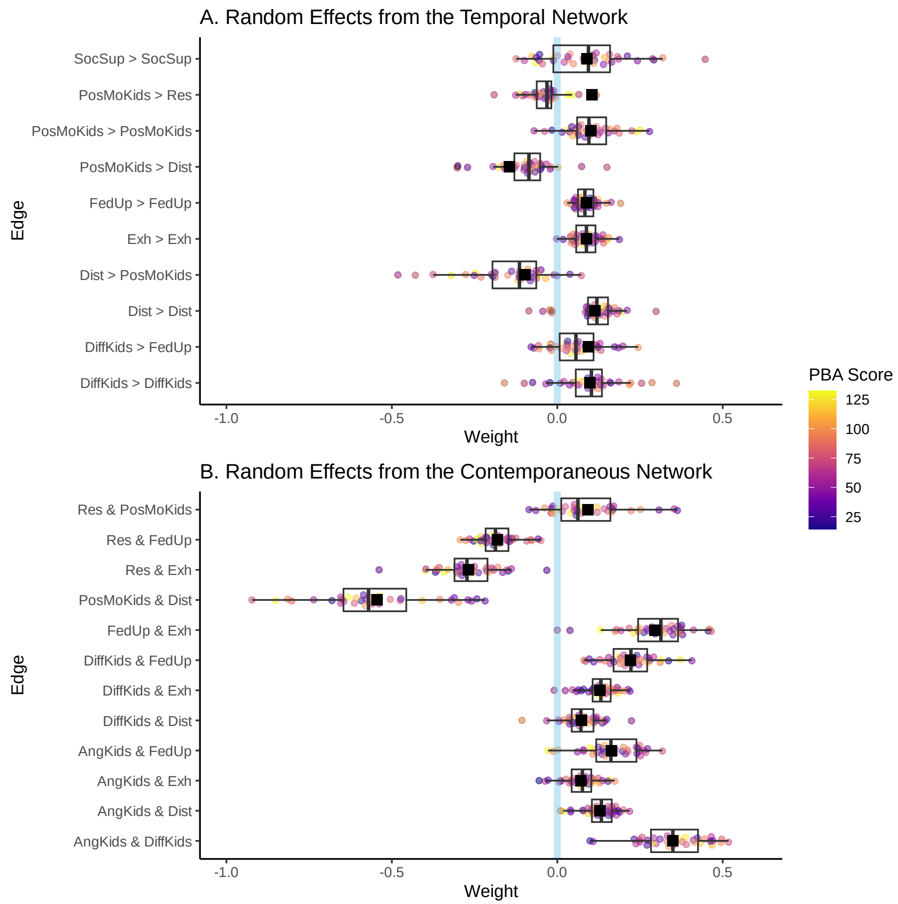


Figure 20. Histograms of Random Effects of the Temporal and Contemporaneous Networks

Note. Solid black squares show the edge values from the network analyses in Figure 19. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support; PBA = Parental Burnout Assessment (with a larger score indicating increased parental burnout severity).

Contemporaneous network

Figure 1B visualizes the partial correlations between variables for the same day, estimated from the residuals of the model represented in Figure 19A. Exact edge values for the contemporaneous network can be found here: https://osf.io/ypcdj. This network has more edges than the temporal network, but the same associations (e.g., between Positive Moments with Kids and Distance, and between Positive Moments with Kids and resources) are still present. Of note, all parental burnout variables (i.e., Exhaustion, Distance and Fed Up) are positively associated with Anger towards Kids and Finding Kids Difficult to Manage, while both Exhaustion and Fed Up are negatively association with Resources. Interestingly, Social Support has no associations with any other variables. When examining random effects (i.e., individual coefficient estimates; see Figure 20B), the group-level estimate (found in Figure 19B) is always near the mean of the individual estimates. However, there are still multiple coefficients with associated random effects that are quite spread out and even a few with different signs/valences.

Between-subjects network

Figure 19C shows the partial correlations between the mean levels of variables. Exact edge values for the contemporaneous network can be found here: https://osf.io/w4bns. This network contains few edges, with the thickest negatively connecting Distance with Positive Moments with Kids. The other edges show that parents who on average report having resources are less likely to report feeling exhausted; and those parents in turn who on average report greater daily exhaustion also endorse higher daily amounts of Feeling Fed Up. Similarly, Fed Up is positively associated with Finding Kids Difficult to Manage, which in turn is positively associated with Anger toward Kids.

Sensitivity analysis

We conducted three sensitivity analyses with different variations of included participants and/or nodes, to examine how these changes impacted the network results (full results and discussion can be seen in Appendix F). First, we examined a model with all parents (n=40), including the four fathers (see Figure S25), and results were overall similar to the main results (Figure 19; just mothers, n=36).

Second, we estimated a network with the two coparenting nodes added in (nodes=10), and thus without the 10 single parents (n=30). The contemporaneous and between-subject networks (see Figure S26) had all the thick edges from the main analyses (Figure 19) also present, although the between-subjects network did have some additional edges connecting to Resources. The temporal network, however, had more divergences. Some changes directly related to Partner Conflict and Partner Support (e.g., very strong autoregressive effects that led to thinner edges for all other connections), but the addition of partner-related variables also impacted how other variables interacted with Positive Moments with Kids (i.e., less connections to other variables in the temporal network) and Resources (i.e., predicting increased Distance in the temporal network). Note, however, that for this sensitivity analysis, there were initial convergence issues when estimating a model with correlated random effects (likely due to poor power with the reduced sample), which could also have impacted results; in the end, we estimated a model with orthogonal random effects.

Third, we examined a network with the "Time with Kids" variable included (i.e., how many hours the parent spent with their children that day, not including sleep hours), for all parents (n=40). In general, adding Time with Kids did not change the structure of the networks (see Figure S27) compared with the main analyses (Figure 19): all thick edges were still present.

However, for the temporal network, adding Time with Kids led to a very strong autoregressive effect for Time with Kids and much thinner edges for all other connections. In addition, Time with Kids now predicted Positive Moments with Kids (which no longer had an autoregressive loop), suggesting that the amount of time parents spent with their kids predicted the shared positive moments the next day, but not any other variables (e.g., finding children difficult to manage, feeling exhausted, distant, or fed up). For the contemporaneous network, Time with Kids shared connections with a few more variables (e.g., not only Positive Moments with Kids but also Finding Children Difficult to Manage, Fed Up and Exhausted), but in the between-subjects network, none of these connections remained.

8.4 Discussion

Based on the theoretical framework of parental burnout (e.g., Roskam et al., 2018), we had expected that this sample of parents in burnout would show less daily fluctuations of parental burnout features then a previous unselected sample (Blanchard, Hoebeke, et al., 2023b). When comparing groups with *t*-tests, we did not find support for this hypothesis, but when we examined the correlation between the intraindividual standard deviations of daily reports of parental burnout features (i.e., exhaustion, distance and feeling fed up) and the total PBA score on the other, we did find significant correlations for all three features. This therefore supports that the more extreme the parental burnout, the less parents experience daily fluctuations of their experiences of burnout features. It also demonstrates that within the 'unselected' and 'parental burnout' samples, there is significant variation of parental burnout severity. This is also seen in the PBA scores for this sample: 70% were above the moderate severity cut-off for burnout, but the range of PBA scores was wide. Interestingly, daily variations of exhaustion had the smallest correlation with PBA score. In contrast, distance had the largest correlation, which aligns with the hypothesis that emotional distance is particularly crucial to maintaining parental burnout: the higher the parental burnout score, the more parents experience constant emotional distance, with little daily variation. In other words, the environment and family context would have less of an impact on parents' experience of distance (as well as fed up and exhaustion, although to a lesser extent) as the severity of parental burnout increases.

This limited impact of context on parenting experience is also reflected in the results from the temporal network, which is generally sparse: overall, there are limited interactions between the family context and parental burnout features. We did not find evidence for our hypothesis that Distance would be strongly connected to Exhaustion and Fed Up, although it was the only parental burnout feature predicting a child-related variable (specifically, Positive Moments with Kids), or indeed any other variable in the network. We would therefore posit that Distance might still be the parental burnout variable with the most key role in the network for this sample of burned-out parents, although Positive Moments with Kids is the variable with the most connections. In addition, all three parental burnout features have strong autoregressive associations (with Distance exhibiting the thickest loop). This implies that if a parent is experiencing exhaustion, distance or feeling fed up, they likely also will the next day (which was not the case for the unselected sample in Blanchard, Hoebeke, et al., 2023).

Thus, three aspects of the results together suggest that parental burnout is a state less impacted by the family context than general parenting: the sparsity of the network, the high autocorrelation of parental burnout features, and the association of parental burnout severity with less daily variation of exhaustion, distance, and feeling fed up. In other words, once a parent is in a state of burnout, they will likely stay in that state of burnout, regardless of changes in their family context.

The main interactions likely to impact the parent's experience are their interactions with their children. For example, the only variable to predict a decrease of a parental burnout feature is sharing positive moments with children, which predicts less emotional distance from children in the next timepoint. This result is clinically relevant, especially when taking into account the sensitivity analysis including Time with Kids (Figure S27), where the same edge is still present: this means that sharing positive moments involves not just spending more time with children, but specifically the quality of the interaction (supporting previous research; e.g., Hsin, 2009).

One other child-related variable predicts a parental burnout feature: if parents find their children difficult to manage one evening, they are likely to feel more fed up the next day. This association is also present in the temporal network of the unselected sample (i.e., Blanchard, Hoebeke, et al., 2023). From the current data collection method (i.e., asking the parent if they find their children difficult to manage), we cannot know if this item reflects mainly the parent's perception of the child's behavior as difficult to manage, or mainly the actual behavior of the child—most likely, it reflects both to some degree. Nonetheless, this finding highlights the importance of ensuring parents have adequate childrearing skills and strategies (as discussed by Mikolajczak & Roskam, 2018), as well as general resources when in difficult situations (e.g., convenient/free childcare situations when the parent is overwhelmed or unavailable), so that parents are less likely to (regularly) find their children difficult to manage, and then feel fed up. Although these proposals center on helping individual parents, they implicate large-scale institutional endorsement (e.g., governmental funding), and their implementation should be focused on community building and support networks (Webster-Stratton, 1997).

It is crucial to note, however, that there is substantial individual variation in results: the magnitude—and in some cases even the direction—of edges varies for different parents, as shown via the random effects in Figure 20. This high level of individual variation highlights the extent to which parenting is dependent on its specific context—for example, the resources parents have available will be vastly different from parent to parent, including social support, the relationship with the partner (if applicable), and so on. This also meshes with systemic thinking as applied to families, which generally emphasizes "uncertainty in understanding how a system works," as family systems are overwhelmingly complex with their many interactions within the family and interacting forces from outside (Dallos & Draper, 2015, p.

148). We thus cannot designate one pattern that always occurs within one family, let alone within all families. We can, however, seek patterns that are likely to occur and re-occur across certain types of families that are experiencing specific problems (such as parental burnout), even while acknowledging and accepting that there will always be idiosyncrasies (e.g., individual variations). As such, identifying the temporal edges that are generally similar for most participants in Figure 2 grants insight into which patterns are common across the different families included in the present sample: the autoregressive effects for the parental burnout features and the edge from Positive Moments with Kids toward Distance.

Nonetheless, we believe that further research must be conducted before implementing these findings directly into clinical practice, or assuming that these results hold true for all parents with burnout (which is unlikely in any case). For one thing, our sample is extremely specific: mothers living in Belgium. Although the sensitivity analysis including the four fathers who participated in the study shows very similar results (see Figure S25), it is still possible that a sample of all fathers would show different overall patterns. Since the only features impacting the parental experience in the temporal network were child-related variables, we also call for future research on parenting experience, particularly when focused on severe parental burnout, to include dyadic (parent-child) and whole family studies. In addition, cultural factors influence how parents understand their role as parents and interact with their broader context (e.g., depending on their childrearing goals, understanding of a “good” parent, expectations of community parenting support; . Future research should therefore also examine daily experiences of parental burnout within the family context in other samples, to see if and how patterns diverge. This is not only necessary to understand the generalizability of results, but also to explore how experiences of parental burnout differ according to the specific cultural norms surrounding parenting and family life.

Chapter 9.

Affective dynamics

Our emotional trajectories make up our affective experience—but these can be disrupted during mental illness. This study focuses on affect anchored to the parenting context (i.e., daily emotional exhaustion, emotional distance from children, and feeling fed up) to assess whether the way parenting affect fluctuates relates to dysfunction: parental burnout severity. We focus on three specific patterns (i.e., affective dynamic indices): inertia (i.e., persistence across days), variation (i.e., magnitude of change), and covariation (i.e., whether affect variables fluctuate together). We reanalyzed multiple datasets (from Belgium and the U.S.) yielding 180 parents who had rated their parenting affect daily for either three or eight weeks. We computed a regression model with all affective indices as predictors (controlling for mean levels), with parental burnout severity as the outcome variable. Results indicate that inertia of emotional distance predicts parental burnout severity across most sensitivity models (i.e., even with varied operationalizations of affective indices). No other temporal pattern (i.e., variation or covariation) robustly predicted parental burnout severity, although the mean levels of emotional distance and emotional exhaustion did. Results from sensitivity analyses emphasize that operationalization choices for affective indices can yield varying values and impact results.

Currently under review, with preprint at:

Blanchard, M. A., Kerr, M.L, Kirkorian, H., Barr, R., Hoebeke, Y., & Heeren, A. (2023). *The Affective Dynamics of Parenting: Inertia of Emotional Distance as a Key Characteristic of Parental Burnout Severity.* Manuscript submitted for publication. Preprint available: https://doi.org/10.31234/osf.io/ekp49.

9.1 Introduction

Parenting represents a gargantuan task, encompassing the weighty responsibility of raising children and preparing them to navigate their adulthood. The impact parenting has on children's development and well-being is well-established (Masten & Palmer, 2019; Ulferts, 2020), but there is overall less focus placed on the experience of parenting and its impact on the parents themselves. One area of research that does examine when parenting becomes dysfunctional for the parent does so via the construct of parental burnout: How do parents become overwhelmed and exhausted by parenting, to the point where they might no longer feel close to their children or want to be a parent (Mikolajczak et al., 2019)? Parental burnout involves severe emotional exhaustion, emotional distance from children, feeling fed up with parenting, and a sense of contrast with the previous parental self (Roskam et al., 2018). We know from previous research that these parenting experiences fluctuate from day to day for all parents (Blanchard, Hoebeke, et al., 2023b), and we will hereafter refer to the daily variations in parents' experiences of emotional exhaustion, emotional distance, and being fed up with parenting as negative "parenting affect." We term these variables (negative) 'parenting affect' because we use affect as an umbrella term for emotion, mood and stress response (following Gross, 2015) [1], here applied specifically to the context of parenting. Note that these three affect items are connected to parental burnout, and in fact were developed to represent daily fluctuations in the three main dimensions of parental burnout (Blanchard et al., 2023). What is still unknown, however, is whether the *way* parenting affect fluctuates—its "trajectories, patterns, and regularities" (Kuppens, 2015)—can meaningfully shed light on the experience of parenting and how it might morph into the dysfunctional extremes of parenthood, such as parental burnout.

Research into how (and why) affect fluctuates falls under the heading of 'affect dynamics' (Kuppens, 2015). This research area focuses on the dynamic nature of emotions and affect, such as how affect ebbs and flows across time. Much research into affective dynamics has focused on broad composite variables of 'positive affect' and 'negative affect' (Gray & Watson, 2007), but experts also point to the importance of specificity and context (Koval et al., 2021). For example, examining specific discrete emotions often provides more information on distinct patterns than when joining variables together into broad categories (e.g., irritability and shame vs. combining them with other emotions into a broad "negative affect" construct; Mneimne et al., 2018). The context of emotional reactions and affect changes are also key: there is a higher signal-to-noise ratio when emotional assessments are anchored to a particular context, as they are more focused and capture less random fluctuations of affect (Dejonckheere & Mestdagh, 2021). Moreover, the study of affect dynamics is not limited to emotions but is also relevant to other processes such as state rumination (Bean et al., 2021) and solitude (Elmer et al., 2020);

---

[1] In fact, when discussing affect dynamics, it is also common to use the terms 'emotion dynamics' and 'affect dynamics' interchangeably, as the concept encapsulates everything from "short-term emotional fluctuations to longer term mood swings, involving one single emotion (component) or the interrelation or concordance across multiple emotions (or components) across time" (Kuppens, 2015, p. 298).

the crucial component involves examining the patterns and trajectories of internal experiences. Hence, we adopted a similar approach in this study and sought to examine the fluctuations of affect anchored to the parenting context, and to specifically explore patterns of three distinct parenting components: emotional exhaustion, emotional distance, and feeling fed up.

Most research on affect dynamics specifically examines whether certain dynamic patterns characterize psychological (mal)adjustment, or even specific psychological disorders. However, much of this research has substantial methodological flaws, as detailed by Dejonckheere et al. (2019): the majority of studies only examine the association between specific affective dynamic indices and psychopathology without taking into account the mean or overlap with other related indices. For many of these dynamic affect indices, the mean level best predicted (mal)adjustment, and once the mean level (or other affective dynamic indices) was controlled for, any association disappeared (Dejonckheere et al., 2019; Koval et al., 2021). In other words, if someone has an overall high level of negative affect, this is typically more informative about their general well-being than the specific patterns in which they experience negative affect. This does not mean that these patterns are not interesting, but that methodological robustness is crucial for understanding the precise usefulness of a specific pattern. As Dejonckheere et al. (2019, p. 485) highlight, an affective dynamic measure should "summarize[s] affective time series in a unique way" and have "significant additional explanatory power in the prediction of psychological well-being."

We wanted to build on these recent developments in the study of affective dynamics and apply them to the parenting domain, to search for patterns and trajectories in the daily experience of parenting. Although researchers have examined affective dynamics in a family context, it has typically been from the lens of developmental psychopathology, focusing on interactions between people on short timescales of seconds to minutes. Research that has focused on within-person processes over hours and days in a family context often focus on children and adolescents (Maciejewski et al., 2019; Van Lissa et al., 2017; Vannucci et al., 2019). Few studies directly pertain to the ebb and flow of parents' affect, even though parenting affect patterns not only shed light on the processes and difficulties of parenting—relevant to a large swathe of the population—but also impact the entire family system (Kerig, 2019). Studies that have directly examined affective dynamics of parenting experiences (e.g., Kerr, Rasmussen, et al., 2021; Somers et al., 2020) have thus far included only general affect measures, and not measures of affect specific to the parenting context. In addition, no study thus far has explored whether and how parenting affect patterns predict psychological maladjustment specific to the context of parenting, such as parental burnout.

To better understand how specific patterns of negative parenting affect might predict parental burnout, we examined three specific and commonly used affective indices: variability, inertia, and covariation. *Variability* represents the average change from the mean level per person over the entire time period investigated (Dejonckheere et al., 2019). We expected that less variability in each of the three negative parenting affect variables would predict more severe parental burnout. This follows the overarching theoretical framework of parental burnout: that severe parental burnout occurs through a chronic accumulation of parenting stress, which results in increased severity and frequency of parental burnout symptoms (Mikolajczak & Roskam, 2018). We hypothesized that this increased frequency and severity of symptoms would also yield less fluctuation in negative parenting affect. In addition, one previous study found similar results: smaller standard deviation of parental affect predicted more severe parental burnout (Blanchard, Hoebeke, et al., 2023a).[2] However, in light of the importance of distinguishing the impact of affective dynamic indices from the mean affect level (as described above and in Dejonckheere et al., 2019), we specifically examined relative variability, which takes into account the maximum possible variability given the mean affect level (Mestdagh et al., 2018).

We also sought to capture the *inertia* of negative parenting affect, which can be defined as "a propensity for emotions to linger across time, reflecting slower recovery from emotional perturbations" (Koval et al., 2021).[3] In other words, inertia represents resistance to change or "getting stuck" (in the apt words of Elmer et al., 2020b) in a way of experiencing the world (e.g., in a specific emotion or mood, but also more broadly, such as in a state of solitude). We expected that high levels of negative parenting affect inertia—characterizing parents who linger in experiences of exhaustion, distance and feeling fed up from one day to the next without changing in response to new situations—would be associated with more severe parental burnout.

Lastly, we investigated the *covariation* of negative parenting affect: whether specific affective experiences (here: emotional exhaustion, emotional distance, and feeling fed up) vary together. If parental burnout is a unitary construct, as theorized (Roskam et al., 2018), then we expected that for parents with severe parental burnout, the negative parenting affect variables would fluctuate in sync with one another.

The current study thus analyzes these three patterns of parenting affect: relative variability, inertia, and covariation. Or in other words: how much parents' affect varies overall, whether parents linger in states of negative parenting affect, and whether their experiences of different components of negative parenting affect synchronize with one another. To examine these distinct affect patterns, we reanalyze three datasets with repeated daily measures of parents' emotional exhaustion, emotional distance, and feeling fed up. We aim to pinpoint which affective dynamic patterns characterize the daily experience of parenting when it becomes maladjusted and dysfunctional—or in other words, when it takes the form of parental burnout.

---

[2] Note that this dataset is in fact one part of the data reanalyzed in the current study. However, the specific index of variation was different (here, we use relative SD), and it also did not account for the mean level.

[3] Note that inertia specifically examines persistence from one timepoint to the next, and so takes the order of temporal fluctuations into account, while variation refers only to the overall magnitude of change, regardless of moment-to-moment fluctuations (Koval et al., 2021). This means that someone can have both high variability (a broad range in their general level of affect) and high inertia (persistence of emotions from one day to the next), such as through slow but large changes in affect levels.

9.2 Methods

Participants

To maximize the sample size, we include multiple datasets from other projects: two from Belgian participants (original projects: Blanchard, Hoebeke, et al., 2023b, 2023a), and one from American participants (the original study, titled the Ecology of Media Use project, is not yet published). Details regarding these samples, including their original projects, where to find the data, and key differences between the samples, can be found in Table 14.

Table 14. Information on the three datasets reanalyzed for the current study

Characteristic	Dataset #1	Dataset #2	Dataset #3
Location	Belgium	Belgium	US
Original project	https://doi.org/10.1037/fam0001070	https://psyarxiv.com/hs6tb/ *[preprint]*	EMU study *[not yet published]*
Data	https://osf.io/pshdn/	https://osf.io/9zgbw/	*[will be uploaded after data collection]*
Response Scale	0-100	0-100	0-10
Recruitment criteria	Min. 1 child < 16 years	Min. 1 child < 16 years	Min. 1 child aged 1-2 years (11-23 months)
Language	French	French	English
% Mothers	86%	90%	91.3%
Max. # of timepoints	56	56	21
Dummy code in Model	Belgium	Belgium	US
Total # of participants	50	40	104
# of participants (after excluding if SD = 0)	49	40	91

Further demographic information regarding the three samples can be found either in the original publications or (in the case of the US sample) in Appendix G (also at https://osf.io/xc5rd). The original number of participants included 90 from the combined Belgian sample and 104 from the US sample.[4] However, since multiple indices required variance for daily responses, we also excluded participants that had at least one daily variable that they answered identically for the entire data collection.[5] This excluded 1 participant from the Belgian sample and 13 from the US dataset, yielding a final sample of 180 participants.

[4] As preregistered, we only included samples from the US dataset that answered the Parental Burnout Assessment and at least 50% (i.e., minimal 11 out of 21) of the daily surveys.

[5] This was not included as part of the preregistration but was required since relative SD requires a non-zero standard deviation to compute it. Only one participant had no variation for all three daily parenting affect variables, so technically it could have been possible to compute inertia estimates for variables with variability for the other participants. However, this could be problematic for some of the inertia analyses (e.g., for multilevel vector autoregressive models, which include cross-lags between daily parenting affect variables, autoregressive estimates would still be impacted), and in any case these participants would have been removed in the regression analysis via listwise deletion. Moreover, if a participant answered the exact same (in our datasets, participants with no variation answered "0" every day), there are doubts whether the participant put thought into answering the item.

Parental burnout severity: Parental Burnout Assessment

Parental burnout severity was assessed at the start of data collection via the Parental Burnout Assessment (PBA; Roskam et al., 2018). This questionnaire includes questions about the four dimensions of parental burnout (emotional exhaustion, emotional distance, feeling fed up and a sense of contrast with the previous parental self), all answered on frequency Likert scales from 0 (*never*) to 6 (*every day*).[6] The number of items in each subscale and example items are as follows: emotional exhaustion (9 items; e.g., *When I get up in the morning and have to face another day with my child(ren), I feel exhausted before I've even started*), emotional distance toward child(ren) (3 items; e.g., *I'm no longer able to show my child(ren) how much I love them*), feeling fed up (5 items; e.g., *I can't take being a parent anymore*), and a sense of contrast with the previous parental self (6 items; e.g., *I'm ashamed of the parent I've become*). Scores could range from 0 to 138, with clinical cut-offs established by Brianda et al. (2023): 52.7 for moderate parental burnout severity and 86.3 for severe. We used the summed total score as an index of parental burnout severity.

Parenting affect: Daily diary survey

Daily parenting experience for each of the three fluctuating parental burnout features was assessed via a daily diary item. These three items read as follows: "I felt exhausted while caring for my children" (emotional exhaustion); "I felt close to my children in both good and bad situations (emotional distance; a reversed item); and "I felt overwhelmed caring for my children" (feeling fed up). All items were answered on a sliding scale from "Not at all" to "Absolutely" (on a 0-100 scale for Datasets 1 and 2, and on a 0-10 scale for Dataset 3). The development procedure and psychometric properties of these items, as well as their translation between English and French, are described extensively in previous research (Blanchard et al., 2023).

Recruitment and procedure

Dataset 1 (Belgium: General population)

Parents attended a one-on-one briefing with a researcher to go over the study procedure and daily diary items. The parent answered a few baseline questionnaires afterwards, and then began the daily diary items (including some not covered in this study) the next day. Daily diary items were randomized, and set up via the *formr* software (Arslan et al., 2020); further details are described in Blanchard, Hoebeke, et al. (2023). Over eight weeks, parents answered daily items in the evening, sometime between 6PM and 9PM (at a time of their choosing), with five hours to respond. Participants took a median time of 1.32 minutes to complete the daily diary. Parents received up to 35€ if they answered over 80% of daily diary prompts. All parents

[6] Note that the US sample only saw the scale values (e.g., "never") and not any numeric values, but Qualtrics assigned a value from 1-7 to each option. We therefore recoded data from the US sample from 0-6 during preprocessing so that it matched the Belgian sample.

provided written informed consent, and the project received approval from the UCLouvain biomedical review board (approval date: May 11, 2020; protocol title: PBNET). The eight weeks of data collection ran from April through June of 2021.

Dataset 2 (Belgium: Parental burnout)

The procedure for Dataset 2 was identical to Dataset 1, except that the recruitment targeted a specific population: parents with severe parental burnout. These participants were recruited by contacting all psychologists and clinical workers in French-speaking Belgium that had completed training on parental burnout and requesting that those who agreed transmit information about the study to any relevant patients. Recruitment also took place via local relevant social media groups (e.g., about parental exhaustion and burnout). Recruitment criteria also included feeling exhausted and overwhelmed as a parent. Parents received up to 50€ if they answered over 80% of daily diary prompts. Further information can be found in (Blanchard, Hoebeke, et al., 2023a).

All parents provided written informed consent, and the project received approval from the UCLouvain biomedical review board (approval date: May 11, 2020; protocol title: PBNET). In total, data collection spanned from March 2022 to March 2023. Participants took a median time of 1.38 minutes to complete the daily diary.

Dataset 3 (US)

The US sample began data collection in February 2023, and only participants that had finalized their data collection by July 2023 are included in this study. Participants did an introductory video call to discuss all study procedures. Parents answered end-of-day daily surveys for 21 days. These surveys were sent each evening at 8PM, and parents had three hours to answer[7], with reminders after one and two hours. Surveys were administered via Survey Signal, which administered a text message (and potential reminders) containing a link to the survey on Qualtrics. For these daily surveys, parents took a median time of 3.1 minutes to answer. This survey included first the three daily parenting affect items, and then up to 29 additional items, depending on if they said 'yes' to certain questions (e.g., about screen media use).

This daily survey was collected as part of a larger study focused on families' media usage and parents' experiences. This larger study included experience sampling (i.e., 5 signals sent at random times throughout the day) and a time diary at the end of the day regarding their children's daily activities and media use. Parents received compensation for their participation: up to $65 USD in the form of an Amazon gift card, depending on how many tasks they completed for the study (i.e., enrollment, baseline survey packet, end of day surveys, experience sampling prompts, etc). For the present study, we only use the parents' responses to the PBA completed at baseline and their answers to the three daily parenting affect items collected over

[7] Note that the time window increased to six hours partway through data collection to allow parents additional time to answer the surveys.

the next 21 days. All parents provided written informed consent, and the project received approval from the IRB at Georgetown University (GU IRB protocol # 2017-0591).

Data analysis

We performed all analyses using R statistical software (v4.3.0; R Core Team, 2021). Throughout our code, we made use of tidyverse packages (Wickham et al., 2019). We preregistered all analyses and main R code (https://osf.io/hzqe7), and we mention in footnotes where we deviated from this preregistration.

Affective dynamic indices

Variability: Relative SD

To operationalize variability, we computed each participant's relative standard deviation (Mestdagh et al., 2018) for each parenting affect. Relative variability is defined as "the ratio of the variability divided by the maximum possible variability given the mean [...] resulting in a variability measure that is not confounded by the mean" (Mestdagh et al., 2018, p. 691).

Inertia: Autoregression

To operationalize inertia, we used the autoregressive slope: we regressed each parent's affect from one day onto their affect from the previous day (Koval et al., 2021), computing an autoregressive coefficient for each parenting affect variable separately. We did so using a multilevel autoregressive model of order 1, or AR(1) model, to take into account that parents are likely similar to one another; this model assumes that participants come from a common distribution. However, since we have two very distinct samples, we generated two separate multilevel AR(1) models: one for each sample, so that we did not statistically assume the two samples came from the same population.

Autoregression coefficients typically range from 0 to 1.[8] A larger autoregressive coefficient represents greater inertia (i.e., responses that are more stable and staying closer to the mean value from one day to the next), while a coefficient close to zero signals that the previous day's response has little impact on today's response.

Covariation: ICC

To operationalize covariation between daily ratings of exhaustion, distance and feeling fed up for each participant, we used the intraclass correlation coefficient (ICC), which examines the degree to which ratings on different variables converge (Dejonckheere et al.,

[8] Autoregressive coefficients can range from -1 to 1, but a negative autoregressive coefficient is unusual for most psychology variables, as it represents a variable that is rapidly cycling—for psychology variables, this is usually related to intake (e.g., smoking or eating disorders). Technically, an autoregressive coefficient can also be greater than 1, but this would mean that the responses are constantly increasing (i.e., violating the assumption of stationarity; Koval et al., 2021).

2019).[9] Specifically, we calculated the ICC measuring consistency (P. E. Shrout & Fleiss, 1979) between the three parenting affect measures, following previous authors examining covariation of ESM measures (Dejonckheere et al., 2019; Erbas et al., 2014; Ottenstein & Lischetzke, 2020). We chose the consistency ICC because it considers correlations among variables but not mean levels, and we sought to differentiate affect dynamics from within-person means as much as possible. ICC scores range from 0 to 1, with a high score reflecting high convergence between ratings. Note that negative ICC scores can occur, but their interpretation is not clear in the context of emotion differentiation or covariation (Dejonckheere et al., 2019). We had preregistered that we would set to zero any negative ICC scores (following Dejonckheere et al., 2019), but all participants had positive ICC scores.

Preprocessing

To ensure data is on the same scale for both samples, we divided the responses from Datasets 1 & 2 (collected on a 0-100 scale) by 10, so that they matched the responses from Dataset 3 (collected on a 0-10 scale).[10] We also removed participants that had no variability (i.e., SD = 0) on any of the three daily affect questions, as their intensive longitudinal data must have variability to compute relative SD (Mestdagh et al., 2018).[11]

Regression models

We checked assumptions for the regression models with the performance package in R and the 'check model' function (Lüdecke et al., 2021). We found no evidence of outliers, nor collinearity (for models without interactions), but almost all models were non-normal and heteroscedastic. Following our preregistration and the method of Dejonckheere et al. (2019), we performed a non-parametric bootstrap procedure, randomly resampling participants (5000 times), fitting the regression model to each resample, and thereby deriving a sampling distribution and p-value. Models with raw data are reported in Appendix G (see Figure S35).

As we are interested in the predictive value of each affective index while controlling for the others and for the mean-value, we performed our analyses in two steps. First, we ran a regression model with only the intra-individual means of the three daily parenting affect variables (i.e., exhaustion, distance, and feeling fed up) predictive parental burnout severity. This model therefore had three predictors of interest. Since Dataset 3 had multiple important characteristics that differ from Datasets 1 and 2 (i.e., different scale, different country, different

---

[9] Note that intraclass coefficient (ICC) is also used as an affective dynamic index to examine "emotional granularity" or "differentiation" for researchers interested in the distinction between emotional experiences (the inverse of covariation; Dejonckheere et al., 2019). In this case, they would interpret the ICC inversely, with low ICC scores representing high emotion differentiation.

[10] All affective dynamic indices are already on a common scale (0 to 1 for ICC, around 0 to 1 for autoregression, and 0 to 1 for relative SD).

[11] We had preregistered including all participants in analyses and then removing those with no variability for a sensitivity analysis, but since variability is required to compute relative SD we only included in the main analyses participants who had variability for all three daily parenting affect variables.

age range of children, as summarized in Table 14), we also included a dummy variable representing sample (coded as Belgium: 0 or 1). We then ran a second regression model, where we added all affective dynamic predictors (i.e., inertia and variability for each of the three parenting affect variables, as well as their covariation). We again included a dummy variable representing sample, as well as interactions between this dummy variable and all affective dynamic predictors. This allowed us to directly examine if sample type impacted how affective dynamic indices predict parental burnout. This second model therefore included 19 predictors in total, of which 15 were of interest (the dummy variable, the seven dynamic affect predictors, and the seven interactions between the dummy variable and dynamic affect predictors).

We first examined the change in explained variance (i.e., R squared) between the two models, to examine if the dynamic indices predicted additional variance in parental burnout severity compared to just the mean levels of parental affect. We then examined the specific dynamic affect predictors (and their interactions with the dummy variable representing sample type), to assess which particular dynamic affect indices predict parental burnout severity. To correct for multiple testing and control the false discovery rate, we used the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995) on all beta coefficients in the regression models.

Power (post-hoc for secondary data analysis)

As part of the preregistration for secondary data analysis, we estimated a likely effect size (medium to large) and used G*Power to conduct post-hoc power analyses (i.e., knowing we would have access to at least 170 participants across the three datasets) to estimate the likely statistical power (Erdfelder et al., 2009) in identifying a change in R squared between the two models. As such, with an alpha of 0.05, a conservative sample size estimate of 170, and a medium effect size ($f2 = .15$), the power is .89. With a large effect size ($f2 = .32$), the power increases to 0.99.

Sensitivity analyses

As there were multiple ways we could have computed the affective dynamic indices and combined the samples, we also conducted a series of sensitivity analyses to examine what impact alternate decisions could have on the results. We decided (during data analysis) to compute a sensitivity model for all possible combinations of affective indices described below, yielding a limited multiverse analysis of sorts (Steegen et al., 2016). We did so to better understand the impact of affective index operationalization and preprocessing choices on results.

Inertia: AR(1)

We also wanted to examine inertia in isolation, instead of deriving autoregressive coefficients from a multilevel model, which brings the assumption that subject-level parameters from a common normal distribution (while also handling missing data better). We therefore also computed autoregressive indices per variable per participants using an autoregressive model of order 1: an AR(1) model.

Inertia: mlVAR

Since the three parenting affect items are at least partially related, we also derived autoregressive coefficients from a multilevel vector autoregressive model with all three variables, where each variable at timepoint *t-1* predicted itself and all other variables at time *t* (Epskamp et al., 2019). Using such a multilevel model allows us to account for associations with the other parenting affect variables.[12]

Variability: SD

We also sought to examine how using the typical standard deviation (SD), and not the relative version as proposed by Mestdagh et al. (2018) that removes the overlap with the mean, impacted results. To ensure that the SD would be on the same scale across samples (i.e., Datasets 1 & 2 with a scale of 0-100 and Dataset 3 with a scale of 0-10), we first divided all parental affect responses for Datasets 1 & 2 by 10 so that they were on the same scale as Dataset 3 before computing SD.

Sample: Separate models per model type

Following our preregistration, since the main model showed significant interactions of predictors of interest with a dummy variable representing sample, we also conducted sensitivity analyses where all regression models were computed separately by sample (i.e., US vs Belgium).

Exploratory analysis

We also wanted to examine whether the mean level of a parental affect variable moderates its associated affective dynamic index. For example, it could be the case that parents with low relative variability (i.e., few fluctuations) also have low overall mean levels of daily exhaustion, distance and feeling fed-up, and therefore low relative variability would not be predictive of parental burnout severity for these parents. This has previously been demonstrated, with negative affect mean levels moderating the associations between negative affect variability and depressive symptoms (Maciejewski et al., 2022). We thus also included an additional exploratory (preregistered) regression model, where we added an interaction between each

[12] Note that this was the original main method we had preregistered to use to operationalize inertia. However, when computing autoregressive indices with the mlVAR model, we received an error message for the US sample, as 91 subjects had less than 20 measurements which could lead to biased estimates, most notably for autoregressive estimates (as can be seen in Figure 4). We therefore still include autoregressive estimates from mlVAR models as sensitivity analyses, but focused our main analyses on multilevel AR(1) models, which have less parameters and so required less timepoints (and in Figure 4, less autoregressive values are negative for the multilevel AR(1) model). Ideally, however, we would still have substantially more timepoints per person.

affective dynamic index and its associated variable's mean (thus for all affective dynamic indices except ICC, which is associated with all three variables).

Transparency and openness

We report how we determined our sample size, all data exclusions, all manipulations (i.e., note) and all measures in the study. The experience sampling data from Datasets #1 and #2 (i.e., the Belgian sample) are available on the Open Science Framework with their original projects: https://osf.io/pshdn/ and https://osf.io/9zgbw/. Dataset #3 is not yet publicly available. The study's preregistration, all R analysis code, and the supplementary materials can all be found on the OSF page for this project: https://osf.io/5nfbu/.

9.3 Results

Descriptive statistics

For the Belgian sample, the mean compliance rate for daily parenting affect surveys was .93 (equivalent to around 52 surveys answered per person), while for the US sample, the mean compliance rate was .82 (equivalent to an average of 17 daily surveys answered per person). Figure 1 contains the mean and standard deviation of the three daily variables per sample, and Figure 2 contains the distribution per sample of parental burnout scores.

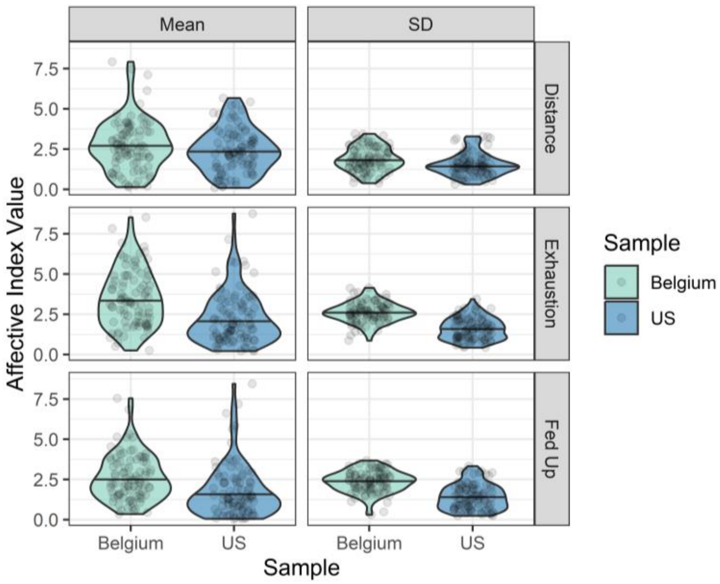


Figure 21. Violin Plots of the Mean and SD of Each Daily Variable per Sample

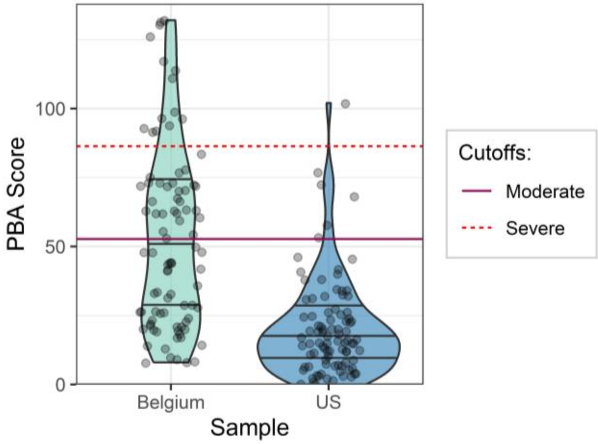


Figure 22. Violin plots of Parental Burnout Severity per Sample

Note. Cutoffs for Moderate and Severe parental burnout were taken from Brianda et al. (2023). PBA = Parental Burnout Assessment Score. Note that the Belgian sample includes both a community sample (Dataset 1) and a sample focused on parents with burnout (Dataset 2), while the US sample is a community sample (Dataset 3).

Main model: Bootstrapped regression model

For our main analysis, we first examined a regression model with only mean-levels of three daily parenting affect variables predicting parental burnout severity (see Model 1 in Table 15). This model explained a large percentage of variance in parental burnout severity, with an adjusted R^2 of .56. When we add in the affective indices as predictors (see Model 2 in Table 15), the model explains slightly more variance, with an adjusted R^2 of .61. When examining the predictors, however, only one predictor related to the affective indices was significant: inertia (i.e., autoregressive parameter) of Distance. Moreover, the interaction between the autoregressive parameter for Distance and dummy variable for sample was also significant, implying a different association across samples. However, when conducting separate regression models per sample (as preregistered as sensitivity analyses), neither the US nor the Belgian sample has any affective indices robustly[13] predicting parental burnout severity (see Figure S35 in Appendix G).

[13] In the Belgian sample, Inertia of Emotional Distance significantly predicts parental burnout severity in the two sensitivity models with mlVAR indices; no affective indices are significant predictors in the US sample.

Table 15. Bootstrapped regression models

	Model 1 (Only Means)			Model 2 (Affective Indices)		
Predictors	***Estimates***	***CI***	***p***	***Estimates***	***CI***	***p***
BE	**21.28**	**14.74 – 27.96**	**<0.001**	**-38.11**	**-73.73 – -0.54**	**0.048**
mean Exh	**5.35**	**0.96 – 8.51**	**0.017**	**4.67**	**0.81 – 8.09**	**0.019**
mean Dist	**6.69**	**4.51 – 8.79**	**<0.001**	**10.50**	**7.29 – 14.09**	**<0.001**
mean FedUp	1.84	-1.53 – 7.51	0.316	-0.08	-3.45 – 5.52	0.965
rsd Exh				-1.80	-27.11 – 22.67	0.897
rsd Dist				-8.18	-22.56 – 6.33	0.259
rsd FedUp				-3.22	-21.30 – 16.44	0.715
AR 1ml Exh				10.42	-27.99 – 47.12	0.574
AR 1ml Dist				**-254.13**	**-382.40 – -138.29**	**<0.001**
AR 1ml FedUp				20.13	-21.09 – 63.01	0.342
ICC				13.93	-2.08 – 30.12	0.088
BE × rsd Exh				35.52	-43.33 – 120.62	0.404
BE × rsd Dist				19.62	-22.39 – 60.82	0.361
BE × rsd FedUp				23.65	-37.36 – 80.13	0.430
BE × AR 1ml Exh				-36.37	-332.85 – 279.60	0.804
BE × AR 1ml Dist				**250.78**	**92.87 – 419.73**	**0.004**
BE × AR 1ml FedUp				91.74	-52.86 – 233.86	0.214
BE × ICC				-4.11	-36.25 – 27.38	0.803
Observations	180			180		
R^2 / R^2 adjusted	0.567 / 0.557			0.645 / 0.605		

Note. Text in bold represents a significant p-value. Since both models were non-normal and heteroscedastic, we followed our preregistration and performed a non-parametric bootstrap procedure, randomly resampling participants 5000 times, fitting the regression model to each resample, and thereby deriving a sampling distribution and p-value. We corrected for multiple comparisons within each model via the Benjamini-Hochberg procedure. BE = dummy predictor indicating sample (Belgium or US); Exh = emotional exhaustion; Dist = emotional distance; FedUp =feeling fed up; rsd = relative SD; AR 1ml = autoregressive parameter from a multilevel AR(1) model; ICC = intraclass correlation coefficient.

Sensitivity analyses: Bootstrapped models

As the main model included specific operationalizations of inertia (i.e., from a multilevel AR(1) model) and variation (i.e., relative SD), we conducted multiple sensitivity analyses where we varied the specific method to compute each index.[14] These are summarized in Figure 23, which shows bootstrapped regression models with varying operationalizations of affective indices.[15] Specifically, we examined inertia (i.e., simple autoregression model, multilevel autoregression model, or multilevel vector autoregression model), variation (i.e., relative SD vs. SD), and covariation (i.e., ICC).

[14] To be clear, we had preregistered conducting most of these sensitivity analyses, but we decided to extend them to a constrained multiverse analysis, where we would examine models crossing the different operationalizations per index, to gain a more robust view of how the choice of index operationalization can impact results.

[15] Sensitivity models without bootstrapped models and by sample are summarized in Figure S35.

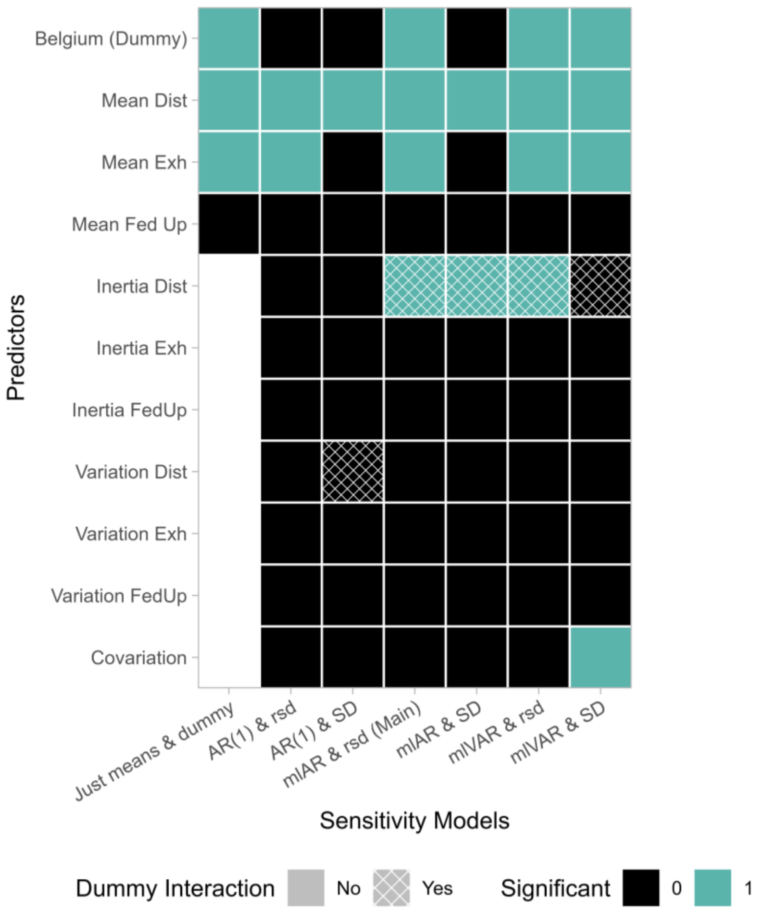


Figure 23. Summary of Significant Predictors Across Sensitivity Models

Note. Belgium (Dummy) = dummy predictor indicating sample (Belgium or US); Dist = emotional distance; Exh = emotional exhaustion; FedUp =feeling fed up; AR(1) = autoregressive model; rsd = relative SD; mlAR = autoregressive parameter from a multilevel AR(1) model; mlVAR = multilevel vector autoregressive model.

Across all models, the most robust predictor of parental burnout severity is Mean Emotional Distance. Among other predictors, Inertia of Emotional Distance (or its interaction with sample type) predict parental burnout severity in four out of six sensitivity analyses: all except those which operationalized inertia with a simple AR(1) model. Note that when Inertia of Emotional Distance is significant, so typically is its interaction with sample type. The distributions of different inertia indices do vastly differ by sample (see Figure 24), and for certain indices (i.e., particularly autoregressive parameters from the AR(1) model and the multilevel VAR model) the US sample includes many more negative autoregressive values (likely due in part to many participants having fewer than 20 timepoints, at least for the mlVAR model).

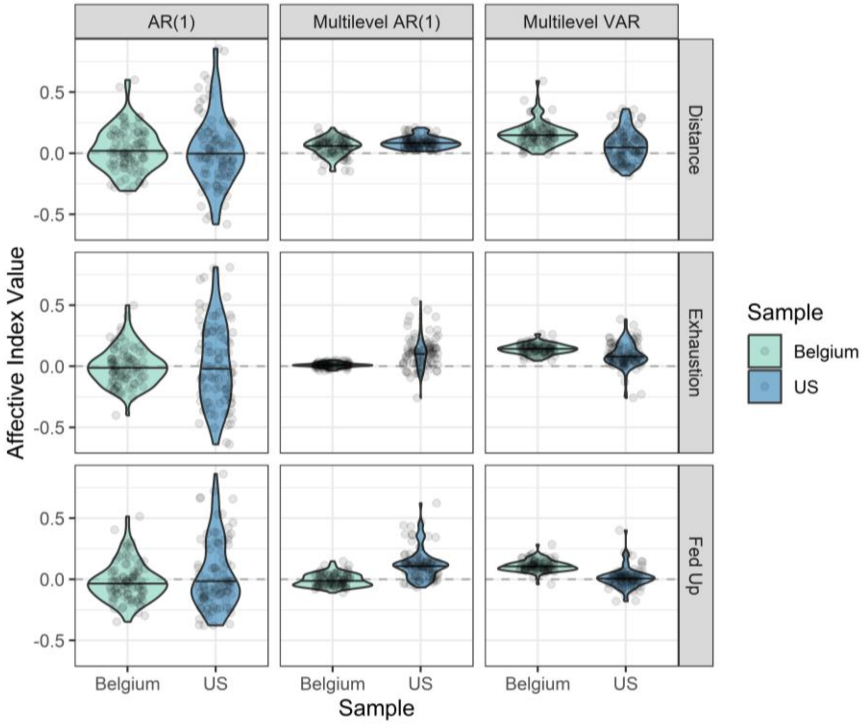


Figure 24. Distribution of Autoregressive Values across Operationalization Type and Sample

Note. Distance = emotional distance; Exhaustion = emotional exhaustion; Fed Up =feeling fed up; AR(1) = autoregressive model of order one; mlAR = multilevel autoregressive model of order one; VAR = multilevel vector autoregressive model of order one.

Amongst sensitivity analyses in Figure 23, Mean Emotional Exhaustion significantly predicts parental burnout severity mainly in models where variation is operationalized as relative SD (i.e., less dependent on mean levels). Lastly, Covariation predicts parental burnout in only one sensitivity model, as does the interaction of Variation of Distance with Sample Type.

Exploratory Model: Moderation by the mean

Since the dummy variable representing sample was significant in previous models, we followed our preregistration and examined possible moderation by mean level separately by sample. Across both samples and estimation types (regular vs. bootstrapped), the only significant predictors were Mean Emotional Distance and the interaction of Mean Emotional Distance and Variation (i.e., relative SD) of Distance in the U.S. sample (see Figure S35 in Appendix G), perhaps due to the low levels of parental burnout severity in the U.S. sample. The mean levels therefore do not appear to moderate the relationship between affective indices and parental burnout severity.

9.4 Discussion

The results emphasize the key role Emotional Distance—both its average intensity and its persistence across time—plays in parental burnout severity. Indeed, across sensitivity models, the most important predictor of parental burnout was the average intensity of how emotionally distant a parent feels from their children from day to day. We had predicted that all affective indices would predict parental burnout, but the only affective index that predicted parental burnout severity across most of sensitivity models was inertia of Emotional Distance. Indeed, this index (or its interaction with sample type) was a significant predictor in models where it was operationalized as a multilevel autoregressive model of some type (with or without additional covariates; see Fig. 3). Perhaps this is because multilevel autoregressive models have greater predictive accuracy by avoiding overfitting due to their multilevel nature (i.e., shrinking estimates to the population-level mean; Lafit et al., 2022). If future research further supports the role inertia of emotional distance plays in parental burnout severity, this would imply that parents who report severe burnout are more likely to get "stuck" in feeling emotionally distant from their children, which should therefore be specifically targeted in clinical research and interventions.

The other two parenting affect variables seem to play less of a role in describing parental burnout severity. For example, intensity of Emotional Exhaustion predicted parental burnout severity in some sensitivity models, but none of its affective indices were significant in any sensitivity model. Although we had expected affective indices of all three daily parenting affect variables—including Emotional Distance—to significantly predict parental burnout severity, these results fit with conceptualizations of burnout that view emotional exhaustion as kickstarting parental burnout, and emotional distance as maintaining it (Blanchard, Hoebeke, et al., 2023b; Roskam & Mikolajczak, 2021). Within this framework, temporal patterns of emotional exhaustion would not predict parental burnout severity but instead might play a role in its development. Feeling Fed Up, for its part, did not predict parental burnout severity in any form. Future research could focus on what specific role it plays in parental burnout, and if severe parental burnout is always characterized by the parent feeling fed up. And interestingly, only one sensitivity model included a significant predictor for covariation, suggesting that parental burnout is *not* characterized by the three symptoms consolidating and fluctuating simultaneously. This finding highlights the importance of investigating each component of parental burnout separately, without aggregating them by rote and potentially losing nuanced trajectories.

This study has clinical implications regarding the conceptualization and treatment for parental burnout. First, although we expected that parental burnout severity would be equally characterized by all three main features, it appears that they instead play different roles. From the sensitivity analyses in this study, emotional distance emerges as the heart of parental burnout severity. It is possible that exhaustion plays a role in the development of parental burnout (Blanchard, Hoebeke, et al., 2023b; Roskam et al., 2018; Roskam & Mikolajczak, 2021), and so is not robustly related to parental burnout severity or maintenance. A common treatment for burnout is distance from the provoking context to allow time and space to rediscover an ideal balance between stress and restoration, such as via an extended sick leave from work for job burnout (Van Dam, 2021). Although such physical removal from parenting is more complicated, parents typically report feeling like they have more resources the day after they report feeling distant from their children (Blanchard, Hoebeke, et al., 2023b). Thus, emotional distance might be a way for the parent suffering from burnout to attempt to find the distance they need to recover and rest, since a leave of absence from parenting is less possible than from a job. However, emotional distance can impact the children's well-being, particularly if the feelings of emotional distance persist long enough (Masten & Palmer, 2019). Emotional distance also creates repercussions throughout the family unit if the parent does not recover (Kerig, 2019; Mikolajczak, Raes, et al., 2018). As such, relevant interventions should consider ways to provide parents with time to recover apart from their children, such as through additional free childcare, and future research should investigate other coping strategies parents utilize to get small amounts of rest, such as media use.

Emotional distance also relates to the values of parenthood in European and American cultures, which particularly centers outward demonstrations of emotional warmth (Bornstein, 2012; Cheah et al., 2015; Dupont et al., 2022). Indeed, when parents were asked to name five adjectives to describe an ideal mother or father, words relating to warmth (e.g., "empathy", "loving", "tenderness") emerged to describe an ideal parent (and even more so an ideal mother), particularly from the perspective of European and American parents (Lin et al., 2023). Parental burnout also appears more prevalent in many European and American countries than Asian or African ones (Roskam et al., 2021). Perhaps this contradiction between a cultural imperative to feel and behave with warmth toward your children but an actual experience of emotional distance (potentially a short-term protective mechanism for the parent) could be at the root of the experiences of shame and suffering linked with parental burnout (Hubert & Aujoulat, 2018), and help explain its prevalence patterns across cultures. This is just a supposition based on the central role of emotional distance found in this study and the cultural context of the included samples, however. Future research should investigate in detail how emotional distance and parenting affect generally, as well as their role in parental burnout, vary depending on the cultural and geographic context.

This study also has important implications for research on dynamic affective indices. First, the results further support the importance of examining affective indices while controlling for mean levels (Dejonckheere et al., 2019). When examining the relationship between affective indices and parental burnout severity via bivariate correlations, many had significant associations. However, most no longer held when controlling for the mean level, nor when conducting sensitivity analyses. It can also be relevant to examine if the mean level moderates the relationship between an affective index and dysfunction, as discussed by Nook (2021), although we found no strong evidence for that in the present study (see Figure S35 in Appendix G).

Second, the results from the sensitivity models emphasize the key role of understanding what impact operationalization can have on results. For example, if we had only included a model with the multilevel VAR autoregressive parameter for inertia, and the SD parameter for variation, we would likely have concluded that covariation played a key role in parental burnout severity. However, no other model supported this conclusion. Another example includes operationalization of variation: if the variation index was independent from the mean (i.e., relative SD), then the mean-level of Emotional Exhaustion significantly predicted parental burnout severity. This could have occurred due to the computation and assumptions regarding the relative variability index[16], or perhaps the overlap between SD and mean-level of exhaustion played a role, or for some other reason. Regardless, this pattern of results further highlights how the specific operationalization of an affective index can impact results. Thus, in light of the nuances uncovered from our sensitivity analyses, we strongly encourage other researchers investigating affective indices to examine the robustness of their results with sensitivity analyses, particularly if there is no theoretical reason to choose one particular operationalization.

In addition, through the sensitivity analyses, we were able to see the importance of data parameters. For example, the number of timepoints clearly impacted the inertia values across samples: the US sample answered a maximum of 21 daily surveys, which resulted in many more negative autoregressive values (i.e., difficult to interpret for psychology research; Koval et al., 2021). Although previous research has used far fewer timepoints to examine inertia (e.g., 10 timepoints from a laboratory emotion rating task; Koval et al., 2013), guidelines on appropriate data parameters (i.e., number of timepoints and sample size, response type) are beginning to emerge. For example, Pirla and colleagues (2022) suggest a minimum of 40 timepoints to robustly recover inertia (although they examined autocorrelation), and less to recover indices not tied to fluctuations (i.e., variation). More comprehensive guidelines on what type of data and which operationalizations can be used to compute affective indices is still needed, however.

This study has limitations. In particular, although the sample size was sufficient based on post-hoc power analyses, we only reached this sample by combining multiple datasets which ended up including important differences. These differences included geographic location, language, and stage of parenting (i.e., parents of children 1- 2 years old for the US sample, and parents of children under 17 years old for the Belgian sample), and it is not clear whether differences between samples specifically related to any of these distinctions. The number of timepoints included in the study procedure also differed between samples (maximum 21 vs maximum 56). This last point is perhaps most unfortunate, as we know that it impacted estimation of inertia indices (in particular, of multilevel vector autoregression indices, which typically require minimum 20 timepoints but many participants in the US sample answered fewer timepoints), and perhaps of other indices as well. Another key difference between the two samples involves the distribution of parental burnout severity: the Belgian sample

[16] For example, the relative variability index assumes the participant considers the full span of the given scale when responding (Mestdagh et al., 2018). This is potentially not the case for some of the daily parenting affect variables, particularly Emotional Distance (see Figure 1), such as if parents feel uncomfortable endorsing to the maximum part of the scale due to cultural beliefs involving parenting.

included parents spanning the entire parental burnout scale (due to sampling strategy, as one dataset specifically targeted parents with burnout), whereas the US sample mainly had lower parental burnout, as it did not oversample for parental burnout (see Figure 22). These two points in particular—differences in number of timepoints, and in distribution of parental burnout scores—likely contributed to the many significant interactions of affective index predictors with sample type. Although we investigated regression models in both samples separately, no affective indices significantly predicted parental burnout, perhaps due to lower power when the samples were tested independently. Lastly, all data was collected during or shortly after lockdowns related to the Covid pandemic, which impacted parenting experiences and parental burnout (Kerr, Fanning, et al., 2021), but it is not clear how they specifically impacted these samples. Future research should combine intensive longitudinal designs with more conventional longitudinal studies to track cohort effects and disentangle cumulative effects of untreated burnout.

Conclusion

Average intensity of emotional distance and its inertia (i.e., persistence across time) robustly predicted parental burnout severity across multiple sensitivity analyses. Average intensity of emotional exhaustion also predicted parental burnout severity across most sensitivity analyses, but no other variable or temporal pattern did. Researchers and clinicians should therefore prioritize these aspects in the treatment of parental burnout. This study also highlights the importance of methodological robustness when examining affective indices, including controlling for mean levels and examining the impact of operationalization on results.

Section 4.

General discussion

Chapter 10.

Discussion

10.1 Summary of main findings

This dissertation aimed to investigate the experience of parental burnout as embedded in the family context, using mainly a network approach.

In **Chapter 4**, two types of network models (both reanalyzing the same dataset of N=1551 parents) indicated emotional distance as a central variable, as well as the variable with the strongest partial associations with neglect toward children. Results did not change when comparing mothers to fathers or based on the number of children. Note that I conducted a second study along with other research collaborators using the same methods but with the Parental Burnout Assessment to measure parental burnout (developed based on the qualitative experiences of parents in burnout, instead of adapted from occupational burnout), reanalyzing data regarding parental burnout and child maltreatment from 3144 parents (Kalkan et al., 2023)[1]. This study suggested emotional exhaustion as driving activation in the network, while emotional distance remained the main connection to neglect toward children.

Chapter 5 detailed the development of 11 experience sampling items to assess parental burnout features, interactions with the partner, children, and broader family environment. Parents' responses to these items showed sufficient variance, meaning the items successfully captured daily variations in parenting experiences. Additionally, items showed convergent validity by correlating with retrospective questionnaire scores on parental burnout, and divergent validity by weakly or not correlating with retrospective scores for depression, anxiety, and stress. **Chapter 6** surveyed typical research practices (including how researchers collected, processed and analyzed data) related to temporal network analyses via a scoping review. These two methodological chapters laid the groundwork for the last section of the thesis: empirical investigations into how parental burnout features fluctuate.

Chapter 7 presented a temporal network analysis of parents from the general population (focusing on 43 mothers), examining how daily exhaustion, distance and feelings of being fed up interrelate with interactions with the partner, children, and broader environment. Emotional exhaustion was the only parental burnout variable to self-predict itself from one day to the next (i.e., to have a significant autoregressive coefficient), and it also predicted finding

[1] Note that this article is included as Appendix A. It has shared first authorship, by both Rana Begum Kalkan and M. Annelise Blanchard.

children difficult to manage, which in turn predicted feeling fed up. In this network of parents from the general population, emotional distance only connected to one other variable: sharing positive moments with children. **Chapter 8** presented the same methods and analyses but with a sample of parents in burnout (focusing on 36 mothers). In the resulting temporal network, all three parental burnout features self-predicted (show significant autoregressive coefficients) from one day to the next. In addition, the parental burnout features showed few associations with other variables—except for interactions with children, which strongly predicted emotional distance and feeling fed up.

Lastly, **Chapter 9** reanalyzed the data from Chapter 7 and Chapter 8 (N=90), along with data from research collaborators in the U.S. (N=90), to examine the affective dynamics of parental burnout. In particular, this chapter examined if the *way* emotional exhaustion, emotional distance, and feeling fed up fluctuate characterizes parental burnout severity. Results indicated that only inertia (i.e., persistence, operationalized via autoregressive models) of emotional distance predicted parental burnout severity.

10.2 Theoretical implications for parental burnout

When compiling all of these results, three conclusions immediately emerge:

First, each feature of parental burnout plays a distinct role: every study demonstrates that different features have distinct associations with each other and with family-related variables. In addition, the three features of parental burnout that fluctuate in time do not co-vary together to indicate parental burnout severity, as demonstrated with the affective index for covariation in Chapter 9. Until now, relatively few studies examine parental burnout not (only) as a unitary construct but via its distinct features (exceptions include Le Vigouroux et al., 2022; Roskam & Mikolajczak, 2021). However, in light of the findings in this dissertation, solely focusing on parental burnout as a unitary construct likely hinders understanding of how each feature might change in interaction with the family context and during different stages of parental burnout (i.e., development or maintenance). Therefore, future research should routinely investigate parental burnout via its different components, to pin down their possible different roles.

Second, to understand parental burnout—as well as the parenting experiences of any parent—requires situating it within its context: the family. Indeed, all chapters that empirically investigate parental burnout demonstrate that the experiences of parental burnout relate to and fluctuate with interactions with the partner, children, and family environment. Moreover, when compiling the results from Section 3 on Patterns across time, parents seem less influenced by and responsive to their family context when in a state of severe parental burnout—with the key exception of interactions with their children. The family environment therefore centers the experience of parental burnout: the family provides the context in which parental burnout develops and persists, and *how* parenting experiences relate to the family context appears to change during the progression of parental burnout.

Lastly, like with most clinical phenomena, parental burnout exists on a continuum. Indeed, since parents from the general population show daily fluctuations in their levels of emotional exhaustion, emotional distance and feeling fed up (in both Chapter 5 and Chapter 7), clearly the features of parental burnout resonate for all parents. Indeed, just as most clinical phenomena involves general processes that have become problematic through their changes in frequency, intensity, or duration (Wright & Zimmermann, 2019), so does parental burnout involve general parenting experiences but with greater intensity and different temporal patterns. Therefore, although severity thresholds (e.g., Brianda et al., 2023) hold important clinical utility, research should investigate parental burnout along all aspects of the continuum: parents with little overall exhaustion or distance, parents beginning to feel exhausted, and so on. This would allow a nuanced vision of parenting processes and how they might change across the spectrum of parental burnout. This thesis accomplishes such a continuum-based view for a few specific aspects of parental burnout, such as demonstrating that emotional distance persists longer and is less influenced by the family environment the more severe the parental burnout.

Next, I shall discuss the roles of specific features of parental burnout.

Emotional exhaustion

Within the burnout literature, researchers accept emotional exhaustion as the core dimension of burnout (Halbesleben & Bowler, 2007). This is partly because it ties otherwise disparate definitions of burnout together: from definitions with three distinct dimensions anchored to the work context (Maslach et al., 2010), to two dimensions (i.e., exhaustion and disengagement) focused on the work context (Demerouti et al., 2001), to only one dimension (i.e., exhaustion) not restricted to any context but instead context-free (Pines & Aronson, 1988). In addition, several studies suggest that emotional exhaustion is the first step in the burnout process (Lee & Ashforth, 1993; Leiter, 1993; Leiter & Maslach, 1988). Parental burnout research also supports this: Roskam and Mikolajczak (2021) examined two separate cross-lagged datasets, each with three waves months apart. They found that parental burnout starts with emotional exhaustion, which then contributes to emotional distance at the next wave.

Findings from this dissertation offer further support. First, emotional exhaustion plays a central role in both cross-sectional network analyses with parents from the general population in Kalkan et al. (2023), suggesting that it sparks parental burnout. (Note that this was not found in the cross-sectional network analyses in Chapter 4, which identify emotional distance as central to the network; but this network analysis used the older conceptualization of parental burnout which did not include feeling fed up or a sense of contrast as dimensions, and where emotional distance was adapted from the MBI construct of depersonalization). Second, in Chapter 7 on unselected parents, only emotional exhaustion among the three parental burnout features sports a self-predictive autoregressive loop, suggesting that it is the only feature with inertia (i.e., persistence) for unselected parents. In addition, exhaustion predicted finding children difficult to manage in the temporal network, which in turn predicted feeling fed up (and these variables form a feedback loop when all parents, including n=4 fathers, are included in the network analysis in Appendix E; see Figure S21). This again **supports exhaustion as the parental burnout variable most likely to kickstart the development of parental burnout.**

These findings stand out from previous research since they take the family context into account: exhaustion appears as particularly impactful for unselected parents even when taking into account their interactions with their children, partner, and wider context. However, further research on exhaustion's role in the development of burnout remains needed (ideally while also taking the family context into account), particularly longitudinal research that follows large cohorts of parents and tracks which parents develop parental burnout, to robustly assess which features of burnout (and of the family environment) predict its development.

Emotional distance

From the findings in this dissertation discussed above, emotional distance seems implicated in parental burnout severity. In Chapter 7 examining parents from the general population, when parents feel distant from their children, they tend to perceive regained resources the next day (see Figure 18). This fits with occupational burnout literature where distance from the stressful context is prescribed as treatment, to allow time and space to recover (Van Dam, 2021). Although such physical separation from the stressful context is more complicated for parenting than for work contexts, parental burnout researchers have similarly posited that detachment from children might serve to “protect parents from negative affect and cognitions about parenting,” as theoretically proposed by Roskam and Mikolajczak (2021). In other words, exhausted parents might emotionally distance themselves from their children in an attempt to recover and rest when they have few other options. From the findings in Chapter 9, it seems likely that parents who have severe burnout end up “stuck” in a state of emotional distance toward their children, while other parents may feel distant here and there but not consistently over time. This emphasizes that emotional distance, and particularly its persistence, characterizes parental burnout.

What exactly is emotional distance, though? Lubiewska and colleagues (2023) investigate this question (via the term "distanced mothering") across several cultures: distance in parenting is primarily associated with a lack of warmth (poor expression of positive emotions including love; also low support/closeness/attention) and low involvement in daily activities. Dutch mothers viewed this type of parenting as neglectful and shameful, but mothers in Turkey and especially Poland viewed it through a lens of strict, rule-oriented parenting, sometimes framed as a strategy to foster independence. The authors interpret this distinction as related to a traditional-modern cultural continuum, with traditionalism in family contexts linked with societal acceptance of hierarchy, order, authority; patriarchy (power relations embedded in male and older family members); and cultural endorsement of restraint over indulgence. They also frame results from Poland in the context of a cultural prioritization of fostering early childhood independence, particularly in the post-Communist period. These results both indicate that the concept of distanced mothering (or presumably distanced parenting) can vary by culture, based on its cultural acceptance and how people interpret it: as uninvolved and disengaged parenting (perhaps to the extent of “mild” neglect), or as parents who care about their children but in a hands-off and control-oriented way. Thus, emotional distance in parenting is imbued with different meaning in different cultures, depending on the values and needs of the specific culture: in some cultures, it is viewed as normative and supporting child development, while in others it is viewed as harmful and neglectful.

In addition to the cultural aspects detailed by Lubiewska and colleagues (2023), another reason Dutch mothers likely interpret distanced parenting as neglectful and shameful is due to European and North American parenting values prioritizing demonstrations of emotional warmth (Bornstein, 2012; Cheah et al., 2015; Dupont et al., 2022). This importance placed on expressions of warmth is strengthened by ideologies of positive parenting embraced by many Western cultures, which view parenting warmth and support as essential components of good parenting, often accompanied by a rejection of any disciplinary consequences (Larzelere et al., 2017). The Council of Europe even codified this emphasis on emotional warmth in parenting in their "Policy to Support Positive Parenting" (Rodrigo, 2010). This emphasis on emotional warmth as central to parenting for some cultures but not all can also be seen when assessing parents' beliefs about the ideal parent (Lin et al., 2023). Parents in Western European and North American countries emphasize words relating to warmth such as "empathy", "love", and "kindness"; although some of these terms also come up in other cultures' good-parent beliefs, they are less homogeneous (e.g., not appearing together & not appearing across income levels). As such, parenting warmth appears as a critical facet of "good" parenting in many Western European & North American countries, but not around the world. Thus, for parents in cultures centering warmth as a cornerstone of good parenting, they will likely view lack of warmth, or emotional distance from one's children, as an incontrovertible sign of "bad" and neglectful parenting.

These cultural parenting norms govern not just behavior (e.g., external actions) but also emotions (e.g., internal experiences; Vishkin & Tamir, 2023). For example, parents in France, Belgium, and Luxembourg report feeling that they should show positive emotions while interacting with their children, including specifically "love," "joy," "happiness," "pride," "compassion," "attention," "enthusiasm," and "satisfaction" (Lin, Hansotte, et al., 2021). If individuals do not adhere to norms, their well-being can be impacted (Vishkin & Tamir, 2023), leading to negative emotions, and in particular intense feelings of shame (Schaumberg & Skowronek, 2022). Research on emotion norms within parenting similarly finds that not adhering to emotion norms relates to low well-being (Lin, Hansotte, et al., 2021). This is particularly the case when simultaneously *experiencing* emotions discordant with norms while also attempting to *express* emotions that are in line with norms (e.g., "parenting with a smile" while sad or angry, otherwise known as "surface acting"), which is linked with parental burnout (Lin, Hansotte, et al., 2021).

So what about parents who become exhausted when faced with the relentless stress of parenting, and start to feel emotionally distant from their children? Perhaps for some parents in collectivist cultures with more norms relating to family support, they can "rely more on caregiving support from their extended family members" (Lubiewska et al., 2023, p. 3) and take distance when they feel the need, spending time again with their children when they feel recovered. This in fact aligns with some preliminary cross-cultural research, which suggests that parental burnout rates are highest in countries with individualistic cultures (Roskam et al., 2021). And maybe parents in cultures that do not prioritize warmth as an absolutely essential component of parenting might feel distant for some time, but do not associate this with a perception of themselves as "bad" parents. But for parents who feel distant from their children, who do not have the resources to pursue time dedicated to rest or recovery, and who espouse parenting and emotion norms centering emotional warmth, they are faced with a devilish internal contradiction: they are stuck between a cultural imperative to both feel and behave with warmth toward their child and a reality of emotional distance. This discrepancy likely yields the intense shame parents experience when in burnout: they feel like they fail at (what they see as) the most important aspect of being a parent (Hubert & Aujoulat, 2018). This shame in turn isolates the parent and maintains their suffering. As Dallos and Draper (2015) discuss (for families dealing with psychosis, but equally relevant to parental burnout): "the extent to which families feel that they have failed, feel stigmatized by the illness, and feel alone and ashamed [are] key ingredients to maintaining the difficulties" (pg. 194). Thus, **emotional distance (particularly when occurring for parents in cultures prioritizing parental warmth) likely works to maintain parental burnout.**

Preliminary evidence from this thesis supports this hypothesis regarding the role of emotional distance maintaining parental burnout, particularly since daily intensity and inertia of emotional distance characterize parental burnout severity (see Chapter 9). However, further research should investigate this in greater depth, including via lengthy longitudinal studies across cultures prioritizing parenting warmth as well as those that do not. After all, this hypothesis suggests that the varied prevalence rates of parental burnout across the world (Roskam et al., 2021) are not due primarily to differences in community support in parenting, but instead to the presence and strength of parenting norms relating to parental warmth. Parents across cultures can feel exhausted (although perhaps to a lesser or greater extent depending on parenting pressures and on community support; Super & Harkness, 2020), but **perhaps the acute suffering and shame associated with parental burnout arises particularly from the discrepancy between cultural norms of warmth and the lived experience of distance.**

Is parental burnout distinct from occupational burnout?

This key role of distance distinguishes parental burnout from occupational burnout even more so than their differing contexts. Indeed, exhaustion is the core construct of occupational burnout, theoretically (e.g., Demerouti et al., 2001; Maslach et al., 2010; Pines & Aronson, 1988) and empirically (e.g., Halbesleben & Bowler, 2007; Lee & Ashforth, 1993; Leiter, 1993; Leiter & Maslach, 1988). For parental burnout, although emotional exhaustion seems implicated in its development, parents' suffering and shame appears closely tied to emotional distance (Hubert & Aujoulat, 2018). As hypothesized above, this might be particularly the case in cultures with strong norms involving parental warmth. This distinctiveness implies that the findings from parental burnout might not generalize to job burnout, for although they share a central construct, their differences in contexts also imply different cultural and emotion norms coming into play.

10.3 Clinical implications for parental burnout

Prevention

As exhaustion seems the first step to developing parental burnout, it should be targeted in prevention efforts. Exhaustion arises from unresolvable stress (Van Dam, 2021), and so prevention can focus on decreasing both exhaustion and parenting stress. This can be very concrete, in the form of additional parenting resources that are easy and free to access (e.g., childcare for all ages, groups to meet other parents, practical childrearing advice about not only safety but also structure and discipline; Larzelere et al., 2017; Masten, 2018; Webster-Stratton, 1997). These parenting resources could be implemented widely but require substantial government endorsement and financial support. Crucially, to be effective as broad prevention strategies to decrease burnout for parents generally, these types of parenting resources should not be needs-based or require parents to pay, but instead be freely and widely available to all parents. As discussed in the Discussion section on Emotional exhaustion above, intensive societal parenting standards also add to parents' stress and exhaustion. Rendering these parenting standards more realistic can also be a target for prevention of parental burnout. Although changing parenting standards would be difficult as it implicates many different fields (i.e., political, legal, and childcare-related domains, as well as social and popular media on childcare), a first step could involve practitioners who have regular contact with parents (e.g., healthcare providers, childcare providers, therapists) discussing realistic parenting standards with parents. For example, substantial research has shown that even one supportive and trusted adult figure in the lives of children boosts their resilience and well-being (Bellis et al., 2017). Thus, even if a parent is going through a hard time and is not able to be emotionally close to their child for a period of time, they can ensure that their child has other caring and trusted figures around them: another parent, a family member, a teacher, and so on. In other words, the parent is not the only person in a child's life, and if a parent needs a break, they can rely on others for a time.

As this dissertation inscribes itself within family systems theory, prevention can also involve boosting the resilience of a family as a whole. Indeed, researchers propose that instead of looking at resilience as pertinent to just one individual, resilience should be thought of as a multilevel system: including not just the family, but also the wider community, culture, government, laws, and so on (Henry & Harrist, 2022; Masten, 2018). All of these are interconnected, as all systems are always interlinked with others (Dallos & Draper, 2015). Therefore, to boost the resilience of a family, one can intervene directly on the parent (e.g., education regarding parenting strategies or discipline), on the family (e.g., interventions to improve family functioning, family cohesion, routines), but also much more broadly (Masten, 2018). For example, interventions at the community level could involve ensuring accessible parenting support, social groups, and education on child development and childrearing strategies (Webster-Stratton, 1997). At the cultural level, rituals and beliefs can impact how families interpret and make meaning of their daily lives (Masten, 2018). And at an even broader social level, ease of access and price of healthcare and housing, as well as the quality and availability of schooling for children of very young ages, all have direct and substantial impacts on family functioning (Masten, 2018). Thus, efforts to boost the resilience of family systems could operate at any and all of these levels.

Treatment

According to findings from this dissertation, treatment efforts should particularly focus on emotional distance and the associated shame parents feel. As inertia of emotional distance characterizes severe parental burnout (see Chapter 9), treatment efforts should focus on breaking this persistence. One possible mechanism behind inertia involves difficulty down-regulating or disengaging from negative emotional states once initiated (Koval et al., 2021). Applying this to parental burnout means that for parents in a severe state of burnout, they might "get stuck" in feeling distant from their kids and have difficulties getting out of this state. One way that appears effective at lessening parents' emotional distance, and potentially breaking them out of that state, involves their interactions with children; after all, these were the only aspects of the family context to impact the daily experiences of emotional distance for parents in burnout (see Chapter 8). Treatments for parents in burnout could thus focus on encouraging interactions between parent and child in a way that is not stressful, tiring or pressuring for the parent. Prior research supports that even small, spontaneous and everyday interactions (e.g., chatting while running errands) can be quality time for parent and child, as long as both parties enjoy the time spent together (Hsin, 2009; Kremer-Sadlik & Paugh, 2007).

In addition, parents often need distance from their parental role to recover (both physically and emotionally) from burnout. Indeed, treatment from burnout includes taking physical and emotional distance from the initiating context (i.e., often the workplace), such as through medical leave (Van Dam, 2021). This allows the individual to "restore a healthy balance between effort and rest" (pg. 737) for both their body and their mind, and to have space away from the unresolvable stress they faced, as well as time to recover from their exhaustion (Van Dam, 2021). Such distance can be more complicated for parenting, but free childcare (for children of all ages, including during evenings and weekends) would be an important resource, allowing parents to have time to themselves and (at least momentary) distance from their children, either for routine tasks such as grocery shopping or for taking time to enjoy their hobbies and otherwise rest.

Along with emotional distance, treatment should also target the accompanying shame and isolation. One way to do so can be via contact with other families with parental burnout, either informally or via some type of family group therapy: if parents can share their experiences with others living through similar experiences, this can lessen feelings of isolation, stigma, and shame. Such therapy in a group setting allows for more perspectives and advice for parents than just from the one "expert" voice of the therapist (Dallos & Draper, 2015, p. 194).

Indeed, the only study currently evaluating treatments for parental burnout specifically highlights that a group-based non-directive intervention, allowing parents in burnout to share their experiences in a space led by psychologists trained in active listening, was effective for treating parental burnout (Brianda, Roskam, Gross, et al., 2020). This also suggests that psychologists do not have to be experts in parental burnout specifically to deliver effective treatment and support for parents, if in a group therapy setting where parents can also offer support, advice, and understanding to one another. However, no research has yet specifically evaluated multiple group family therapy for parental burnout, even though this might be a particularly useful treatment tool (especially if it reunites families with different configurations), as families reflect they often can better observe processes in families different from their own (Dallos & Draper, 2015, p. 196). Dallos and Draper (2015) identify that key to this approach are a warm and easy-going atmosphere, humor, flexibility about rules, families being able to bring and share food, and reflection about what parts of the group felt safe. Cognitive behavioral family therapy can also help parents to "identify and interrupt dysfunctional child-related cognitions" (Dallos & Draper, 2015, p. 199), such as thinking they are a failure as a parent. Such therapy can help parents learn and employ relaxation techniques and cognitive coping statements, as well as teach parents general knowledge of behavioral principles, so that they can understand when and how parenting strategies could work, instead of just memorizing a list of techniques (Dallos & Draper, 2015, p. 194).

10.4 Methodological Implications

Reassessing concepts from a network approach

As discussed in Chapter 2, a network approach to psychopathology involves key tenets about the network's structure and organization and how these can be empirically investigated. In the following paragraphs, I discuss whether and how these concepts fit with a family systems approach.

Centrality

Adopting a network approach, adapted from other fields and therefore inheriting the developments from these fields (Barabási, 2016), also brings with it "a whole new toolbox to analyze the interrelations of symptoms" (Bringmann et al., 2019, p. 893). One such tool revolves around the structure of the network: identifying which node has particular structural importance to the network, otherwise known as centrality (see the Centrality section on page 29 for more detail). A central node typically denotes a node strongly connected to other nodes in the network, with the assumption that a central node has a greater impact on the entire network than a poorly connected node. Within a network approach to psychopathology (Borsboom, 2017), researchers hypothesize that central symptoms work to activate the network, and similarly can be targeted by interventions to turn off the whole network of symptoms (McNally, 2016). Within a family systems approach, some clinical researchers hypothesize similarly: that by focusing on one core part of the system, the entire system can be changed, since effects will echo and rebound throughout the system as a whole (Lebow, 2020).

However, the centrality indices used in psychology networks come from social network analysis, and doubts arise about how well they translate. In social networks (and, indeed, most other networks), connections between nodes are observable and reflect the raw data (e.g., emails sent, or gossip exchanged between people; power lines connecting power stations; mail delivered by mail carriers). Centrality measures can be conceptualized as capturing how something (e.g., emails, gossip, electricity, mail packages) propagates through the system (Borgatti, 2005). Different types of propagation include parallel flow (spreading simultaneously to many other nodes, like an email virus spreading to all contacts in a mailbox), serial flow (spreading from one activated node to the next, like a sexually transmitted disease), or transfer flow (moving from one node to the next, like a package passing from the mail carrier to its destination; Bringmann et al., 2019). And in social networks, a centrality measure designed for a specific type of flow process should be applied only to that type of network, in order to generate appropriate and accurate centrality estimates (Borgatti, 2005).

However, in psychology networks, connections between psychological constructs are unknown, and are thus estimated using statistical modeling techniques (e.g., mostly some type of correlation or regression; Bringmann et al., 2019). As psychological edges are not known quantities, we cannot identify what type of propagation characterizes a psychological network (Bringmann et al., 2019)—or, indeed, exactly what flows through it. “Activation”? Symptoms? What would a network concretely resemble if these flowed through it? These questions make it clear that although one can theorize about how psychological constructs propagate through a system, little concrete knowledge exists, and so applying specific centrality metrics is likely often inappropriate. Additional assumptions implicated in many centrality measures are problematic for psychology networks. For example, the assumption of exchangeability—that all nodes are exchangeable—does not hold for many psychological constructs, particularly symptoms: not only are symptoms qualitatively different, but some symptoms are more severe than others (e.g., suicidal ideation; Bringmann et al., 2019).

Another crucial issue involves negative edges, which are not present in social networks but often are in psychology networks. Centrality measures from social networks that are adapted to allow for negative edges often take the absolute value of edges (e.g., strength centrality), implying that any connection between variables (whether positive or negative) has an equivalent impact on the network. However, another approach involves decreasing the centrality index of a node if it has a negative edge (e.g., expected influence; Robinaugh et al., 2016), thus implying that only positive edges denote structural importance to the network, since nodes with high centrality indices are designated as most central.[2] This could make sense for a network with only symptoms if a researcher is interested in examining which symptoms most strongly activate others, but many psychology networks (including those with symptoms;

[2] For example: if two nodes have the same amount and weight of positive connections, but one also has an additional negative edge, this second node would have a lesser expected influence value.

Jones et al., 2017) involve external contextual variables, which are likely not all straightforwardly "negative" or "harmful." This is why I did not include expected influence indices for the temporal network analyses in Chapter 7 and Chapter 8, while I did for the cross-sectional network analysis in Chapter 4, which included only straightforwardly negative variables. These examples demonstrate that the way centrality indices are calculated carry assumptions about how variables and edges contribute to the network as a whole—and these assumptions are rarely made explicit or discussed.

Moreover, although in psychopathology networks central nodes are often discussed as important targets of prevention and treatment, we still know little about whether centrality indices predict onset or severity of psychopathology (Bringmann et al., 2019). One of the only studies that has investigated this, Rodebaugh et al. (2018), found that although strength centrality (i.e., the number and weight of edges connecting to any given node) did predict change processes and the severity of social anxiety, it had less predictive power than simply the number of times a patient endorsed a given symptom. Moreover, a network approach to psychopathology proposes central nodes as intervention targets with the assumption that "turning off" a central node could deactivate the whole network, thus proposing treatments that are more efficient and targeted. However, most interventions in psychology are "fat-handed" and impact many symptoms at once, since most psychology symptoms are related and interacting (Eronen, 2020; McNally, 2021).

In light of these flaws, I agree with other researchers (e.g., Bringmann et al., 2019) that instead of focusing on specific variables, psychology research should pivot to examining the network and its behaviors as a whole. This also follows the call of dynamic systems theory and family systems research generally: networks are useful conceptualization tools precisely *because* they offer a cohesive view of many interacting variables together forming a larger whole (Dallos & Draper, 2015), instead of focusing on individual variables and connections. One method to examine structural importance while taking the entire network into account (adapted from methods using time series, such as econometrics) involves impulse response functions, or IRFs (Bringmann et al., 2019; McNally, 2021). IRFs simulate how a shock (i.e., increase or decrease in value) to a specific variable would reverberate throughout the network (Blaauw et al., 2017). In other words, IRFs describe the structure of the network in a manner that takes into account all interactions between all variables. Although IRFs better suit psychological temporal networks and systems approaches than centrality measures adapted from social network analysis, they are currently implemented mainly for idiographic (and not multilevel) networks, and are not easily executable via an active R package (Blaauw et al., 2017).

Connectivity

A second important concept key to the network approach to psychopathology involves connectivity, or the view that increased connectivity between symptoms denotes more severe psychopathology (see Connectivity section on page 31). In this dissertation, I do not directly examine whether network connectivity (i.e., denser networks with more connections between nodes) relates to parental burnout severity, mainly because the networks I analyze do not only include symptoms but also family contextual factors—and so increased connectivity would not imply closer connections between symptoms, but instead that parents' experiences of burnout are closely entwined with their family environment. However, Spiller and colleagues (2021) have examined network connectivity and its relation to occupational burnout severity. They recruited 43 participants who answered experience sampling prompts on their stress, exhaustion, frustration and worry five times a day for 17 days, and then generated idiographic temporal network analyses for each participant via the random effects from multilevel VAR models. To compute network density, they averaged the absolute values of all edges (both cross-lagged and autoregressive). When correlating density values with burnout, they found a correlation of r = .32 (with CIs ranging from .09 to 1.0). This suggests that increased connections from one timepoint to the next indeed relates to occupational burnout severity for this sample.

Three important limitations constrain their results, however. First, COVID-19 related lockdowns began in the middle of data collection, with part of the recruitment process delayed for a few months; the anxiety and uncertainty around COVID-19 likely impacted the collected variables (stress, exhaustion, frustration, and worry). Second, only considering absolute values in connectivity might lose important information for these specific variables, as negative edges connecting two variables (e.g., exhaustion toward frustration, or stress toward worry, both shown in participant networks in Figure 2 from Spiller et al. 2021) would symbolize a dampening of the network, rather than its activation. This reflects a general lack of agreement on *how* to assess network density in psychological networks, even though different methods yield different results (e.g., de Vos et al., 2017) and imply different assumptions about how edges impact the network. Ideally, Spiller and colleagues (2021) would have investigated and compared multiple methods to compute density using sensitivity analyses, to assess whether the method chosen to compute density impacted their results. Lastly, as Spiller and colleagues (2021) discuss in their limitations section, multilevel VAR models do not allow for examination of how networks and their density change over time. This is a crucial limitation, as the connectivity hypothesis suggests that an individual's network will change over time: specifically, that as a mental disorder (here: occupational burnout) worsens, the individual's symptom network will become more connected.

From these results, we cannot extrapolate to parental burnout, both in light of these limitations, and because parental burnout likely operates somewhat differently from occupational burnout (particularly with the variable of "distance," as described in the section on Emotional distance on page 174). For example, future research could investigate connectivity and parental burnout severity in a network that only included straightforwardly "negative" parenting experiences (e.g., emotional exhaustion, emotional distance, and feeling fed up; as well as perhaps other relevant variables such as shame and feelings of isolation). However, all results from this dissertation point to parental experiences changing in close interaction with their family context, and so only examining connections between parenting experiences without also incorporating the family environment likely leaves out essential information.

Evaluating network models for a family systems approach

Another main goal of this dissertation focused on assessing the utility of methods stemming from a network approach to psychology (Borsboom & Cramer, 2013) for examining the family system. Examining and understanding the family system is no simple endeavor. Each individual human being is fantastically complex, from the many intertwined levels of their biological system (e.g., molecular, cellular, and hormonal), to the emotional, cognitive, and behavioral parts of the human system (Dallos & Draper, 2015). Adding in the level of the family—attempting to understand and account for how individuals interact and influence each other over time—renders the task exponentially more complicated. As Dallos and Draper (2015) put it: "Debates raged about whether families were too complex to be able to develop causal explanations of pathology" (pg. 148). Within therapy, this complexity can be circumnavigated by focusing on solutions and helpful changes, even without a deeper understanding of causes and difficulties (Dallos & Draper, 2015, p. 149). For research, however, understanding the causes and mechanisms in detail is the very goal.

Therefore, to adequately grasp the complexity of family systems, research should:

1. **Be grounded in its context** rather than stripped of it (as is commonly done for psychological research, e.g., many experimental designs). This relates to the concept of **holism**: the system is more than the sum of its parts, and so cannot be understood by examining the parts individually (this concept appears in general systems theory, von Bertalanffy, 1967; as well as specifically in family systems theory, Kerig, 2019).
2. **Examine the structure of the system.** In other words: what are the various components of the system and how do they interact (Granic & Hollenstein, 2015)?
3. **Allow for circular causality**: components should be able to have not just linear relationships but also feedback loops and other non-linear interactions, otherwise known as "multiple reciprocal interactions among system elements" (Granic & Hollenstein, 2015, p. 892). This relates to the concept of **interdependency** (Kerig, 2019), which is crucial when examining a construct as embedded within a context: one factor cannot impact just one other factor independently of other variables and the context, as interactions between variables and their contexts are always more complex.
4. **Examine how the system changes over time.** The structure and organization of a system will change over time, either as the individual components and relationships change (Masten & Palmer, 2019) or in the face of internal or external perturbations (Granic & Hollenstein, 2015). Understanding how systems change means understanding how systems evolve within their context.

Specifics of each network analysis type

Based on these essential aspects of investigating family systems, let us evaluate the statistical models used for network analyses in this dissertation (see Table 16).

Table 16. Which criteria to assess family systems do network models fulfill?

Criteria to assess family systems	GGM	DAG	mlVAR
Context	X	X	X
Structure	X	X	X
Reciprocal interactions			X
Change over time			

Note. GGM = Gaussian Graphical Model (generated with cross-sectional data); DAG = Directed Acyclic Graph (generated with cross-sectional data); mlVAR = multilevel Vector Autoregressive model (generated with intensive longitudinal data).

First, Graphical Gaussian models (GGMs) are simplistic, essentially visualizing (partial) correlations with some regularization (e.g., pruning of small edges). This can be useful to examine statistical dependencies between many cross-sectional variables. Thus, GGMs allow a simultaneous examination of parenting and family context variables and present a possible structure for how variables interrelate. However, their utility peaks for investigating psychological phenomenon in early stages of research, when the structure between variables is still poorly known (Ryan et al., 2022).

Second, Directed Acyclic Graphs (DAGs) offer greater complexity by visualizing conditional dependencies between cross-sectional variables, thus offering a glimpse at causal dependencies between variables. In other words, researchers can use DAGs to "make inferences about the structure and nature of directed causal relationships" (Ryan et al., 2022, p. 1) from cross-sectional data. This enormous utility is tempered by strict and strong assumptions (e.g., inclusion of all relevant variables; Rohrer, 2018), the need for careful interpretation mindful of pitfalls (e.g., collider bias; see Rohrer, 2018)[3], and the impossibility for feedback loops (essential for family systems theory).

Third, multilevel vector autoregressive (VAR) models allow reciprocal interactions (i.e., feedback loops) between variables, estimated from intensive longitudinal data. Multilevel VAR models assume stationarity (i.e., stability over time; see Assumptions section on page 97), and so they suit well to investigating a system in a stable state, with repeatable and recognizable patterns and cycles (Kerig, 2019). With this assumption of stationarity, however, VAR models cannot investigate how a system changes over time.

Thus, by combining these three different types of networks analyzing different data sets, a more comprehensive picture can form of how parenting experiences embed into the structure of a family system. As Rohrer (2018, pp. 39–40) put it: "Different research designs are neither mutually interchangeable nor rivals, but can contribute unique information to help answer common research questions." Each network adds slightly different information: the GGM shows statistical dependencies in cross-sectional data; the DAG visualizes (potential)

[3] As Rohrer (2018) explains regarding colliders: "A collider for a certain pair of variables is any variable that is causally influenced by both of them. Controlling for, or conditioning analysis on, such a variable (or any of its descendants) can introduce a spurious (i.e., noncausal) association between its causes."

causal dependencies in cross-sectional data; and the multilevel VAR model demonstrates patterns appearing over time, including reciprocal ones. This is also why I examined affective dynamics in Chapter 9: to provide different information from a different angle (here, whether the *way* variables fluctuated related to parental burnout severity), so as to build a more comprehensive picture of parents' experiences of parental burnout. However, this general assessment of each of the three statistical network models does not answer the broader question of whether a network approach suits the examination of family systems.

Strengths of network models to assess family systems

The psychological network models used in this thesis are existing statistical models (Bringmann et al., 2019) that clearly visualize connections between variables, thereby allowing a coherent view of the variables as a network system. This places the emphasis on the system itself instead of on connections between individual variables, which is a major component of family systems theory (e.g., holism; Kerig, 2019). This also allows a focus on context, as any variables believed relevant can be included in the network. Note that some researchers believe only symptoms should be included in a network investigating psychopathology (Borsboom, 2017), although others disagree, preferring to include any variables that are "plausible causal candidates in the etiology and maintenance of disorders" (Jones et al., 2017, p. 2). In any case, though, using network analyses to investigate family systems requires investigating context. Moreover, since both cross-sectional and temporal networks assess associations between all variables, this allows for the investigation of many simultaneous interactions at once. Temporal networks go a step further and allow the examination of the interdependency and reciprocal interactions between variables, since edges can point in either direction. All network models used in this thesis can thus be useful exploratory tools for investigating the basic structure of a family system.

In addition, temporal network analyses (and specifically, multilevel VAR models) assess patterns appearing between variables over time. Thus, they can answer dynamic questions—such as: how does a parent' experience and their interaction with their environment, and particularly their children, (mutually) influence each other over time? Investigating such patterns that establish themselves over time reflects how parenting experiences and practices are not static, but instead form through countless small moments which build up to form an overall "tone" of parenting. Parents react in different ways in different situations, and their parenting behaviors can show a lot of variability from one day to the next—but their behaviors and experiences within all of these moments add up over time to form a distribution, and the parameters (e.g., average, variation) of this distribution demonstrate their parenting traits. This concept comes from Whole Trait Theory to explain how personality traits form a relatively solid distribution from many varied moments (Fleeson & Jayawickreme, 2015), and is also key to family systems theory: "...the notion of a stable, invariant, biologically determined personality is challenged by the notion of people's actions and experiences as shaped by the processes of feedback in different social contexts and co-creation of shared meanings" (Dallos & Draper, 2015, p. 266). Thus, intensive longitudinal data is particularly useful to capture such parenting moments as they build into distributions; and analyzing this data with multilevel VAR models can assess the parenting patterns that form over time in interaction with the family environment.

Flaws of network models to assess family systems

However, multilevel VAR models, as well as cross-sectional models, cannot assess how systems change over time. At most, one could compute networks at different stages of someone's life or treatment and compare network structures (e.g., Hofmann et al., 2020). However, this would only provide static snapshots of a system (albeit snapshots potentially built on temporal data) before and after different events, but not examine *how* the system changes, which is essential to a systemic approach (Granic & Hollenstein, 2015). As Davies and colleagues put it (2004): "According to family systems theory, the family unit is an open system in which family perturbations and stressors (e.g., parent psychopathology, relationship disturbances, disruptive family events) may provoke subsequent reorganizations in family functioning and function as architects of interdependencies between family subsystems." As such, parental burnout—both a form of parental dysfunction and a disruptive family event—should alter the structure of the family system; but the network models used in this dissertation cannot directly capture such change. Indeed, although the temporal network analyses can demonstrate patterns occurring within parents' experiences, they do not on their own provide robust evidence to implicate exhaustion in the development of parental burnout, nor distance in its maintenance (see Chapter 7 and Chapter 8). This is because VAR models can only show which patterns reliably occur and re-occur over time, but not how systems change.

Another (although more general) limitation of network analyses: the many variables and interactions modeled render it difficult to form confirmatory hypotheses unless they involve the whole network. One could easily hypothesize and then assess the density of a network, for example (especially if comparing multiple networks), but to hypothesize about two specific variables diminishes the reason behind a network approach, as well as faces increasing difficulty in specific predictions as the number of nodes increases. Although hypotheses about which nodes most influence the network should be straightforward, the many complications of applying centrality measures to psychological networks (see Section on Reassessing Centrality on page 179) render this challenging as well. For example, I theorized the emotional distance would be structurally important in the network representing experiences of parents in burnout—but without a reliable centrality metric, translating this hypothesis into a concrete prediction proved difficult. My strategy was to assume distance should be strongly related to the other parental burnout variables (which was not the case), but the final network (see Figure 19) did contain emotional distance possessing the strongest autoregressive loop as well as a strong reciprocal connection to "sharing positive moments with children." Combined with other findings (e.g., inertia of emotional distance from Chapter 9), I concluded that emotional distance likely does play an important role in the maintenance of parental burnout. But is the network analysis in Figure 19 sufficient on its own to conclude that emotional distance plays an important structural role in the parents' network? Currently, without having robust objective centrality metrics, the answer is no, as researchers can subjectively interpret and explain many nodes in a network as particularly structurally important.

Moreover, multilevel VAR models come with particular flaws impacting their suitability for examining family systems. First, when modeling many variables, they can result in overfitting (i.e., the coefficients capturing not only the effects of interest but also part of the error) due to too little data for the number of variables (Bulteel et al., 2018). This can be addressed in future research by either using regularization (e.g., setting some edges to zero such as with LASSO shrinkage) or reducing variables to fewer dimensions (e.g., with principal component analysis), both of which can lead to better predictions (Bulteel et al., 2018). Second, VAR models assume a discrete flow of time—that is, that variables change from one timepoint to the next in distinct steps—and do not reflect how time continues flowing between separate timepoints (Ryan et al., 2018). This means that results from VAR models are dependent on the time interval. As an example, consider taking medication for a headache: there is little effect right away, a small effect after some minutes, a larger effect after a few hours, and no effect a day later (Ryan et al., 2018). Although this leads to a curved effect evolving over time, a VAR model would show different edges depending on time interval: no association when looking at the level of seconds, a small association for minutes, a large association for hours, and no association for days. In other words, coefficients change size, sign and relative ordering based on the time interval, and so results do not typically generalize to other time intervals (Ryan et al., 2018). Thus, discrete time models such as VAR models should ideally only be used if there is only one relevant time interval, and only that interval is measured and modeled. Although this can be argued for parenting (i.e., childrearing routines revolve around daily routines, and so likely parenting experiences follow as well), it is still a strong assumption that no other time intervals matter for assessing how parental experiences embed within the family system.

In light of these general flaws in using a network approach to assess family systems, as well as the specific limitations of VAR models, I will discuss a number of alternative models that could be used for future research.

Alternative statistical models

Time-varying models

To allow models to assess change over time, a crucial facet of dynamic systems, future research can use newly developed time-varying VAR models, which allow parameters to vary over time. Multiple model types exist. First, time-varying VAR models can allow parameters to *gradually* shift over time based on a generalized additive model (GAM) framework using non-parametric smoothing functions (Bringmann et al., 2017). Second, they can allow *sudden* shifts between two states, otherwise known as regime shifts. This could be done, for example, via models using kernel change point (KCP) detection (Cabrieto et al., 2021). Investigating sudden shifts between two stable states is an important aspect of family systems theory (Granic & Hollenstein, 2015). In addition, psychopathology researchers have previously examined the clinical usefulness of detecting an impending shift from a stable "healthy" state to a stable "disordered" state (i.e., early warning signals; van de Leemput et al., 2014; Wichers et al., 2016), although such signals are likely not generic or independent of models, but instead specific to the system in question (Dablander et al., 2022). Time-varying models also exist that combine both methods, and so can assess both sudden and gradual changes in parameters (Albers & Bringmann, 2020). Moreover, detailed tutorials explaining time-varying models and how to use them (including descriptions of R packages and example data) are available (e.g., Haslbeck et al., 2020).

Continuous time models

To address how VAR models depend on the specific time-interval chosen, researchers can use continuous time models instead of discrete time ones. These models are similar (and in some cases, mathematically equivalent to) discrete time models, but use differential equations to model time as continuously flowing instead of separated into distinct timepoints (Ryan et al., 2018). This not only solves the problem of time interval dependency, but also allows simultaneous modeling of variables that fluctuate on different timeframes (e.g., sleep and mood), which can be more complicated for discrete time models (Bringmann et al., 2022). Discrete time models can be more difficult to implement and interpret, but tutorials and R packages (at least concerning simple models) already exist (e.g., Ryan & Hamaker, 2022). Moreover, researchers have proposed centrality measures specifically for continuous time networks, to examine how an increase in a specific variable (either at a specific timepoint or over an interval of time) impacts the rest of the system (Ryan & Hamaker, 2022). Extending continuous time models to allow a multilevel framework, regime-switching, non-linear dynamics, and higher order models are all possible, although most extensions are not yet easily implemented in accessible ways (e.g., via open-source R packages).

State-space models

State space models originate from engineering and economics with historically little carry over into behavioral sciences; their main goal was to "track changes in dynamic systems and to implement diverse time series models" (Chow et al., 2010, p. 306). These models describe temporary stable states as *attractors* (working to 'attract' the system from other nearby temporary states) in the *state space* of a system, which encompasses all possible states of a system (Lewis et al., 1999). An offshoot of this approach is typical within the field of developmental psychopathology research, particularly to examine dyadic interactions between parent and child through state-space grids (Granic & Hollenstein, 2015). Using these visual representations of state-space grids, behavioral habits are operationalized as attractor states, and all possible behaviors (typically coded by researchers) form two ordinal variables which make up the axes of the grid (Lewis et al., 1999). However, the general class of state-space models are much broader (Slipetz et al., 2023) but rarely used in psychology research (with some notable exceptions, such as to examine affective dynamics; Oravecz et al., 2011).

Slipetz and colleagues (2023, p. 4) define the state-space modeling framework as "a general framework that represents the observed time-varying variables as measurements of a set of underlying, unobserved state variables." Because this framework involves underlying (i.e., latent[4]) variables, it can model measurement error; it also allows for non-linear and time-varying relations as well as continuous time models (Slipetz et al., 2023). However, as these models come from engineering where most parameters are known, it can be finicky to estimate parameters as required for psychology models (Slipetz et al., 2023). Relatedly, state-space models are currently optimal for analyzing continuous data (or at most, data collected on a 0-100 scale) from single participants, as they are not yet suitable for ordinal data or multilevel extensions (Slipetz et al., 2023). Thus, most applications currently feasible for psychology research involve linear models for one participant; such straightforward forms are already easily applicable via the *dynr* (Ou et al., 2019) or *dlm* (Petris, 2010) R packages. Nonetheless, state-space models hold important potential, particularly as further extensions become possible, since they are "particularly conducive for representing intensive within-person changes" (Chow et al., 2010, p. 326).

Are these models ready for psychological applications yet?

Applying models from other fields to psychology means that we do not have to reinvent the wheel: we can use tools already developed in other fields and apply them to psychological constructs. This not without difficulties, of course. First, not all models and tools apply easily or make sense with psychological constructs (as discussed above for propagation processes in social networks vs. psychological networks and implications for centrality; see Discussion section on Centrality). Moreover, when adapting from other fields, psychology researchers do not have all the context and expertise from that field, which can lead to improper usage (e.g., see discussion by Brown et al., 2013). Nonetheless, the ability for methodological advances to easily permeate from one field to another based on the shared concepts of a systems approach has long been viewed as one of the major benefits of a general systems theory (von Bertalanffy, 1950). As such, systemic approaches in psychology can undoubtedly benefit from models adapted from other fields: "Our standard models for multivariate timeseries (i.e., ones with linear dynamics) are not capable of capturing any complex system behaviors. Non-linear models allow for complex systems behavior, but require careful specification and construction" (Slipetz et al., 2023, p. 26).

However, in many respects these models from other fields are promising but not yet ready for widespread use within psychology research. They currently can be applied relatively easily (e.g., with accessible tutorials and R packages or other open-source analysis software) only in simplistic forms, often suitable for analyzing data from one participant in a linear fashion. However, extending these models to multilevel forms (i.e., examining multiple participants at once while allowing some parameter heterogeneity between participants) is often not yet (easily) possible (Chow et al., 2010; Ryan & Hamaker, 2022). Moreover, most of the models

---

[4] A latent approach and a network approach are not as diametrically opposed as sometimes asserted by researchers, and can be combined (Bringmann & Eronen, 2018). But in any case, a family systems approach is not theoretically opposed to modeling latent variables, particularly if this allows for modeling of measurement error.

require many more timepoints, particularly when examining time as continuous or including time-varying parameters (Haslbeck et al., 2020; Ryan & Hamaker, 2022; Slipetz et al., 2023). In sum, these models present key pathways for future research—modeling how systems change over (continuously flowing) time—but their full potential is not yet accessible to the average psychology researcher. Nonetheless, many researchers are currently actively developing these models and adapting them for accessible use in psychology research (Albers & Bringmann, 2020; Bringmann et al., 2017; Haslbeck et al., 2020; Ryan & Hamaker, 2022; Slipetz et al., 2023). With such advances, these models should be well suited to investigating family systems.

Computational models and formal theories

Further methodological advances for family systems theory and parental burnout could take the form of more precise models and theories. If a model exactly defines the system it investigates, articulating its separate components and how they interact—not in a vague verbal way, but via mathematical formulas or other formal language—then more precise findings can follow (Smaldino, 2017). This is because formal models constrain possible options, and so it is easier to rule out some of these options (Newell, 1973). In addition, formal models make assumptions explicit. Some assumptions will always be incorrect—because models attempt to simplify the world and therefore can never perfectly capture it—but formal models allow us to examine which assumptions cannot be true, as well as what would happen if all assumptions would be true (Smaldino, 2017). In other words: the purpose of a (formal) model is to be wrong, as it presents a simple explanation of a complex phenomenon, and thereby allows researchers to closely examine what the model captures accurately and what it ignominiously fails to.

Such formal models would be relevant in two ways for concepts covered in this dissertation. First, although the concept of family systems comes with a number of agreed upon aspects (e.g., reciprocal interactions, change over time; see Discussion section above on Evaluating network models for a family systems approach), it still is a relatively vague concept, as is the case for most conceptualizations of psychological systems. Further work is required to more precisely specify what is actually meant by the phrase "psychological systems are complex systems," at both the level of psychological theory and the level of quantitative models (Slipetz et al., 2023, p. 26). Second, a computational model representing parental burnout could strictly articulate a theorized pathway for how parental burnout develops. Such a model could for example include a parent in a stable "healthy" state that experiences a fast-moving process (e.g., stressors that pile up) that eventually "turns on" a slower-moving process (e.g., emotional exhaustion and then emotional distance) which leads to the parent entering a second stable state: a "burned out" state. One possibility to represent the two stable states is via a regime-switch (Dablander et al., 2022). A similar model has been described (although only verbally, not formally) for occupational burnout, based on allostatic load theory: that stressful situations eventually disrupt the stable state of physiological systems, causing an alternative equilibrium of physiological states in the form of burnout (Grave et al., 2022). If such a model were formally specified for parental burnout (as has been done for panic disorders; see for example

Robinaugh, Haslbeck, et al., 2019), the specific predictions could be further investigated via simulations as well as empirical data, to determine which aspects hold up to investigation and which do not.

Formal models also have their limitations, however, particularly for psychological constructs: they require particularly well-defined constructs as well as valid ways of measuring these constructs, which is notoriously difficult and controversial across all subfields of psychology research (Eronen & Bringmann, 2021). Moreover, even formalizing theories into mathematical models cannot escape the vagueness of psychology concepts—or so argue Hutmacher and Franz (2024). Indeed, Hutmacher and Franz (2004) assert that since psychological constructs irrevocably entwine with people's everyday lives, including "the way we talk about and the way we experience our human mental reality" (pg. 12), psychological constructs change and shift along with how people experience and discuss them, unlike technical constructs in other fields such as physics that are independent from lived experience. Thus, psychological constructs are dependent and shifting depending on their temporal, geographic, and cultural context, and so they cannot be defined as immutably and technically as would be optimal for formal mathematical models.

With this framework of formal theories, and how they are useful in their shortcomings (Smaldino, 2017), it becomes clear that none of the statistical models (used in this thesis, or proposed as future directions) can fully capture the intricacies of the family system. This is because even when allowing complexity, quantitative research must always simplify to some extent. The hope is that even though these models are doubtless wrong in some crucial ways, they yield sufficient information to further advance research, including about what the model can accurately capture and what it cannot.

10.5 Broader methodological considerations

To further discuss what this dissertation managed to capture and what it failed to, I now discuss some broader methodological considerations.

Group versus individual

In this dissertation, I focus especially on group-level results. This was a deliberate choice to look for common patterns amongst different families, in an attempt to improve general understanding of how the family context interacts with parental burnout. However, this means that this dissertation does not examine each parent's individual patterns (with a partial exception for the study of parental affective dynamics in Chapter 9), even though we know that substantial variation exist between parents, including in the samples recruited for this dissertation (as can be seen when looking at the random effects in Figure 20). This tension between commonality (i.e., what patterns and experiences different families share) and uniqueness (i.e., each family's idiosyncrasies) can often be found in family systems therapy (Dallos & Draper, 2015, p. 203). This tension also reverberates throughout psychology research generally, and particularly research on systems: can we assume that in specific circumstances (e.g., parental burnout), there are patterns and structures in a family system that reliably occur across families (or at least families in a specific culture)? When again looking at the random effects for parents in burnout in Figure 20, some similar edges can be found (e.g., the three autoregressive coefficients, as well as the edge connecting positive moments with kids toward distance). On the one hand, this supports that common patterns occur across families experiencing parental burnout—but on the other hand, these associations were estimated via a multilevel model which shrinks estimates toward one another to some degree.

Idiographic analyses (i.e., investigating just one person's symptoms or emotions over time) are possible, with potential clinical utility in tracking processes of change during times of transition[5] (e.g., treatment; Hofmann et al., 2020). Researchers propose that such idiographic networks focusing on symptoms or emotions could be used for individualized interventions (e.g., Bak et al., 2016). In practice, however, no consensus exists on how to interpret network structure and node centrality, as demonstrated by the study which sent the same data from one patient to multiple research groups, none of whom identified a common treatment target or process (Bastiaansen et al., 2020). Thus, using idiographic networks in clinical practice to identify specific treatment cannot (yet) be implemented. In any case, however, this dissertation did not aim for such family-specific idiographic insights, instead seeking to examine common patterns across families to deepen understanding of parental burnout—at least as it occurs within (Belgian) family systems.

Within family systems theory, this tension between group patterns versus individual ones reflects a tension between expert truths and uncertainty. Here, expert truths can be understood as viewing systemic structures and patterns as rules that are "truth" and can be identified, such as was proposed by early systemic thinkers (Dallos & Draper, 2015, p. 203). Uncertainty, meanwhile, involves viewing family systems as built by the families themselves, when they form meanings and identities via interaction. This latter concept of families brings with it not only astounding variation between families (constrained mainly by cultural narratives about families and gender and parenting), but also uncertainty about how much commonality exists across families (Dallos & Draper, 2015). The studies in this dissertation mainly veer on the side of "commonality," searching for patterns that can be found again and again across (Belgian) families. However, I want to—at least in this discussion—allow a "space where uncertainty about the truth [can] be tolerated" (Dallos & Draper, 2015, p. 182). Indeed, research, particularly on human beings, cannot contain strict certainty. Moreover, uncertainty is critical for systemic theory, as open systems (and especially human systems) are relentlessly complex and can always expand outwards as they interact with other systems. We cannot thus designate one pattern that always occurs even in one family, let alone across all families; but we can seek patterns that are likely to occur and re-occur across certain types of families that are experiencing specific problems, even while acknowledging and accepting that there will always

[5] Although note that most idiographic network analyses use VAR models and assume stationarity, and therefore are only able to examine network structures before and after transitions, as done by Hofmann and colleagues (2020).

be idiosyncrasies and individual variations. In other words, the uncertainty in this dissertation reflects an inability to dive into each individual's context and family structure in sufficient depth to fully understand each specific family. This therefore reflects not only the tension between group and individual, but also the constraints of quantitative methods, which do not allow the nuanced insight into meaning and identity that qualitative methods can more easily pursue.

Reflections on measurement

Hollenstein (2021) presents a useful metaphor to describe experience sampling methodology: instead of watching a full movie, we see a series of screenshots. In other words, these screenshots show glimpses of the movie over time, but do not shed much light on the continuity of the movie or its broader context. The quest for complexity at the heart of this dissertation is thus limited by its data: brief snapshots that do not contain the full meaning or trajectory of parents' experiences. This reduction of the rich lived experience of a person to small snapshots is part of all quantitative research: "Quantitative methods [are] reductionist, in that attempts are made to reduce the phenomena to small, focused, and manageable components" (Dallos & Draper, 2015, p. 238). Moreover, the precise method we used to assess parents' affect also constrained their responses: we asked parents to reduce their parenting experience to a number, thus masking meaning and nuance. For example, when a parent feels emotionally distant from their child, that might relate to the situation at that moment (e.g., a temper tantrum), the rest of their day (e.g., they just got laid off work), their general mood (e.g., overwhelmed and angry), their amount of sleep the night before, and so on. In addition, feeling emotionally distant likely involves other emotions (perhaps shame, as discussed by Hubert & Aujoulat, 2018), but this was not captured by one sole numerical rating. Although we attempted to capture some of this context with additional questions (e.g., interactions with children and co-parent; see Table 5 in Chapter 5), the parent was not able to communicate much of the nuance of their affective experience or environment. However, there was at least high between-participant reliability for the three parental burnout items, as well as substantial within-participant reliability (see the Results section from Chapter 5). This indicates that different parents showed distinct patterns of responses, but also that each parent showed relative consistency in their own responses, as expected for experience sampling data.

The quantitative data collection for much of this thesis involved parents translating their experiences into a number on a scale from 0 to 100. This comes with a number of assumptions. For example: can we actually translate our subjective experience into a number on a large scale, in a way that easily distinguishes between the different assessed constructs (Gray & Watson, 2007)? Are we able to accurately judge our own internal states, let alone quantify them? One answer focuses on how affect is a subjective experience, and so the only way to accurately assess it is by asking participants to report it. "One could argue with some justification that measures of current affect are valid as long as participants are not deliberately distorting their assessment. Put differently, the respondents' ongoing synthesis of physiological, situational, and cognitive information *is* their current mood" (Gray & Watson, 2007, p. 178). For the purposes of this dissertation, therefore, we assume that participants can interpret their internal state and represent it with a point on a scale in a relatively consistent manner.

Considering the process of translating affect into a number, Teoh and colleagues (2023) propose that we can think of this as a decision process: interpreting signals from our body and context, and then choosing from constrained options (i.e., point on a scale) to represent it. Future research could directly use such a framework to investigate affect ratings via decision-making frameworks, to more precisely investigate *how* participants decide on a specific number. Such a decision is not easy (as reported by participants in this dissertation; see Appendix C). One method to ease such a decision for participants is to frame it as a comparison with their previous answer, such as via a visual marker placed on the scale indicating their last response (Dejonckheere & Mestdagh, 2023). In this case, their decision does not involve each time repeatedly translating their experience into a number, but instead to decide whether their current experience is more or less intense than their past report. Participants reported that such a reference point helped them more easily and accurately rate their emotions in experience sampling prompts (Dejonckheere & Mestdagh, 2023).

That does not mean, however, that participants' responses do not contain measurement error or other important flaws—far from it. For example, in-the-moment ratings of affect are most accurate, as they do not involve retrospective biases such as duration neglect (i.e., focusing on most intense experiences but forgetting the length of these) or recency effects (i.e., remembering best what just occurred; Gray & Watson, 2007). However, in-the-moment ratings (often assessed multiple times a day) are problematic when assessing parenting affect, as most parents do not spend all moments with their children, and parenting-related interactions often follow a daily routine organized around work/school (see Daily Interval section in Appendix C). We thus made a considered trade-off, prioritizing an intuitive sampling schedule following the daily rhythms of parenting instead of a more frequent schedule less suitable for capturing parenting affect and parent-child dynamics. Moreover, the signal-to-noise ratio (i.e., if a signal captured in data is substantial enough to outweigh measurement noise) of emotional time-series data is often particularly poor (Dejonckheere & Mestdagh, 2021). One way to improve this is to assess affective reactions to specific stimuli rather than global ones (Dejonckheere & Mestdagh, 2021; Lapate & Heller, 2020), which was done in this thesis: rather than assessing global affect, data collections focused on stimuli anchored in the context of parenting.

Culture and parenting

Although the goals of this thesis did not involve examining culture, "virtually all aspects of parenting children are informed by culture" (Bornstein & Cheah, 2006, p. 7); findings should thus be grounded in their cultural context, and the role of culture discussed. To grasp just how much of parenting is conditioned by culture, consider this statement by Bornstein and Cheah (2006, p. 7): "Culture influences when and how parents care for children, the extent to which parents permit children freedom to explore, how nurturant or restrictive parents are, which behaviors parents emphasize, and so forth…" In other words, culture influences not only how parents interact with their children, but also what they consider appropriate or maladaptive behavior and activities for their children, and appropriate or maladaptive behavior for themselves when encouraging or disciplining their children. As such, parenting cannot be understood apart from its cultural context; when research attempts to do so, it accepts variables as natural and immutable when they actually range considerably across cultures (Bornstein & Cheah, 2006). This is still too often the case, as most research on parenting and child development still stems from North America and Europe, and so what is all-too-often considered universally "typical" parenting and child-rearing practices are actually culturally specific to these regions (Bornstein & Cheah, 2006). Since all parents recruited for this dissertation were living in Belgium, it is particularly important to ground findings in their Western European cultural context, involving intensifying parenting norms centering parental closeness and warmth (Dupont et al., 2022; Hays, 1996). As such, findings likely do not generalize to cultures with different parenting and child-rearing norms.

Parental burnout also inextricably intertwines with culture: It involves a preponderance of parenting stressors with too few parenting resources to compensate (Mikolajczak & Roskam, 2018), and culture impacts what parents consider to be stressful or revitalizing. In addition, parents often experience cultural expectations and norms about parenting—and in particular, the resulting feeling that they are not able to meet those expectations or adhere to those norms—as stressful and shameful. Even the suffering related to parental burnout, which many parents link with shame and isolation from feeling distant toward their children (Hubert & Aujoulat, 2018), likely varies depending on cultural norms around parental warmth and closeness, as discussed above regarding varying interpretations on "distanced mothering" (Lubiewska et al., 2023).

The potential for family stress exists "in most if not all culturally structured family environments" (Super & Harkness, 2020, p. 191), so the question then becomes: when might this family stress turn into parental burnout? And since what constitutes stress relates to the environment and cultural norms and interpretations of behavior, what might contribute to parental burnout in different cultures? From the prior discussion based on results in this thesis, limited help available to parents beginning to feel exhausted likely contributes to the development of parental burnout, and cultural norms around distanced parenting likely contribute to the suffering experienced in severe parental burnout. However, these aspects are specific to the culture of the parents recruited for this thesis. Might other aspects be involved in parental burnout in other cultures with other parental norms? Could parental burnout still appear in cultures without such intensive parenting norms that center warmth and closeness? If yes, would it still be recognizable as parental burnout—as parental suffering in such contexts would likely not revolve around the discrepancy between cultural norms of parental warmth and lived experience of emotional distance, as it seems to for Belgian parents (see Discussion section on Emotional distance)? Such questions on the role of culture in parental burnout require further research.

Furthermore, much research on parental burnout has focused on middle-class or highly-educated parents, particularly amongst populations not typically represented in research (e.g., South American, African; see Super & Harkness, 2020). Moreover, nuances on specific cultural milieus were not collected, and so understanding particularities amongst results (e.g., why parental burnout related to high income for mothers in some countries but not others; Super & Harkness, 2020) was not possible without more detailed contextual information. Moreover, many countries have experienced an "emergence of distinctive lifestyles accompanying urbanization, as well as the increase in women's education and employment options" (Super & Harkness, 2020, p. 189). This signals that in some cultural milieus (e.g., likely middle-class) across countries, shifting norms likely bring additional pressures on parents, who potentially have to balance between multiple and conflicting parenting norms. This highlights that variability in parenting norms occurs not only *between* cultures, but also *within* cultures, and research should investigate parental burnout not only cross-culturally, but also within specific cultural milieus within one culture.

As Super and Harkness (2020) assert, to understand how parental burnout relates to culture requires details of a parent's context and their interactions with family members. This could be examined through qualitative research, but also through methodological tools such as experience sampling and network analyses. In particular, future research could use similar analyses as used in this thesis (i.e., investigating daily experiences of parenting and parental burnout grounded in the family system) but in cultures with marked variations in parenting norms and expectations, to examine how interactions between family members would vary. Even more informative would be to complement such quantitative research with qualitative methods on how parents and family members interpret their family dynamics and parental stress and exhaustion, to provide qualitative nuance and meaning to quantitative complexity.

10.6 Limitations and future directions

In addition to limitations and future directions already discussed, a number of additional limitations merit particular attention.

First, the samples in this dissertation mostly comprise of mothers, with only few fathers participating. This gave us valuable information on the parenting experiences of mothers, but possibly does not generalize to how fathers experience parenting. This is because cultural parenting norms differ for fathers compared to mothers, relating to gender norms within each culture. This leads to differences in how fathers and mothers parent and engage with their children: "Mothers and fathers interact with and care for children in complementary ways; that is, they tend to divide the labors of caregiving and engage children emphasizing different types of interactions…" (Bornstein & Cheah, 2006, p. 16). Moreover, in most cultures mothers spend on average more time with their children than fathers do, including in "Western industrialized nations [that] have witnessed relative increases in the amount of time fathers spend with their children" (Bornstein & Cheah, 2006, p. 16). Thus, when compared to mothers, fathers tend to spend less time with their children, interact with them in distinct ways, and engage with different parenting expectations and norms. It is thus possible—even likely—that fathers experience parenting, parental burnout, and their family system, in a different manner than mothers. Although some results in this thesis show no differentiations between fathers and mothers, too few fathers participated in intensive longitudinal data collections for robust conclusions regarding their experiences of parental burnout and the family system.

Secondly, this dissertation does not account for developmental shift. Across studies, parents were all bundled together regardless of the ages of their children. Although this approach of assuming common family processes across developmental stages is typical in research relating to family systems theory, family dynamics and interactions vary intensively within a family depending on the age of the children (Dallos & Draper, 2015). Moreover, it ignores how family systems can and should change over time: processes that are197daptivee at one stage (e.g., close monitoring of infants) can be maladaptive at later stages (e.g., late adolescence). However, this dissertation does not directly focus on the type of interactions between parents and children, but instead on the affective experiences of the parent and how they interpret their interactions with their children (e.g., as "difficult"). This focus on parental experiences across developmental stages of children was purposeful when creating the experience sampling items (see Chapter 5): we aimed to develop items applicable to parents of both infants and teenagers. This was because we sought patterns of parenting experience common across different types of families and interactions. Nonetheless, this deliberate grouping together of families means that this dissertation cannot offer nuanced findings about how the developmental stage of the child might impact parenting affect.

Thirdly, this dissertation focused on parental experiences only; we thus did not assess how the other family members interpreted interactions. This could be done in future research via observational coding of interactions (as commonly done in developmental psychopathology research; Granic & Hollenstein, 2015) or by having multiple family members answer daily questions about their interactions with one another. This last would also allow insight into how much parents' interpretations depend on their context. For example, when a parent in a state of burnout rated how "difficult to manage" they found their child that day, would this be similar the experience of their co-parent or their child?

Lastly, the methods we used imposed normative definitions of what constitutes a family. In our statistical models, we assumed one or maximum two parenting figures (they did not have to be married or partners, but did have to be living together, based on recruitment criteria). This ignores the multiplicity of family configurations: in many families, grandparents take on a major share of childrearing work, as might close family friends in some situations. Where parenting support could come from in families was also defined in constraining ways: multiple questions assessed support (or lack thereof) from the coparent, while all other support—from friends, family, neighbors—was lumped under "social support." This categorizing of support implies that the main source of help for a parent is their coparent, while all help from other people can be seen grouped as "other." This exemplifies how normative assumptions on family creep into and constrain research (Dallos & Draper, 2015). Thus, even when using tools to allow as much complexity as possible, research always involve narrowing, defining, and simplifying of concepts.

10.7 Conclusion

From this discussion, we can derive some closing thoughts and general conclusions:

First, future research should examine parental burnout via its separate components, as each component has distinct associations and interactions with the family context. Moreover, from the findings in this thesis supplemented by additional research, **I propose that emotional exhaustion kickstarts the development of parental burnout, while emotional distance works to maintain it**. This should be the case particularly when cultural norms on parenting prioritize emotional warmth, as parents who feel distant from their children likely experience intense shame and suffering from the discrepancy between cultural expectations of emotional warmth but a lived experience of emotional distance.

Second, parental burnout should be investigated within its family context. We embraced a family systems approach to parenting, which asserts that to understand parenting means to situate it within the family. Even for researchers not explicitly using a family systems approach, however, all of our findings highlight that parental burnout interacts with different aspects of the family environment in different ways during the course of burnout. As such, all research on parental burnout should explicitly attempt to assess the parent's family context and include it in conceptualizations and analyses. Although we endeavored to examine parents' experiences and their contexts while avoiding reductionism, this was still a quantitative thesis that comes with the associated drawbacks: we did not assess the meaning behind participants' numerical answers. Further insight and nuances into how experiences of parental burnout interact with the family context—and how these depend on cultural norms of parenting and gender—requires additional qualitative as well as quantitative research, to form a more well-rounded picture.

Third, although network analyses seem well suited to examining family systems, they are not quite ready yet. They do allow an intuitive conceptualization of the family as a system, and easily model interactions between components and (at least some aspects of the) context. However, none of the network models currently used can show how a system changes over time, and many statistical and conceptual limitations hinder commonly used models. The promise of a general systems theory—that models can easily be exchanged between disciplines—has not yet fully borne out, as adapting statistical models from other fields to psychology requires intensive conceptual and statistical work. However, many researchers are doing this work, and accessible network analyses that can examine how systems change over time appear just over the horizon.

As this thesis made clear when discussing the substantial impact of culture on parenting and parental burnout, however, context includes many expanding layers, and to fully understand a system requires understanding how it interacts with all these layers. Although the ideal approach to investigate family systems involves a multi-tiered analysis that can take into account all different levels from biology to culture, this might never be possible—particularly as we still face key problems when modeling even just one level. However, reaching for complexity—while clearly acknowledging where we fall short—is a worthwhile goal, and I hope future research on parenting, families, and, more globally, psychology, pursues it.

References

Aalbers, G., McNally, R. J., Heeren, A., de Wit, S., & Fried, E. I. (2019). Social media and depression symptoms: A network perspective. *Journal of Experimental Psychology: General, 148*(8), 1454–1462. https://doi.org/10.1037/xge0000528

Abdullah, N., Kamaruddin, M. A., Goh, Y.-X., Othman, R., Dauni, A., Jalal, N. A., Yusuf, N. A. M., Kamat, S. A., Basri, N. H., & Jamal, R. (2021). Participants attrition in a longitudinal study: The Malaysian Cohort Study experience. *International Journal of Environmental Research and Public Health, 18*(14), Article 14. https://doi.org/10.3390/ijerph18147216

Abegaz, F., & Wit, E. (2013). Sparse time series chain graphical models for reconstructing genetic networks. *Biostatistics, 14*(3), 586–599. https://doi.org/10.1093/biostatistics/kxt005

Aguiar, J., Matias, M., Braz, A. C., César, F., Coimbra, S., Gaspar, M. F., & Fontaine, A. M. (2021). Parental burnout and the COVID-19 pandemic: How Portuguese parents experienced lockdown measures. *Family Relations, 70*(4), 927–938. https://doi.org/10.1111/fare.12558

Albers, C. J., & Bringmann, L. F. (2020). Inspecting gradual and abrupt changes in emotion dynamics with the time-varying change point autoregressive model. *European Journal of Psychological Assessment, 36*(3), 492–499. https://doi.org/10.1027/1015-5759/a000589

Aldao, A. (2013). The future of emotion regulation research: Capturing context. *Perspectives on Psychological Science, 8*(2), 155–172. https://doi.org/10.1177/1745691612459518

Arikan, G., Üstündağ-Budak, A. M., Akgün, E., Mikolajczak, M., & Roskam, I. (2020). Validation of the Turkish version of the Parental Burnout Assessment (PBA). *New Directions for Child and Adolescent Development, 2020*(174), 15–32. https://doi.org/10.1002/cad.20375

Arslan, R. C., Tata, C. S., & Walther, M. P. (2018). *formr: A study framework allowing for automated feedback generation and complex longitudinal experience sampling studies using R.* (Version v0.18.3) [Computer software]. https://doi.org/10.5281/zenodo.3229668

Arslan, R. C., Walther, M. P., & Tata, C. S. (2020). formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods, 52*(1), 376–387. https://doi.org/10.3758/s13428-019-01236-y

Aunola, K., Sorkkila, M., & Tolvanen, A. (2020). Validity of the Finnish version of the Parental Burnout Assessment (PBA). *Scandinavian Journal of Psychology*. https://doi.org/10.1111/sjop.12654

Average income in Belgium reached 19,105 euros in 2019 | Statbel. (2021, October 26). Statbel. https://statbel.fgov.be/en/news/average-income-belgium-reached-19105-euros-2019

Average income in Belgium reached 19,671 euros in 2020 | Statbel. (2022, October 27). Statbel. https://statbel.fgov.be/en/news/average-income-belgium-reached-19671-euros-2020

Bak, M., Drukker, M., Hasmi, L., & Van Jim, O. S. (2016). An n=1 clinical network analysis of symptoms and treatment in psychosis. *PLoS ONE, 11*(9), 1–15. https://doi.org/10.1371/journal.pone.0162811

Balsamo, M., Romanelli, R., Innamorati, M., Ciccarese, G., Carlucci, L., & Saggino, A. (2013). The state-trait anxiety inventory: Shadows and lights on its construct validity. *Journal of Psychopathology and Behavioral Assessment*, *35*(4), 475–486. https://doi.org/10.1007/s10862-013-9354-5

Barabási, A. L. (2012). The network takeover. *Nature Physics, 8*(1), 14–16. https://doi.org/10.1038/nphys2188

Barabási, A. L. (2016). *Network science*. Cambridge University Press. http://networksciencebook.com/

Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*, 509–512.

Barabási, A. L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: A network-based approach to human disease. *Nature Reviews Genetics, 12*(1), 56–68. https://doi.org/10.1038/nrg2918

Bastiaansen, J. A., Kunkels, Y. K., Blaauw, F. J., Boker, S. M., Ceulemans, E., Chen, M., Chow, S.-M., de Jonge, P., Emerencia, A. C., Epskamp, S., Fisher, A. J., Hamaker, E. L., Kuppens, P., Lutz, W., Meyer, M. J., Moulder, R., Oravecz, Z., Riese, H., Rubel, J., … Bringmann, L. F. (2020). Time to get personal? The impact of researchers' choices on the selection of treatment targets using the experience sampling methodology. *Journal of Psychosomatic Research, 137,* 110211. https://doi.org/10.1016/j.jpsychores.2020.110211

Bavelas, J. B., & Segal, L. (1982). Family systems theory: Background and implications. *Journal of Communication, 32*(3), 99–107.

Bayani, A., Hadaeghi, F., Jafari, S., & Murray, G. (2017). Critical slowing down as an early warning of transitions in episodes of bipolar disorder: A simulation study based on a computational model of circadian activity rhythms. *Chronobiology International*, *34*(2), 235–245. https://doi.org/10.1080/07420528.2016.1272608

Bean, C. A. L., Heggeness, L. F., & Ciesla, J. A. (2021). Ruminative inertia, emotion regulation, and depression: A daily-diary study. *Behavior Therapy*, *52*(6), 1477–1488. https://doi.org/10.1016/j.beth.2021.04.004

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. Psychological Corporation.

Beck, A. T., Steer, R. A., & Brown, G. K. (1998). *Inventaire de dépression de Beck, 2e édition (BDI-II)*. Éditions du Centre de psychologie appliquée (ECPA).

Bekhuis, E., Schoevers, R. A., van Borkulo, C. D., Rosmalen, J. G. M., & Boschloo, L. (2016). The network structure of major depressive disorder, generalized anxiety disorder and somatic symptomatology. *Psychological Medicine, 46*(14), 2989–2998. https://doi.org/10.1017/S0033291716001550

Bellis, M. A., Hardcastle, K., Ford, K., Hughes, K., Ashton, K., Quigg, Z., & Butler, N. (2017). Does continuous trusted adult support in childhood impart life-course resilience against adverse childhood experiences: A retrospective study on adult health-harming behaviours and mental well-being. *BMC Psychiatry, 17*(1), 110. https://doi.org/10.1186/s12888-017-1260-z

Beltz, A. M., & Gates, K. M. (2017). Network mapping with GIMME. *Multivariate Behavioral Research, 52*(6), 789–804. https://doi.org/10.1080/00273171.2017.1373014

Berger, R. P., Fromkin, J. B., Stutz, H., Makoroff, K., Scribano, P. V., Feldman, K., Tu, L. C., & Fabio, A. (2011). Abusive head trauma during a time of increased unemployment: A multicenter analysis. *Pediatrics, 128*(4), 637–643. https://doi.org/10.1542/peds.2010-2185

Bernstein, E. E., Curtiss, J. E., Wu, G. W. Y., Barreira, P. J., & McNally, R. J. (2019). Exercise and emotion dynamics: An experience sampling study. *Emotion, 19*(4), 637–644. https://doi.org/10.1037/emo0000462

Bernstein, E. E., Heeren, A., & McNally, R. J. (2017). Unpacking rumination and executive control: A network perspective. *Clinical Psychological Science, 5*(5), 816–826. https://doi.org/10.1177/2167702617702717

Bernstein, E. E., Heeren, A., & McNally, R. J. (2019). Reexamining trait rumination as a system of repetitive negative thoughts: A network analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 63,* 21–27. https://doi.org/10.1016/j.jbtep.2018.12.005

Bianchi, R., Truchot, D., Laurent, E., Brisson, R., & Schonfeld, I. S. (2014). Is burnout solely job-related? A critical comment. *Scandinavian Journal of Psychology, 55*(4), 357–361. https://doi.org/10.1111/sjop.12119

Birkeland, M. S., Greene, T., & Spiller, T. R. (2020). The network approach to posttraumatic stress disorder: A systematic review. *European Journal of Psychotraumatology, 11*(1), 1700614. https://doi.org/10.1080/20008198.2019.1700614

Blaauw, F. J., van der Krieke, L., Emerencia, A. C., Aiello, M., & de Jonge, P. (2017). Personalized advice for enhancing well-being using automated impulse response analysis—AIRA. *ArXiv:1706.09268 [Cs]*. http://arxiv.org/abs/1706.09268

Blackwell, S. E. (2019). Mental imagery: From basic research to clinical practice. *Journal of Psychotherapy Integration, 29*(3), 235–247. https://doi.org/10.1037/int0000108

Blanchard, M. A., Contreras, A., Kalkan, R. B., & Heeren, A. (2023). Auditing the research practices and statistical analyses of the group-level temporal network approach to psychological constructs: A systematic scoping review. *Behavior Research Methods, 55*(2). https://doi.org/10.31234/osf.io/dh3cu

Blanchard, M. A., & Heeren, A. (2020). Why we should move from reductionism and embrace a network approach to parental burnout. *New Directions for Child and Adolescent Development, 174,* 159–168. https://doi.org/10.1002/cad.20377

Blanchard, M. A., & Heeren, A. (2022). Ongoing and future challenges of the network approach to psychopathology: From theoretical conjectures to clinical translations. In G. J. G. Asmundson (Ed.), *Comprehensive Clinical Psychology* (Second, Vol. 11, pp. 32–46). Elsevier. https://doi.org/10.1016/B978-0-12-818697-8.00044-3

Blanchard, M. A., Hoebeke, Y., & Heeren, A. (2023a). *Parental burnout and decreased sensitivity to context for mothers: A temporal network approach grounded in the family environment*. PsyArXiv. https://doi.org/10.31234/osf.io/hs6tb

Blanchard, M. A., Hoebeke, Y., & Heeren, A. (2023b). Parental burnout features and the family context: A temporal network approach in mothers. *Journal of Family Psychology, 37*(3), 398–407. https://doi.org/10.1037/fam0001070

Blanchard, M. A., Revol, J., Hoebeke, Y., Roskam, I., Mikolajczak, M., & Heeren, A. (2023). On the temporal nature of parental burnout: Development of an experience sampling methodology (ESM) tool to assess parental burnout and its related ever-changing family context. *Collabra: Psychology, 9*(1), 74614. https://doi.org/10.1525/collabra.74614

Blanchard, M. A., Roskam, I., Mikolajczak, M., & Heeren, A. (2021). A network approach to parental burnout. *Child Abuse & Neglect, 111*, 104826. https://doi.org/10.1016/j.chiabu.2020.104826

Bodner, N., Kuppens, P., Allen, N. B., Sheeber, L. B., & Ceulemans, E. (2018). Affective family interactions and their associations with adolescent depression: A dynamic network approach. *Development and Psychopathology*, *30*(4), 1459–1473. https://doi.org/10.1017/S0954579417001699

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks, 27*(1), 55–71. https://doi.org/10.1016/j.socnet.2004.11.008

Bornstein, M. H. (2012). Cultural approaches to parenting. *Parenting, 12*(2–3), 212–221. https://doi.org/10.1080/15295192.2012.683359

Bornstein, M. H., & Cheah, C. S. L. (2006). The place of "culture and parenting" in the ecological contextual perspective on developmental science. In K. H. Rubin & O. B. Chung (Eds.), *Parental beliefs, parenting, and child development in cross-cultural perspective*. Psychology Press.

Borsboom, D. (2008). Latent variable theory. *Measurement: Interdisciplinary Research & Perspective, 6*(1–2), 25–53. https://doi.org/10.1080/15366360802035497

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry, 16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., & Cramer, O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology, 9*, 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A.-M., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers, 1*(1), 1–18. https://doi.org/10.1038/s43586-021-00055-w

Bradley, R., Whiteside-Mansell, L., & Brisby, J. (1997). Parents' socioemotional investment in children. *Journal of Marriage and the Family, 59*, 77–90. https://doi.org/10.13140/RG.2.1.3911.2803

Brianda, M. E., Mikolajczak, M., Bader, M., Bon, S., Déprez, A., Favez, N., Holstein, L., Le Vigouroux, S., Lebert-Charron, A., Sánchez-Rodríguez, R., Séjourné, N., Wendland, J., & Roskam, I. (2023). Optimizing the assessment of parental burnout: A multi-informant and multimethod approach to determine cutoffs for the parental burnout inventory and the parental burnout assessment. *Assessment*, 10731911221141872. https://doi.org/10.1177/10731911221141873

Brianda, M. E., Roskam, I., Gross, J. J., Franssen, A., Kapala, F., Gérard, F., & Mikolajczak, M. (2020). Treating parental burnout: Impact of two treatment modalities on burnout symptoms, emotions, hair cortisol, and parental neglect and violence. *Psychotherapy and Psychosomatics*, 1–3. https://doi.org/10.1159/000506354

Brianda, M. E., Roskam, I., & Mikolajczak, M. (2019). *Comment traiter le burn out parental ? Manuel d'intervention clinique*. De Boeck Superieur.

Brianda, M. E., Roskam, I., & Mikolajczak, M. (2020). Hair cortisol concentration as a biomarker of parental burnout. *Psychoneuroendocrinology, 117,* 104681. https://doi.org/10.1016/j.psyneuen.2020.104681

Bringmann, L. F., Albers, C., Bockting, C., Borsboom, D., Ceulemans, E., Cramer, A., Epskamp, S., Eronen, M. I., Hamaker, E., Kuppens, P., Lutz, W., McNally, R. J., Molenaar, P., Tio, P., Voelkle, M. C., & Wichers, M. (2022). Psychopathological networks: Theory, methods and practice. *Behaviour Research and Therapy, 149,* 104011. https://doi.org/10.1016/j.brat.2021.104011

Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., Wigman, J. T. W., & Snippe, E. (2019). What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology, 128*(8), 892–903. https://doi.org/10.1037/abn0000446

Bringmann, L. F., & Eronen, M. I. (2018). Don't blame the model: Reconsidering the network approach to psychopathology. *Psychological Review, 125*(4), 606–615. http://dx.doi.org/10.1037/rev0000108

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological Methods, 22*(3), 409–425. http://dx.doi.org/10.1037/met0000085

Bringmann, L. F., Lemmens, L. H. J. M., Huibers, M. J. H., Borsboom, D., & Tuerlinckx, F. (2015). Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychological Medicine, 45*(4), 747–757. https://doi.org/10.1017/S0033291714001809

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., Tuerlinckx, F., & Kuppens, P. (2016). Assessing temporal emotion dynamics using networks. *Assessment, 23*(4), 425–435. https://doi.org/10.1177/1073191116645909

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE, 8*(4). https://doi.org/10.1371/journal.pone.0060188

Bronfenbrenner, U. (1979). *The ecology of human development*. Harvard University Press.

Brown, N. J. L., Sokal, A. D., & Friedman, H. L. (2013). The complex dynamics of wishful thinking: The critical positivity ratio. *American Psychologist, 68*(9), 801–813. https://doi.org/10.1037/a0032850

Bryant, R. A., Creamer, M., O'Donnell, M., Forbes, D., McFarlane, A. C., Silove, D., & Hadzi-Pavlovic, D. (2017). Acute and chronic posttraumatic stress symptoms in the emergence of posttraumatic stress disorder: A network analysis. *JAMA Psychiatry, 74*(2), 135–142. https://doi.org/10.1001/jamapsychiatry.2016.3470

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016). Using raw VAR regression coefficients to build networks can be misleading. *Multivariate Behavioral Research, 51*(2–3), 330–344. https://doi.org/10.1080/00273171.2016.1150151

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2018). Improved insight into and prediction of network dynamics by combining VAR and dimension reduction. *Multivariate Behavioral Research, 53*(6), 853–875. https://doi.org/10.1080/00273171.2018.1516540

Cabrieto, J., Meers, K., Schat, E., Adolf, J., Kuppens, P., Tuerlinckx, F., & Ceulemans, E. (2021). KcpRS: An R package for performing kernel change point detection on the running statistics of multivariate time series. *Behavior Research Methods*. https://doi.org/10.3758/s13428-021-01603-8

Caspi, A., & Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *American Journal of Psychiatry, 175*(9), 831–844. https://doi.org/10.1176/appi.ajp.2018.17121383

Cham, H., Reshetnyak, E., Rosenfeld, B., & Breitbart, W. (2017). Full information maximum likelihood estimation for latent variable interactions with incomplete indicators. *Multivariate Behavioral Research, 52*(1), 12–30. https://doi.org/10.1080/00273171.2016.1245600

Cheah, C. S. L., Li, J., Zhou, N., Yamamoto, Y., & Leung, C. Y. Y. (2015). Understanding Chinese immigrant and European American mothers' expressions of warmth. *Developmental Psychology, 51*(12), 1802–1811. https://doi.org/10.1037/a0039855

Cheng, H., Wang, W., Wang, S., Li, Y., Liu, X., & Li, Y. (2020). Validation of a Chinese version of the Parental Burnout Assessment. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.00321

Chow, S.-M., Ho, M. R., Hamaker, E. L., & Dolan, C. V. (2010). Equivalence and differences between structural equation modeling and state-space modeling techniques. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(2), 303–332. https://doi.org/10.1080/10705511003661553

Coffey, J. K., Xia, M., & Fosco, G. M. (2020). When do adolescents feel loved? A daily within-person study of parent–adolescent relations. *Emotion,* 861. https://doi.org/10.1037/emo0000767

Coiro, M. J., & Emery, R. E. (1998). Do marriage problems affect fathering more than mothering? A quantitative and qualitative review. *Clinical Child and Family Psychology Review, 1*, 23–40.

Contreras, A., Nieto, I., Valiente, C., Espinosa, R., & Vazquez, C. (2019). The study of psychopathology from the network analysis perspective: A systematic review. *Psychotherapy and Psychosomatics, 88*(2), 71–83. https://doi.org/10.1159/000497425

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. J. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality, 54,* 13–29. https://doi.org/10.1016/j.jrp.2014.07.003

Costenbader, E., & Valente, T. W. (2003). The stability of centrality measures when networks are sampled. *Social Networks, 25*(4), 283–307. https://doi.org/10.1016/S0378-8733(03)00012-1

Cramer, A. O. J., Lourens, J. W., Van Der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences, 33,* 137–193. https://doi.org/10.1017/S0140525X09991567

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., Kendler, K. S., & Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality, 26*(4), 414–431. https://doi.org/10.1002/per.1866

Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin*, *32*(7), 917–929. https://doi.org/10.1177/0146167206287721

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems*, *1695*. http://igraph.sf.net/

Curtis, T., Miller, B. C., & Berry, E. H. (2000). Changes in reports and incidence of child abuse following natural disasters. *Child Abuse & Neglect, 24*(9), 1151–1162. https://doi.org/10.1016/S0145-2134(00)00176-9

Curtiss, J., Fulford, D., Hofmann, S. G., & Gershon, A. (2019). Network dynamics of positive and negative affect in bipolar disorder. *Journal of Affective Disorders, 249*(January), 270–277. https://doi.org/10.1016/j.jad.2019.02.017

Dablander, F., Pichler, A., Cika, A., & Bacilieri, A. (2022). Anticipating critical transitions in psychological systems using early warning signals: Theoretical and practical considerations. *Psychological Methods,* 765–790. https://doi.org/10.1037/met0000450

Dalege, J., Borsboom, D., van Harreveld, F., van den Berg, H., Conner, M., & van der Maas, H. L. J. (2016). Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model. *Psychological Review*, *123*(1), 2–22. http://dx.doi.org/10.1037/a0039802

Dallos, R., & Draper, R. (2015). *An introduction to family therapy: Systemic theory and practice* (4th edition). McGraw-Hill Education (UK).

Davies, P. T., Sturge–Apple, M. L., & Cummings, E. M. (2004). Interdependencies among interparental discord and parenting practices: The role of adult vulnerability and relationship perturbations. *Development and Psychopathology, 16*(03). https://doi.org/10.1017/S0954579404004778 de Vos, S., Wardenaar, K. J., Bos, E. H., Wit, E. C., Bouwmans, M. E. J., & De Jonge, P. (2017). An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks. *PLoS ONE, 12*(6), 1–18. https://doi.org/10.1371/journal.pone.0178586

Dejonckheere, E., & Erbas, J. (2021). Designing an experience sampling study. In P. Kuppens & I. Germeys (Eds.), *The open handbook of Experience Sampling Methodology: A step-by-step guide to designing, conducting, and analyzing ESM studies* (pp. 33–70). Center for Research on Experience Sampling and Ambulatory Methods. https://www.kuleuven.be/samenwerking/real/real-book

Dejonckheere, E., & Mestdagh, M. (2021). On the signal-to-noise ratio in real-life emotional time series. In C. E. Waugh & P. Kuppens (Eds.), *Affect Dynamics* (pp. 131–152). Springer International Publishing. https://doi.org/10.1007/978-3-030-82965-0_7

Dejonckheere, E., & Mestdagh, M. (2023). *For better or for worse? Visualizing previous intensity levels improves emotion (dynamic) measurement in experience sampling*. PsyArXiv. https://doi.org/10.31234/osf.io/avbyc

Dejonckheere, E., Mestdagh, M., Houben, M., Rutten, I., Sels, L., Kuppens, P., & Tuerlinckx, F. (2019). Complex affect dynamics add limited information to the prediction of psychological well-being. *Nature Human Behaviour, 3*(5), 478–491. https://doi.org/10.1038/s41562-019-0555-0

Demerouti, E., Nachreiner, F., Bakker, A. B., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology, 86*(3), 499–512. https://doi.org/10.1037/0021-9010.86.3.499

Dupont, S., Mikolajczak, M., & Roskam, I. (2022). The cult of the child: A critical examination of its consequences on parents, teachers and children. *Social Sciences, 11*(3), Article 3. https://doi.org/10.3390/socsci11030141

Eisele, G., Kasanova, Z., & Houben, M. (2021). Questionnaire design and evaluation. In I. Germeys & P. Kuppens (Eds.), *The open handbook of Experience Sampling Methodology: A step-by-step guide to designing, conducting, and analyzing ESM studies* (pp. 71–89). Center for Research on Experience Sampling and Ambulatory Methods. https://www.kuleuven.be/samenwerking/real/real-book

Elliott, H., Jones, P. J., & Schmidt, U. (2020). Central symptoms predict post-treatment outcomes and clinical impairment in anorexia nervosa: A network analysis. *Clinical Psychological Science, 8*(1), 139–154. https://doi.org/10.1177/2167702619865958

Ellison, W. D., Levy, K. N., Newman, M. G., Pincus, A. L., Wilson, S. J., & Molenaar, P. C. M. (2020). Dynamics among borderline personality and anxiety features in psychotherapy outpatients: An exploration of nomothetic and idiographic patterns. *Personality Disorders: Theory, Research, and Treatment*, *11*(2), 131–140. https://doi.org/10.1037/per0000363

Elmer, T., Geschwind, N., Peeters, F., Wichers, M., & Bringmann, L. (2020). Getting stuck in social isolation: Solitude inertia and depressive symptoms. *Journal of Abnormal Psychology, 129,* 713–723. https://doi.org/10.1037/abn0000588

Elovainio, M., Kuula, L., Halonen, R., & Pesonen, A.-K. (2020). Dynamic fluctuations of emotional states in adolescents with delayed sleep phase—A longitudinal network modeling approach. *Journal of Affective Disorders, 276,* 467–475. https://doi.org/10.1016/j.jad.2020.07.050

Epskamp, S. (2020). Psychometric network models from time-series and panel data. *Psychometrika, 85*(1), 206–231. https://doi.org/10.1007/s11336-020-09697-3

Epskamp, S., Borsboom, D., & Fried, E. I. (2017). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods, 50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*(4), 1–18. https://doi.org/10.18637/jss.v048.i04

Epskamp, S., Deserno, M. K., & Bringmann, L. F. (2019). *MlVAR: Multi-level vector autoregression* [Computer software]. https://CRAN.R-project.org/package=mlVAR

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167

Epskamp, S., Kruis, J., & Marsman, M. (2017). Estimating psychopathological networks: Be careful what you wish for. *PLoS ONE, 12*(6). https://doi.org/10.1371/journal.pone.0179891

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika, 82*(4), 904–927. https://doi.org/10.1007/s11336-017-9557-x

Epskamp, S., van Borkulo, C. D., van der Veen, D. C., Servaas, M. N., Isvoranu, A. M., Riese, H., & Cramer, A. O. J. (2018). Personalized network modeling in psychopathology: The importance of contemporaneous and temporal connections. *Clinical Psychological Science*, *6*(3), 416–427. https://doi.org/10.1177/2167702617744325

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian Graphical Model in cross-sectional and time-series data. *Multivariate Behavioral Research, 53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Erbas, Y., Ceulemans, E., Lee Pe, M., Koval, P., & Kuppens, P. (2014). Negative emotion differentiation: Its personality and well-being correlates and a comparison of different assessment methods. *Cognition and Emotion, 28*(7), 1196–1213. https://doi.org/10.1080/02699931.2013.875890

Erdfelder, E., FAul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Eronen, M. I. (2020). Causal discovery and the problem of psychological interventions. *New Ideas in Psychology*, *59,* 100785. https://doi.org/10.1016/j.newideapsych.2020.100785

Eronen, M. I., & Bringmann, L. F. (2021). The theory crisis in psychology: How to move forward. *Perspectives on Psychological Science,* 1745691620970586. https://doi.org/10.1177/1745691620970586

Everaert, J., & Joormann, J. (2019). Emotion regulation difficulties related to depression and anxiety: A network approach to model relations among symptoms, positive reappraisal, and repetitive negative thinking. *Clinical Psychological Science, 7*(6), 1–15. https://doi.org/10.1177/2167702619859342

Faelens, L., Hoorelbeke, K., Soenens, B., Van Gaeveren, K., De Marez, L., De Raedt, R., & Koster, E. H. W. (2021). Social media use and well-being: A prospective experience-sampling study. *Computers in Human Behavior, 114*, 106510. https://doi.org/10.1016/j.chb.2020.106510

Fisher, A. J., & Bosley, H. G. (2020). Identifying the presence and timing of discrete mood states prior to therapy. *Behaviour Research and Therapy, 128*, 103596. https://doi.org/10.1016/j.brat.2020.103596

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences, 115*(27), E6106–E6115. https://doi.org/10.1073/pnas.1711978115

Fisher, A. J., Reeves, J. W., Lawyer, G., Medaglia, J. D., & Rubel, J. A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of Abnormal Psychology, 126*(8), 1044–1056. https://doi.org/10.1037/abn0000311

Fleeson, W., & Jayawickreme, E. (2015). Whole trait theory. *Journal of Research in Personality, 56*, 82–92. https://doi.org/10.1016/j.jrp.2014.10.009

Fombonne, E., Chedan, F., Carradec, A. M., Achard, S., & et al. (1988). Le Child Behavior Checklist: Un instrument pour la recherche en psychiatrie de l'enfant. [The Child Behavior Checklist: A research instrument for pediatric psychiatry.]. *Psychiatrie & Psychobiologie, 3*(6), 409–418.

Forrest, L. N., Sarfan, L. D., Ortiz, S. N., Brown, T. A., & Smith, A. R. (2019). Bridging eating disorder symptoms and trait anxiety in patients with eating disorders: A network approach. *International Journal of Eating Disorders, 52*(6), 701–711. https://doi.org/10.1002/eat.23070

Foygel, R., & Drton, M. (2011). *Bayesian model choice and information criteria in sparse generalized linear models*. https://arxiv.org/abs/1112.5635

Freudenberger, H. J. (1975). The staff burn-out syndrome in alternative institutions. *Psychotherapy: Theory, Research & Practice, 12*(1), 73–82. https://doi.org/10.1037/h0086411

Fried, E. I. (2015). Problematic assumptions have slowed down depression research: Why symptoms, not syndromes are the way forward. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00309

Fried, E. I. (2016, November 26). New paper: Differential variability of variables drives network structure. *Psych Networks*. https://psych-networks.com/new-paper-differential-variability-items-drives-network-structure/

Fried, E. I., & Nesse, R. M. (2014). The impact of individual depressive symptoms on impairment of psychosocial functioning. *PLOS ONE, 9*(2), e90311. https://doi.org/10.1371/journal.pone.0090311

Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders, 172*, 96–102. https://doi.org/10.1016/j.jad.2014.10.010

Friedman, J., Hastie, T., & Tibshirani, R. (2011). *glasso: Graphical lasso-estimation of Gaussian Graphical Models*.

Fülöp, F., Bőthe, B., Gál, É., Cachia, J. Y. A., Demetrovics, Z., & Orosz, G. (2022). A two-study validation of a single-item measure of relationship satisfaction: RAS-1. *Current Psychology, 41*, 2109–2121. https://doi.org/10.1007/s12144-020-00727-y

Furutani, K., Kawamoto, T., Alimardani, M., & Nakashima, K. (2020). Exhausted parents in Japan: Preliminary validation of the Japanese version of the Parental Burnout Assessment. *New Directions for Child and Adolescent Development, 2020*(174), 33–49. https://doi.org/10.1002/cad.20371

Gabriel, A. S., Podsakoff, N. P., Beal, D. J., Scott, B. A., Sonnentag, S., Trougakos, J. P., & Butts, M. M. (2019). Experience sampling methods: A discussion of critical trends and considerations for scholarly advancement. *Organizational Research Methods, 22*(4), 969–1006. https://doi.org/10.1177/1094428118802626

Gannagé, M., Besson, E., Harfouche, J., Roskam, I., & Mikolajczak, M. (2020). Parental burnout in Lebanon: Validation psychometric properties of the Lebanese Arabic version of the Parental Burnout Assessment. *New Directions for Child and Adolescent Development, 2020*(174), 51–65. https://doi.org/10.1002/cad.20383

Gillis, A., & Roskam, I. (2019). Daily exhaustion and support in parenting: Impact on the quality of the parent–child relationship. *Journal of Child and Family Studies, 28*, 2007–2016. https://doi.org/10.1007/s10826-019-01428-2

Gillis, A., & Roskam, I. (2020). Regulation between daily exhaustion and support in parenting: A dyadic perspective. *International Journal of Behavioral Development, 44*(3), 226–235. https://doi.org/10.1177/0165025419868536

Goerg, G. M. (2015). The Lambert way to Gaussianize heavy-tailed data with the inverse of Tukey's *h* transformation as a special case. *The Scientific World Journal, 2015*, 1–16. https://doi.org/10.1155/2015/909231

Granic, I., & Hollenstein, T. (2015). A survey of dynamic systems methods for developmental psychopathology. In *Developmental Psychopathology* (pp. 889–930). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470939383.ch22

Grave, R. B. de, Hasselman, F., & Bijleveld, E. (2022). From work stress to disease: A computational model. *PLOS ONE, 17*(2), e0263966. https://doi.org/10.1371/journal.pone.0263966

Gray, E. K., & Watson, D. (2007). Assessing positive and negative affect via self-report. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 171–183). Oxford University Press. https://books.google.com/books?hl=en&lr=&id=9xhnDAAAQBAJ&oi=fnd&pg=PA171&dq=Assessing+positive+and+negative+affect+via+self-report+gray&ots=nMIAcZkJy-&sig=YTCX5DdB8qwO4-3mmY_DrgY1we0

Greene, T., Gelkopf, M., Epskamp, S., & Fried, E. (2018). Dynamic networks of PTSD symptoms during conflict. *Psychological Medicine, 48*(14), 2409–2417. https://doi.org/10.1017/S0033291718000351

Greene, T., West, M., & Somer, E. (2020). Maladaptive daydreaming and emotional regulation difficulties: A network analysis. *Psychiatry Research, 285,* 112799. https://doi.org/10.1016/j.psychres.2020.112799

Griffith, A. K. (2022). Parental burnout and child maltreatment during the COVID-19 pandemic. *Journal of Family Violence, 37,* 725–731. https://doi.org/10.1007/s10896-020-00172-2

Groen, R. N., Ryan, O., Wigman, J. T. W., Riese, H., Penninx, B. W. J. H., Giltay, E. J., Wichers, M., & Hartman, C. A. (2020). Comorbidity between depression and anxiety: Assessing the role of bridge mental states in dynamic psychological networks. *BMC Medicine, 18*(1), 308. https://doi.org/10.1186/s12916-020-01738-z

Groen, R. N., Snippe, E., Bringmann, L. F., Simons, C. J. P., Hartmann, J. A., Bos, E. H., & Wichers, M. (2019). Capturing the risk of persisting depressive symptoms: A dynamic network investigation of patients' daily symptom experiences. *Psychiatry Research, 271,* 640–648. https://doi.org/10.1016/j.psychres.2018.12.054

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry, 26*(1), 1–26. https://doi.org/10.1080/1047840X.2014.940781

Guloksuz, S., Pries, L. K., & Van Os, J. (2017). Application of network methods for understanding mental disorders: Pitfalls and promise. *Psychological Medicine, 47*(16), 2743–2752. https://doi.org/10.1017/S0033291717001350

Halbesleben, J. R. B., & Bowler, W. M. (2007). Emotional exhaustion and job performance: The mediating role of motivation. *Journal of Applied Psychology, 92*(1), 93–106. https://doi.org/10.1037/0021-9010.92.1.93

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research, 53*(6), 820–841. https://doi.org/10.1080/00273171.2018.1446819

Hansotte, L., Nguyen, N., Roskam, I., Stinglhamber, F., & Mikolajczak, M. (2021). Are all burned out parents neglectful and violent? A latent profile analysis. *Journal of Child and Family Studies, 30*(1), 158–168. https://doi.org/10.1007/s10826-020-01850-x

Haslbeck, J. M. B., Bringmann, L. F., & Waldorp, L. J. (2020). A tutorial on estimating time-varying vector autoregressive models. *Multivariate Behavioral Research, 0*(0), 1–30. https://doi.org/10.1080/00273171.2020.1743630

Haslbeck, J. M. B., Ryan, O., Robinaugh, D. J., Waldorp, L. J., & Borsboom, D. (2022). Modeling psychopathology: From data models to formal theories. *Psychological Methods, 27*(6), 930.

Hasmi, L., Drukker, M., Guloksuz, S., Menne-Lothmann, C., Decoster, J., van Winkel, R., Collip, D., Delespaul, P., de Hert, M., Derom, C., Thiery, E., Jacobs, N., Rutten, B. P. F., Wichers, M., & van Os, J. (2017). Network approach to understanding emotion dynamics in relation to childhood trauma and genetic liability to psychopathology: Replication of a prospective experience sampling analysis. *Frontiers in Psychology, 8,* 1908. https://doi.org/10.3389/fpsyg.2017.01908

Hayes, A. M., & Strauss, J. L. (1998). Dynamic systems theory as a paradigm for the study of change in psychotherapy: An application to cognitive therapy for depression. *Journal of Consulting and Clinical Psychology, 66*(6), 939–947. http://dx.doi.org/10.1037/0022-006X.66.6.939

Hayes, A. M., Yasinski, C., Ben Barnes, J., & Bockting, C. L. H. (2015). Network destabilization and transition in depression: New methods for studying the dynamics of therapeutic change. *Clinical Psychology Review, 41*, 27–39. https://doi.org/10.1016/j.cpr.2015.06.007

Hays, S. (1996). *The cultural contradictions of motherhood*. Yale University Press.

Hays, S. (1998). The fallacious assumptions and unrealistic prescriptions of attachment theory: A comment on "Parents' Socioemotional Investment in Children." *Journal of Marriage and Family, 60*(3), 782–790. https://doi.org/10.2307/353546

Heeren, A., Bernstein, E. E., & McNally, R. J. (2018). Deconstructing trait anxiety: A network perspective. *Anxiety, Stress and Coping, 31*(3), 262–276. https://doi.org/10.1080/10615806.2018.1439263

Heeren, A., Bernstein, E. E., & McNally, R. J. (2020). Bridging maladaptive social self-beliefs and social anxiety: A network perspective. *Journal of Anxiety Disorders, 74*, 102267. https://doi.org/10.1016/j.janxdis.2020.102267

Heeren, A., Jones, P. J., & McNally, R. J. (2018). Mapping network connectivity among symptoms of social anxiety and comorbid depression in people with social anxiety disorder. *Journal of Affective Disorders, 228*, 75–82. https://doi.org/10.1016/j.jad.2017.12.003

Heeren, A., & McNally, R. J. (2016). An integrative network approach to social anxiety disorder: The complex dynamic interplay among attentional bias for threat, attentional control, and symptoms. *Journal of Anxiety Disorders, 42*, 95–104. https://doi.org/10.1016/j.janxdis.2016.06.009

Heeren, A., & McNally, R. J. (2018). Social anxiety disorder as a densely interconnected network of fear and avoidance for social situations. *Cognitive Therapy and Research, 42*, 103–113. https://doi.org/10.1007/s10608-017-9876-3

Heinemann, L. V., & Heinemann, T. (2017). Burnout research: Emergence and scientific investigation of a contested diagnosis. *SAGE Open, 7*(1), 2158244017697154. https://doi.org/10.1177/2158244017697154

Hektner, J., Schmidt, J., & Csikszentmihalyi, M. (2007). Psychometrics of ESM data. In *Experience Sampling Method: Measuring the quality of everyday life* (pp. 103–121). SAGE Publications, Inc. https://www.doi.org/10.4135/9781412984201

Henry, C. S., & Harrist, A. W. (2022). Family resilience theory. In K. Adamsons, A. L. Few-Demo, C. Proulx, & K. Roy (Eds.), *Sourcebook of Family Theories and Methodologies: A Dynamic Approach* (pp. 65–91). Springer International Publishing. https://doi.org/10.1007/978-3-030-92002-9_4

Hirota, M., Holmgren, M., Nes, E. H. V., & Scheffer, M. (2011). Global resilience of tropical forest and savanna to critical transitions. *Science, 334*(6053), 232–235. https://doi.org/10.1126/science.1210657

Hittner, J. B., May, K., & Silver, N. C. (2003). A Monte Carlo evaluation of tests for comparing dependent correlations. *Journal of General Psychology, 130*(2), 149–168. https://doi.org/10.1080/00221300309601282

Hoffart, A., & Johnson, S. U. (2020). Within-person networks of clinical features of social anxiety disorder during cognitive and interpersonal therapy. *Journal of Anxiety Disorders, 76,* 102312. https://doi.org/10.1016/j.janxdis.2020.102312

Hofmann, S. G., Curtiss, J. E., & Hayes, S. C. (2020). Beyond linear mediation: Toward a dynamic network approach to study treatment processes. *Clinical Psychology Review, 76,* 101824. https://doi.org/10.1016/j.cpr.2020.101824

Hofmann, S. G., & Hay, A. C. (2018). Rethinking avoidance: Toward a balanced approach to avoidance in treating anxiety disorders. *Journal of Anxiety Disorders, 55,* 14–21. https://doi.org/10.1016/j.janxdis.2018.03.004

Holland, P. W., & Rosenbaum, P. R. (1986). Conditional association and unidimensionality in monotone latent variable models. *The Annals of Statistics, 14*(4), 1523–1543. JSTOR.

Hollenstein, T. (2021). Affect dynamics and time scales: Pictures of movies. In C. E. Waugh & P. Kuppens (Eds.), *Affect Dynamics* (pp. 117–129). Springer International Publishing. https://doi.org/10.1007/978-3-030-82965-0_6

Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review, 30,* 349–362. https://doi.org/10.1016/j.cpr.2010.01.001

Hsin, A. (2009). Parent's time with children: Does time matter for children's cognitive achievement? *Social Indicators Research, 93*(1), 123–126. https://doi.org/10.1007/s11205-008-9413-6

Hubert, S., & Aujoulat, I. (2018). Parental burnout: When exhausted mothers open up. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.01021

Huckins, J. F., DaSilva, A. W., Hedlund, E. L., Murphy, E. I., Rogers, C., Wang, W., Obuchi, M., Holtzheimer, P. E., Wagner, D. D., & Campbell, A. T. (2020). Causal factors of anxiety and depression in college students: Longitudinal ecological momentary assessment and causal analysis using Peter and Clark Momentary Conditional Independence. *JMIR Mental Health, 7*(6), e16684. https://doi.org/10.2196/16684

Hutmacher, F., & Franz, D. (2024). Approaching psychology's current crises by exploring the vagueness of psychological concepts: Recommendations for advancing the discipline. *The American Psychologist*. https://doi.org/10.1037/amp0001300

Hyland, M. E. (2019). Functional disorders can also be explained through a non-reductionist application of network theory. *Behavioral and Brain Sciences; New York, 42*. http://dx.doi.org/10.1017/S0140525X18001048

Isvoranu, A.-M., & Epskamp, S. (2021). Which estimation method to choose in network psychometrics? Deriving guidelines for applied researchers. *Psychological Methods*.

Isvoranu, A.-M., Epskamp, S., & Cheung, M. (2021). Network models of post-traumatic stress disorder: A meta-analysis. *Journal of Abnormal Psychology*, *130*(8), 841–861. https://psycnet.apa.org/doi/10.1037/abn0000704

Jackson, D. (1965). The study of family. *Family Process, 4*, 1–20.

Jacobson, N. C., Chow, S.-M., & Newman, M. G. (2019). The Differential Time-Varying Effect Model (DTVEM): A tool for diagnosing and modeling time lags in intensive longitudinal data. *Behavior Research Methods, 51*(1), 295–315. https://doi.org/10.3758/s13428-018-1101-0

Jhangiani, R. S., Chiang, I.-C. A., Cuttler, C., & Leighton, D. C. (2019). Reliability and validity of measurement. In *Research Methods in Psychology* (4th ed.). https://kpu.pressbooks.pub/psychmethods4e/chapter/reliability-and-validity-of-measurement/

Ji, L., Chow, S.-M., Schermerhorn, A. C., Jacobson, N. C., & Cummings, E. M. (2018). Handling missing data in the modeling of intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal, 25*(5), 715–736. https://doi.org/10.1080/10705511.2017.1417046

Jiang, H., Fei, X., Liu, H., Roeder, K., Lafferty, J., Wasserman, L., Li, X., & Zhao, T. (2019). *huge: High-dimensional undirected graph estimation* (R package version 1.3.2) [Computer software].

Jones, P. J. (2018). *Networktools: Tools for identifying important nodes in networks* (R Package Version 1.1.1.) [Computer software].

Jones, P. J., Heeren, A., & McNally, R. J. (2017). Commentary: A network theory of mental disorders. *Frontiers in Psychology, 8*, 1305. https://doi.org/10.3389/fpsyg.2017.01305

Jones, P. J., Ma, R., & McNally, R. J. (2019). Bridge centrality: A network approach to understanding comorbidity. *Multivariate Behavioral Research*, 1–15. https://doi.org/10.1080/00273171.2019.1614898

Jones, P. J., Mair, P., & McNally, R. J. (2018). Visualizing psychological networks: A tutorial in R. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.01742

Jongeneel, A., Aalbers, G., Bell, I., Fried, E. I., Delespaul, P., Riper, H., van der Gaag, M., & van den Berg, D. (2020). A time-series network approach to auditory verbal hallucinations: Examining dynamic interactions using experience sampling methodology. *Schizophrenia Research, 215*, 148–156. https://doi.org/10.1016/j.schres.2019.10.055

Jordan, D. G., Winer, E. S., & Salem, T. (2020). The current status of temporal network analysis for clinical science: Considerations as the paradigm shifts? *Journal of Clinical Psychology*. https://doi.org/10.1002/jclp.22957

Jovanović, V., & Lazić, M. (2020). Is longer always better? A comparison of the validity of single-item versus multiple-item measures of life satisfaction. *Applied Research in Quality of Life, 15*(3), 675–692. https://doi.org/10.1007/s11482-018-9680-6

Kaiser, T., & Laireiter, A.-R. (2019). Daily dynamic assessment and modelling of intersession processes in ambulatory psychotherapy: A proof of concept study. *Psychotherapy Research, 29*(8), 1062–1073. https://doi.org/10.1080/10503307.2018.1497213

Kalis, A., & Borsboom, D. (2020). Folk psychology as a causal language. *Theory & Psychology, 30*(5), 723–728. https://doi.org/10.1177/0959354320933940

Kalkan, R. B., Blanchard, M. A., Mikolajczak, M., Roskam, I., & Heeren, A. (2023). Emotional exhaustion and feeling fed up as the driving forces of parental burnout and its consequences on children: Insights from a network approach. *Current Psychology, 42*(26), 22278–22289. https://doi.org/10.1007/s12144-022-03311-8

Kawamoto, T., Furutani, K., & Alimardani, M. (2018). Preliminary validation of Japanese version of the Parental Burnout Inventory and its relationship with perfectionism. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.00970

Kendler, K. S. (2005). "A gene for...": The nature of gene action in psychiatric disorders. *American Journal of Psychiatry, 162*(7), 1243–1252.

Kendler, K. S., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine, 41*(6), 1143–1150. https://doi.org/10.1017/S0033291710001844

Kerig, P. K. (2019). Parenting and family systems. In M. H. Bornstein (Ed.), *Handbook of Parenting: Vol. 3: Being and becoming a parent* (3rd ed., pp. 3–35). Routledge.

Kerr, M. L., Fanning, K. A., Huynh, T., Botto, I., & Kim, C. N. (2021). Parents' self-reported psychological impacts of COVID-19: Associations with parental burnout, child behavior, and income. *Journal of Pediatric Psychology, 46*(10), 1162–1171. https://doi.org/10.1093/jpepsy/jsab089

Kerr, M. L., Rasmussen, H. F., Buttitta, K. V., Smiley, P. A., & Borelli, J. L. (2021). Exploring the complexity of mothers' real-time emotions while caregiving. *Emotion, 21*(3), 545–556. https://doi.org/10.1037/emo0000719

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 617–627. https://doi.org/10.1001/archpsyc.62.6.617

Kirtley, O. J., Lafit, G., Achterhof, R., Hiekkaranta, A. P., & Myin-Germeys, I. (2021). Making the black box transparent: A template and tutorial for registration of studies using experience-sampling methods. *Advances in Methods and Practices in Psychological Science, 4*(1), 2515245920924686. https://doi.org/10.1177/2515245920924686

Kjellberg, S., Haider, J., & Sundin, O. (2016). Researchers' use of social network sites: A scoping review. *Library & Information Science Research, 38*(3), 224–234. https://doi.org/10.1016/j.lisr.2016.08.008

Klippel, A., Viechtbauer, W., Reininghaus, U., Wigman, J., van Borkulo, C., Myin-Germeys, I., & Wichers, M. (2018). The cascade of stress: A network approach to explore differential dynamics in populations varying in risk for psychosis. *Schizophrenia Bulletin, 44*(2), 328–337. https://doi.org/10.1093/schbul/sbx037

Kohl, C., McIntosh, E. J., Unger, S., Haddaway, N. R., Kecke, S., Schiemann, J., & Wilhelm, R. (2018). Online tools supporting the conduct and reporting of systematic reviews and systematic maps: A case study on CADIMA and review of existing tools. *Environmental Evidence, 7*(1), 8. https://doi.org/10.1186/s13750-018-0115-5

Komulainen, K., Airaksinen, J., Savelieva, K., Gluschkoff, K., García Velázquez, R., Elovainio, M., & Jokela, M. (2020). Network dynamics of depressive symptoms in antidepressant medication treatment: Secondary analysis of eight clinical trials. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-020-00884-3

Koval, P., Burnett, P. T., & Zheng, Y. (2021). Emotional inertia: On the conservation of emotional momentum. In C. E. Waugh & P. Kuppens (Eds.), *Affect Dynamics* (pp. 63–94). Springer International Publishing. https://doi.org/10.1007/978-3-030-82965-0_4

Koval, P., Pe, M. L., Meers, K., & Kuppens, P. (2013). Affect dynamics in relation to depressive symptoms: Variable, unstable or inert? *Emotion*, *13*(6), 1132–1141. https://doi.org/10.1037/a0033579

Kraft, B., Jonassen, R., Heeren, A., Harmer, C., Stiles, T., & Landrø, N. I. (2019). Attention bias modification in remitted depression is associated with increased interest and leads to reduced adverse impact of anxiety symptoms and negative cognition. *Clinical Psychological Science*, *7*(3), 530–544. https://doi.org/10.1177/2167702618822480

Kremer-Sadlik, T., & Paugh, A. L. (2007). Everyday moments: Finding "quality time" in American working families. *Time & Society*, *16*(2–3), 287–308. https://doi.org/10.1177/0961463X07080276

Kuppens, P. (2015). It's about time: A special section on affect dynamics. *Emotion Review*, *7*(4), 297–300. https://doi.org/10.1177/1754073915590947

Kuranova, A., Wigman, J. T. W., Menne-Lothmann, C., Decoster, J., van Winkel, R., Delespaul, P., Drukker, M., de Hert, M., Derom, C., Thiery, E., Rutten, B. P. F., Jacobs, N., van Os, J., Oldehinkel, A. J., Booij, S. H., & Wichers, M. (2021). Network dynamics of momentary affect states and future course of psychopathology in adolescents. *PLOS ONE*, *16*(3), e0247458. https://doi.org/10.1371/journal.pone.0247458

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, *54*(1–3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y

Lafit, G., Adolf, J. K., Dejonckheere, E., Myin-Germeys, I., Viechtbauer, W., & Ceulemans, E. (2021). Selection of the number of participants in intensive longitudinal studies: A user-friendly shiny app and tutorial for performing power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science*, *4*(1), 2515245920978738. https://doi.org/10.1177/2515245920978738

Lafit, G., Meers, K., & Ceulemans, E. (2022). A systematic study into the factors that affect the predictive accuracy of multilevel VAR(1) models. *Psychometrika*, *87*(2), 432–476. https://doi.org/10.1007/s11336-021-09803-z

Lamers, F., Hoogendoorn, A. W., Smit, J. H., van Dyck, R., Zitman, F. G., Nolen, W. A., & Penninx, B. W. (2012). Sociodemographic and psychiatric determinants of attrition in the Netherlands Study of Depression and Anxiety (NESDA). *Comprehensive Psychiatry*, *53*(1), 63–70. https://doi.org/10.1016/j.comppsych.2011.01.011

Lane, S. T., & Gates, K. M. (2017). Automated selection of robust individual-level structural equation models for time series data. *Structural Equation Modeling: A*

Multidisciplinary Journal, 24(5), 768–782. https://doi.org/10.1080/10705511.2017.1309978

Lange, J., Dalege, J., Borsboom, D., van Kleef, G. A., & Fischer, A. H. (2020). Toward an integrative psychometric model of emotions. *Perspectives on Psychological Science, 15*(2), 444–468. https://doi.org/10.1177/1745691619895057

Lapate, R. C., & Heller, A. S. (2020). Context matters for affective chronometry. *Nature Human Behaviour, 4*(7), Article 7. https://doi.org/10.1038/s41562-020-0860-7

Larzelere, R. E., Gunnoe, M. L., Roberts, M. W., & Ferguson, C. J. (2017). Children and parents deserve better parental discipline research: Critiquing the evidence for exclusively "Positive" parenting. *Marriage & Family Review, 53*(1), 24–35. https://doi.org/10.1080/01494929.2016.1145613

Lazarus, G., Sened, H., & Rafaeli, E. (2020). Subjectifying the personality state: Theoretical underpinnings and an empirical example. *European Journal of Personality, 34*(6), 1017–1036. https://doi.org/10.1002/per.2278

Le Vigouroux, S., Charbonnier, E., & Scola, C. (2022). Profiles and age-related differences in the expression of the three parental burnout dimensions. *European Journal of Developmental Psychology*, 1–20. https://doi.org/10.1080/17405629.2021.1990749

Lebow, J. L. (2020). Editorial: The systemic in couple and family research and couple and family therapy. *Family Process, 59*(1), 3–9. https://doi.org/10.1111/famp.12529

Lee, R. T., & Ashforth, B. E. (1993). A longitudinal study of burn-out among supervisors and managers: Comparisons between the Leiter and Maslach (1988) and Golembiewski et al. (1986) models. *Organizational Behavior & Human Decision Processes, 54*(3), 369–398.

Leiter, M. P. (1993). Burnout as a developmental process: Consideration of models. In W. Schaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout: Recent developments in theory and research*. Taylor & Francis.

Leiter, M. P., & Maslach, C. (1988). The impact of interpersonal environment on burnout and organizational commitment. *Journal of Organizational Behavior, 9*, 297–308. https://doi.org/10.1002/job.4030090402

Levinson, C. A., Brosof, L. C., Vanzhula, I., Christian, C., Jones, P., Rodebaugh, T. L., Langer, J. K., White, E. K., Warren, C., Weeks, J. W., Menatti, A., Lim, M. H., & Fernandez, K. C. (2018). Social anxiety and eating disorder comorbidity and underlying vulnerabilities: Using network analysis to conceptualize comorbidity. *International Journal of Eating Disorders, 51*(7), 693–709. https://doi.org/10.1002/eat.22890

Levinson, C. A., Vanzhula, I. A., Smith, T. W., & Stice, E. (2020). Group and longitudinal intra-individual networks of eating disorder symptoms in adolescents and young adults at-risk for an eating disorder. *Behaviour Research and Therapy, 135*, 103731. https://doi.org/10.1016/j.brat.2020.103731

Levinson, C. A., Vanzhula, I., & Brosof, L. C. (2018). Longitudinal and personalized networks of eating disorder cognitions and behaviors: Targets for precision intervention a proof of concept study. *International Journal of Eating Disorders, 51*(11), 1233–1243. https://doi.org/10.1002/eat.22952

Lewis, M. D., Lamey, A. V., & Douglas, L. (1999). A new dynamic systems method for the analysis of early socioemotional development. *Developmental Science, 2*(4), 457–475. https://doi.org/10.1111/1467-7687.00090

Lin, G.-X., Hansotte, L., Szczygieł, D., Meeussen, L., Roskam, I., & Mikolajczak, M. (2021). Parenting with a smile: Display rules, regulatory effort, and parental burnout. *Journal of Social and Personal Relationships, 38*(9), 2701–2721. https://doi.org/10.1177/02654075211019124

Lin, G.-X., Mikolajczak, M., Keller, H., Akgun, E., Arikan, G., Aunola, K., Barham, E., Besson, E., Blanchard, M. A., Boujut, E., Brianda, M. E., Brytek-Matera, A., César, F., Chen, B.-B., Dorard, G., dos Santos Elias, L. C., Dunsmuir, S., Egorova, N., Escobar, M. J., … Roskam, I. (2023). Parenting culture(s): Ideal-parent beliefs across 37 countries. *Journal of Cross-Cultural Psychology, 54*(1), 4–24. https://doi.org/10.1177/00220221221123043

Lin, G.-X., Szczygieł, D., Hansotte, L., Roskam, I., & Mikolajczak, M. (2021). Aiming to be perfect parents increases the risk of parental burnout, but emotional competence mitigates it. *Current Psychology*. https://doi.org/10.1007/s12144-021-01509-w

Lindahl Norberg, A. (2007). Burnout in mothers and fathers of children surviving brain tumour. *Journal of Clinical Psychology in Medical Settings, 14*(2), 130–137. https://doi.org/10.1007/s10880-007-9063-x

Lindström, C., Åman, J., & Norberg, A. L. (2011). Parental burnout in relation to sociodemographic, psychosocial and personality factors as well as disease duration and glycaemic control in children with Type 1 diabetes mellitus. *Acta Paediatrica, 100*(7), 1011–1017. https://doi.org/10.1111/j.1651-2227.2011.02198.x

Liss, M., Schiffrin, H. H., Mackintosh, V. H., Miles-McLean, H., & Erchull, M. J. (2013). Development and validation of a quantitative measure of intensive parenting attitudes. *Journal of Child and Family Studies, 22*(5), 621–636. https://doi.org/10.1007/s10826-012-9616-y

López-Nicolás, R., López-López, J. A., Rubio-Aparicio, M., & Sánchez-Meca, J. (2021). A meta-review of transparency and reproducibility-related reporting practices in published meta-analyses on clinical psychological interventions (2000–2020). *Behavior Research Methods*. https://doi.org/10.3758/s13428-021-01644-z

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy, 33*(3), 335–343. https://doi.org/10.1016/0005-7967(94)00075-U

Lubiewska, K., Głogowska, K., Sümer, N., Żegleń, M., & Mądrzycka-Borska, Z. (2023). Is distanced mothering a neglectful or normative parenting strategy? Evidence from three countries. *International Journal of Intercultural Relations, 96*, 101842. https://doi.org/10.1016/j.ijintrel.2023.101842

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), 3139. https://doi.org/10.21105/joss.03139

Lüdtke, O., & Trautwein, U. (2007). Aggregating to the between-person level in idiographic research designs: Personal goal research as an example of the need to distinguish between reliability and homogeneity. *Journal of Research in Personality, 41*(1), 230–238. https://doi.org/10.1016/j.jrp.2006.03.005

Lundgren-Nilsson, Å., Jonsdottir, I. H., Pallant, J., & Ahlborg, G. (2012). Internal construct validity of the Shirom-Melamed Burnout Questionnaire (SMBQ). *BMC Public Health, 12*(1), 1. https://doi.org/10.1186/1471-2458-12-1

Lutz, W., Schwartz, B., Hofmann, S. G., Fisher, A. J., Husen, K., & Rubel, J. A. (2018). Using network analysis for the prediction of treatment dropout in patients with mood and anxiety disorders: A methodological proof-of-concept study. *Scientific Reports, 8*(1), Article 1. https://doi.org/10.1038/s41598-018-25953-0

Lydon-Staley, D. M., Barnett, I., Satterthwaite, T. D., & Bassett, D. S. (2019). Digital phenotyping for psychiatry: Accommodating data and theory with network science methodologies. *Current Opinion in Biomedical Engineering, 9,* 8–13. https://doi.org/10.1016/j.cobme.2018.12.003

Maciejewski, D. F., Keijsers, L., van Lier, P. A. C., Branje, S. J. T., Meeus, W. H. J., & Koot, H. M. (2019). Most fare well—But some do not: Distinct profiles of mood variability development and their association with adjustment during adolescence. *Developmental Psychology, 55*(2), 434–448. https://doi.org/10.1037/dev0000650

Maciejewski, D. F., van Roekel, E., Ha, T., De France, K., Lin, L., Lennarz, H., Trompetter, H., Meeus, W., Lichtwarck-Aschoff, A., & Branje, S. (2022). *Beyond main effects? Affect level as a moderator in the relation between affect dynamics and depressive symptoms.* PsyArXiv. https://osf.io/preprints/psyarxiv/42uzq

Malinen, K., Rönkä, A., Sevón, E., & Schoebi, D. (2017). The difficulty of being a professional, a parent, and a spouse on the same day: Daily spillover of workplace interactions on parenting, and the role of spousal support. *Journal of Prevention & Intervention in the Community, 45*(3), 156–167. https://doi.org/10.1080/10852352.2016.1198121

Martín-Brufau, R., Suso-Ribera, C., & Corbalán, J. (2020). Emotion network analysis during COVID-19 quarantine - a longitudinal study. *Frontiers in Psychology, 11,* 559572. https://doi.org/10.3389/fpsyg.2020.559572

Maslach, C. (1976). *Burned-out. 9*(5), 16–22.

Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior, 2*(2), 99–113. https://doi.org/10.1002/job.4030020205

Maslach, C., Jackson, S. E., & Leiter, M. P. (2010). *Maslach Burnout Inventory Manual.* Mind Garden.

Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology, 52*(1), 397–422. https://doi.org/10.1146/annurev.psych.52.1.397

Masten, A. S. (2018). Resilience theory and research on children and families: Past, present, and promise. *Journal of Family Theory & Review, 10*(1), 12–31. https://doi.org/10.1111/jftr.12255

Masten, A. S., & Palmer, A. R. (2019). Parenting to promote resilience in children. In M. H. Bornstein (Ed.), *Handbook of Parenting: Vol. 5: The Practice of Parenting* (3rd ed., pp. 156–188). Routledge.

Matias, M., Aguiar, J., César, F., Braz, A. C., Barham, E. J., Leme, V., Elias, L., Gaspar, M. F., Mikolajczak, M., Roskam, I., & Fontaine, A. M. (2020). The Brazilian–Portuguese version of the Parental Burnout Assessment: Transcultural adaptation and initial validity evidence. *New Directions for Child and Adolescent Development, 2020*(174), 67–83. https://doi.org/10.1002/cad.20374

McCuish, E., Bouchard, M., & Beauregard, E. (2021). A network-based examination of the longitudinal association between psychopathy and offending versatility. *Journal of Quantitative Criminology, 37*(3), 693–714. https://doi.org/10.1007/s10940-020-09462-w

McNally, R. J. (2016). Can network analysis transform psychopathology? *Behaviour Research and Therapy, 86,* 95–104. https://doi.org/10.1016/j.brat.2016.06.006

McNally, R. J. (2021). Network analysis of psychopathology: Controversies and challenges. *Annual Review of Clinical Psychology, 17*(1), 31–53. https://doi.org/10.1146/annurev-clinpsy-081219-092850

McNally, R. J., Heeren, A., & Robinaugh, D. J. (2017). A Bayesian network analysis of posttraumatic stress disorder symptoms in adults reporting childhood sexual abuse. *European Journal of Psychotraumatology, 8*(sup3), 1341276. https://doi.org/10.1080/20008198.2017.1341276

McNally, R. J., Mair, P., Mugno, B. L., & Riemann, B. C. (2017). Comorbid obsessive-compulsive disorder and depression: A Bayesian network approach. *Psychological Medicine, 47,* 1204–1214.

McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods, 25*(5), 610–635. https://doi.org/10.1037/met0000250

Meeussen, L., & Van Laar, C. (2018). Feeling pressure to be a perfect mother relates to parental burnout and career ambitions. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.02113

Melamed, S., Ugarten, U., Shirom, A., Kahana, L., Lerman, Y., & Froom, P. (1999). Chronic burnout, somatic arousal and elevated salivary cortisol levels. *Journal of Psychosomatic Research, 46*(6), 591–598. https://doi.org/10.1016/S0022-3999(99)00007-0

Meng, J., Wang, X., Wei, D., & Qiu, J. (2020). State loneliness is associated with emotional hypervigilance in daily life: A network analysis. *Personality and Individual Differences, 165,* 110154. https://doi.org/10.1016/j.paid.2020.110154

Mestdagh, M., Pe, M., Pestman, W., Verdonck, S., Kuppens, P., & Tuerlinckx, F. (2018). Sidelining the mean: The relative variability index as a generic mean-corrected variability measure for bounded variables. *Psychological Methods, 23*(4), 690–707. https://doi.org/10.1037/met0000153

Mikolajczak, M., Brianda, M. E., Avalosse, H., & Roskam, I. (2018). Consequences of parental burnout: Its specific effect on child neglect and violence. *Child Abuse and Neglect, 80,* 134–145. https://doi.org/10.1016/j.chiabu.2018.03.025

Mikolajczak, M., Gross, J. J., & Roskam, I. (2019). Parental burnout: What is it, and why does it matter? *Clinical Psychological Science, 7*(6), 1319–1329. https://doi.org/10.1177/2167702619858430

Mikolajczak, M., Gross, J. J., Stinglhamber, F., Lindahl Norberg, A., & Roskam, I. (2020). Is Parental Burnout Distinct From Job Burnout and Depressive Symptoms? *Clinical Psychological Science, 8*(4), 673–689. https://doi.org/10.1177/2167702620917447

Mikolajczak, M., Raes, M. E., Avalosse, H., & Roskam, I. (2018). Exhausted parents: Sociodemographic, child-related, parent-related, parenting and family-functioning correlates of parental burnout. *Journal of Child and Family Studies, 27*(2), 602–614. https://doi.org/10.1007/s10826-017-0892-4

Mikolajczak, M., & Roskam, I. (2018). A theoretical and clinical framework for parental burnout: The balance between risks and resources (BR2). *Frontiers in Psychology, 9*, 886. https://doi.org/10.3389/fpsyg.2018.00886

Minuchin, S. (1974). *Families and family therapy*. Harvard University Press.

Mneimne, M., Fleeson, W., Arnold, E. M., & Furr, R. M. (2018). Differentiating the everyday emotion dynamics of borderline personality disorder from major depressive disorder and bipolar disorder. *Personality Disorders: Theory, Research, and Treatment, 9*(2), 192–196. https://doi.org/10.1037/per0000255

Mofsen, A. M., Rodebaugh, T. L., Nicol, G. E., Depp, C. A., Miller, J. P., & Lenze, E. J. (2019). When all else fails, listen to the patient: A viewpoint on the use of ecological momentary assessment in clinical trials. *JMIR Mental Health, 6*(5), e11845. https://doi.org/10.2196/11845

Mousavi, S. F., Mikolajczak, M., & Roskam, I. (2020). Parental burnout in Iran: Psychometric properties of the Persian (Farsi) version of the Parental Burnout Assessment (PBA). *New Directions for Child and Adolescent Development, 2020*(174), 85–100. https://doi.org/10.1002/cad.20369

Mowrer, O. H. (1960). *Learning theory and behavior* (pp. xii, 555). John Wiley & Sons Inc. https://doi.org/10.1037/10802-000

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology, 18*(1), 143. https://doi.org/10.1186/s12874-018-0611-x

Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., & Reininghaus, U. (2018). Experience sampling methodology in mental health research: New insights and technical developments. *World Psychiatry, 17*(2), 123–132. https://doi.org/10.1002/wps.20513

Nakamura, T., Kim, J., Sasaki, T., Yamamoto, Y., Takei, K., & Taneichi, S. (2013). Intermittent locomotor dynamics and its transitions in bipolar disorder. *2013 22nd International*

Conference on Noise and Fluctuations (ICNF), 1–4. https://doi.org/10.1109/ICNF.2013.6578924

Neal, J. W., & Neal, Z. P. (2013). Nested or networked? Future directions for ecological systems theory. *Social Development, 22*(4), 722–737. https://doi.org/10.1111/sode.12018

Nehrkorn-Bailey, A. M., Reardon, M. S., & Hicks Patrick, J. (2018). Some methodological and analytical issues related to real-time data capture studies. *Translational Issues in Psychological Science, 4*(4), 349–361. http://dx.doi.org.proxy.bib.ucl.ac.be/10.1037/tps0000177

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In *Visual Information Processing* (pp. 283–308). Elsevier. https://doi.org/10.1016/B978-0-12-170150-5.50012-3

Nook, E. C. (2021). Emotion differentiation and youth mental health: Current understanding and open questions. *Frontiers in Psychology, 12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.700298

Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2011). A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods, 16*(4), 468–490. https://doi.org/10.1037/a0024375

Ottenstein, C., & Lischetzke, T. (2020). Development of a novel method of emotion differentiation that uses open-ended descriptions of momentary affective states. *Assessment, 27*(8), 1928–1945. https://doi.org/10.1177/1073191119839138

Ou, L., Hunter, M., D., & Chow, S.-M. (2019). Whats for dynr: A package for linear and nonlinear dynamic modeling in R. *The R Journal, 11*(1), 91. https://doi.org/10.32614/RJ-2019-012

Oude Maatman, F. (2020). Reformulating the network theory of mental disorders: Folk psychology as a factor, not a fact. *Theory & Psychology, 30*(5), 703–722. https://doi.org/10.1177/0959354320921464

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Palmier-Claus, J. E., Myin-Germeys, I., Barkus, E., Bentley, L., Udachina, A., Delespaul, P. a. E. G., Lewis, S. W., & Dunn, G. (2011). Experience sampling research in individuals with mental illness: Reflections and guidance. *Acta Psychiatrica Scandinavica, 123*(1), 12–20. https://doi.org/10.1111/j.1600-0447.2010.01596.x

Papini, S., Rubin, M., Telch, M. J., Smits, J. A. J., & Hien, D. A. (2020). Pretreatment posttraumatic stress disorder symptom network metrics predict the strength of the association between node change and network change during treatment. *Journal of Traumatic Stress, 33*(1), 64–71. https://doi.org/10.1002/jts.22379

Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, *52*(2), 183–199. https://doi.org/10.1016/j.im.2014.08.008

Pavani, J.-B., Vigouroux, S. L., Kop, J.-L., Congard, A., & Dauvier, B. (2017). A network approach to affect regulation dynamics and personality trait-induced variations: Extraversion and neuroticism moderate reciprocal influences between affect and affect regulation strategies. *European Journal of Personality*, *31*(4), 329–346. https://doi.org/10.1002/per.2109

Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal Inference in Statistics: A Primer*. John Wiley & Sons.

Pelsma, D. M., Roland, B., Tollefson, N., & Wigington, H. (1989). Parent burnout: Validation of the Maslach Burnout Inventory with a sample of mothers. *Measurement and Evaluation in Counseling and Development*, *22*(2), 81–87. https://doi.org/10.1080/07481756.1989.12022915

Penz, M., Stalder, T., Miller, R., Ludwig, V. M., Kanthak, M. K., & Kirschbaum, C. (2018). Hair cortisol as a biological marker for burnout symptomatology. *Psychoneuroendocrinology*, *87*, 218–221. https://doi.org/10.1016/j.psyneuen.2017.07.485

Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., & Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *International Journal of Evidence-Based Healthcare*, *13*(3), 141–146. https://doi.org/10.1097/XEB.0000000000000050

Petris, G. (2010). An R package for dynamic linear models. *Journal of Statistical Software*, *36*, 1–16. https://doi.org/10.18637/jss.v036.i12

Pines, A., & Aronson, E. (1988). *Career burnout: Causes and cures*. Free press.

Pirla, S., Taquet, M., & Quoidbach, J. (2022). Measuring affect dynamics: An empirical framework. *Behavior Research Methods*, *55*(1), 285–300. https://doi.org/10.3758/s13428-022-01829-0

Procaccini, J., & Kiefaber, M. (1984). *Parent burnout*. Penguin Publishing Group.

R Core Team. (2021). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rachman, S. (1978). *Fear and courage*. Freeman. http://www.sciencedirect.com/science/article/pii/S0005789484800453

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21–33.

Reichardt, J., & Bornholdt, S. (2006). Statistical mechanics of community detection. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *74*(1), 1–14. https://doi.org/10.1103/PhysRevE.74.016110

Revelle, W. (2020). *psych: Procedures for psychological, psychometric, and personality research* [Computer software]. Northwestern University. https://CRAN.R-project.org/package=psych

Reynolds, B., Robles, T., & Repetti, R. (2015). Measurement reactivity and fatigue effects in daily diary research with families. *Developmental Psychology, 52*. https://doi.org/10.1037/dev0000081

Robinaugh, D. J., Haslbeck, J. M. B., Waldorp, L., Kossakowski, J., Fried, E. I., Millner, A., McNally, R. J., Nes, E. H. van, Scheffer, M., Kendler, K. S., & Borsboom, D. (2019). *Advancing the network theory of mental disorders: A computational model of panic disorder*. PsyArXiv. https://doi.org/10.31234/osf.io/km37w

Robinaugh, D. J., Hoekstra, R. H. A., Toner, E. R., & Borsboom, D. (2019). The network approach to psychopathology: A review of the literature 2008 – 2018 and an agenda for future research. *Psychological Medicine, 50*(3), 1–14. https://doi.org/10.1017/S0033291719003404

Robinaugh, D. J., LeBlanc, N. J., Vuletich, H. A., & McNally, R. J. (2014). Network analysis of persistent complex bereavement disorder in conjugally bereaved adults. *Journal of Abnormal Psychology, 123*(3), 510–522. http://dx.doi.org/10.1037/abn0000002

Robinaugh, D. J., Millner, A. J., & McNally, R. J. (2016). Identifying highly influential nodes in the complicated grief network. *Journal of Abnormal Psychology, 125*(6), 747–757. https://doi.org/10.1037/abn0000181

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-esteem Scale. *Personality and Social Psychology Bulletin, 27*(2), 151–161. https://doi.org/10.1177/0146167201272002

Rodebaugh, T. L., Tonge, N. A., Piccirillo, M. L., Fried, E., Horenstein, A., Morrison, A. S., Goldin, P., Gross, J. J., Lim, M. H., Fernandez, K., Blanco, C., Schneier, F. R., Bogdan, R., Thompson, R. J., & Heimberg, R. G. (2018). Does centrality in a cross-sectional network suggest intervention targets for social anxiety disorder? *Journal of Consulting and Clinical Psychology, 86*(10), 831–844. https://doi.org/10.1037/ccp0000336

Rodrigo, M. J. (2010). Promoting positive parenting in Europe: New challenges for the European Society for Developmental Psychology. *European Journal of Developmental Psychology, 7*(3), 281–294. https://doi.org/10.1080/17405621003780200

Rodriguez, C. M., & Silvia, P. J. (2022). Spotlight on maternal perceptions of child behavior: A daily diary study with child welfare-involved mothers. *Behavioral Sciences, 12*(2), Article 2. https://doi.org/10.3390/bs12020044

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science, 1*(1), 27–42. https://doi.org/10.1177/2515245917745629

Roskam, I., Aguiar, J., Akgun, E., Arikan, G., Artavia, M., Avalosse, H., Aunola, K., Bader, M., Bahati, C., Barham, E. J., Besson, E., Beyers, W., Boujut, E., Brianda, M. E., Brytek-Matera, A., Carbonneau, N., César, F., Chen, B.-B., Dorard, G., ...

Mikolajczak, M. (2021). Parental burnout around the globe: A 42-country study. *Affective Science, 2*(1), 58–79. https://doi.org/10.1007/s42761-020-00028-4

Roskam, I., Brianda, M. E., & Mikolajczak, M. (2018). A step forward in the conceptualization and measurement of parental burnout: The Parental Burnout Assessment (PBA). *Frontiers in Psychology, 9,* 758. https://doi.org/10.3389/fpsyg.2018.00758

Roskam, I., & Mikolajczak, M. (2021). The slippery slope of parental exhaustion: A process model of parental burnout. *Journal of Applied Developmental Psychology, 77,* 101354. https://doi.org/10.1016/j.appdev.2021.101354

Roskam, I., Raes, M. E., & Mikolajczak, M. (2017). Exhausted parents: Development and preliminary validation of the Parental Burnout Inventory. *Frontiers in Psychology, 8,* 163. https://doi.org/10.3389/fpsyg.2017.00163

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91*(434), 473–489. https://doi.org/10.2307/2291635

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., & Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances, 5*(11), eaau4996. https://doi.org/10.1126/sciadv.aau4996

Ryan, O., Bringmann, L. F., & Schuurman, N. K. (2022). The challenge of generating causal hypotheses using network models. *Structural Equation Modeling: A Multidisciplinary Journal*, 953–970. https://doi.org/10.31234/osf.io/ryg69

Ryan, O., & Hamaker, E. L. (2022). Time to intervene: A continuous-time approach to network analysis and centrality. *Psychometrika, 87*(1), 214–252. https://doi.org/10.1007/s11336-021-09767-0

Ryan, O., Kuiper, R. M., & Hamaker, E. L. (2018). A continuous-time approach to intensive longitudinal data: What, why, and how? In K. van Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous Time Modeling in the Behavioral and Related Sciences* (pp. 27–54). Springer International Publishing. https://doi.org/10.1007/978-3-319-77219-6_2

Savelieva, K., Komulainen, K., Elovainio, M., & Jokela, M. (2021). Longitudinal associations between specific symptoms of depression: Network analysis in a prospective cohort study. *Journal of Affective Disorders, 278,* 99–106. https://doi.org/10.1016/j.jad.2020.09.024

Schaufeli, W. B., & Taris, T. W. (2005). The conceptualization and measurement of burnout: Common ground and worlds apart. *Work & Stress, 19*(3), 256–262. https://doi.org/10.1080/02678370500385913

Schaumberg, R. L., & Skowronek, S. E. (2022). Shame broadcasts social norms: The positive social effects of shame on norm acquisition and normative behavior. *Psychological Science, 33*(8), 1257–1277. https://doi.org/10.1177/09567976221075303

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M., & Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature, 461*(7260), Article 7260. https://doi.org/10.1038/nature08227

Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology, 31*(1), 43–53. https://doi.org/10.1016/j.newideapsych.2011.02.007

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*(6), 323–338. https://doi.org/10.3200/JOER.99.6.323-338

Scutari, M. (2010). Learning Bayesian networks with the bnlearn R package. *Journal of Statistical Software, 35*(3), 1–22. https://doi.org/10.18637/jss.v035.i03

Scutari, M., & Denis, J. (2015). *Bayesian networks: With examples in R*. CRC Press: Boca Raton, FL.

Scutari, M., & Nagarajan, R. (2013). Identifying significant edges in graphical models of molecular networks. *Artificial Intelligence in Medicine, 57*, 207–217.

Shang, Y. (2015). Deffuant model of opinion formation in one-dimensional multiplex networks. *Journal of Physics A: Mathematical and Theoretical, 48*(39), 395101. https://doi.org/10.1088/1751-8113/48/39/395101

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*(3/4), 591–611. https://doi.org/10.2307/2333709

Shevlyakov, G., & Vilchevski, N. (2001). *Robustness in data analysis: Criteria and methods*. De Gruyter. https://doi.org/10.1515/9783110936001

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428.

Shrout, P., & Lane, S. P. (2012). Psychometrics. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 302–320). Guildford Press.

Slipetz, L. R., Eberle, J. W., Levinson, C. A., Falk, A., Cusack, C. E., & Henry, T. R. (2023). *Analyzing ecological momentary assessment data with state-space models: Considerations and recommendations* (arXiv:2305.03214). arXiv. http://arxiv.org/abs/2305.03214

Smaldino, P. E. (2017). Models are stupid, and we need more of them. In R. R. Vallacher, A. Nowak, & S. J. Read (Eds.), *Computational social psychology* (pp. 311–331). Psychology Press.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. SAGE Publications, Inc.

Somers, J. A., Kerr, M. L., McNeish, D., Smiley, P. A., Buttitta, K. V., Rasmussen, H. F., & Borelli, J. L. (2020). Quantitatively representing real-time emotion dynamics: Attachment-based differences in mothers' emotion experiences. *Journal of Family Psychology, 34*(4), 480–489. https://doi.org/10.1037/fam0000617

Sorkkila, M., & Aunola, K. (2020). Risk factors for Parental Burnout among Finnish parents: The role of socially prescribed perfectionism. *Journal of Child and Family Studies, 29*(3), 648–659. https://doi.org/10.1007/s10826-019-01607-1

Spanakis, G., Weiss, G., Boh, B., & Roefs, A. (2016). Network analysis of ecological momentary assessment data for monitoring and understanding eating behavior. In X. Zheng, D. D. Zeng, H. Chen, & S. J. Leischow (Eds.), *Smart Health* (pp. 43–54). Springer International Publishing. https://doi.org/10.1007/978-3-319-29175-8_5

Spiller, T. R., Levi, O., Neria, Y., Suarez-Jimenez, B., Bar-Haim, Y., & Lazarov, A. (2020). On the validity of the centrality hypothesis in cross-sectional between-subject networks of psychopathology. *BMC Medicine, 18*(1), 297. https://doi.org/10.1186/s12916-020-01740-5

Spiller, T. R., Weilenmann, S., Prakash, K., Schnyder, U., von Känel, R., & Pfaltz, M. C. (2021). Emotion network density in burnout. *BMC Psychology, 9*(1), 170. https://doi.org/10.1186/s40359-021-00670-y

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092. https://doi.org/10.1001/archinte.166.10.1092

Stalder, T., & Kirschbaum, C. (2012). Analysis of cortisol in hair – State of the art and future directions. *Brain, Behavior, and Immunity, 26*(7), 1019–1029. https://doi.org/10.1016/j.bbi.2012.02.002

Stănculescu, E., Roskam, I., Mikolajczak, M., Muntean, A., & Gurza, A. (2020). Parental burnout in Romania: Validity of the Romanian version of the parental burnout assessment (PBA-RO). *New Directions for Child and Adolescent Development, 2020*(174), 119–136. https://doi.org/10.1002/cad.20384

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science, 11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Steinskog, D. J., Tjøstheim, D. B., & Kvamstø, N. G. (2007). A cautionary note on the use of the Kolmogorov–Smirnov test for normality. *Monthly Weather Review, 135*(3), 1151–1157. https://doi.org/10.1175/MWR3326.1

Super, C. M., & Harkness, S. (2020). Research on parental burnout across cultures: Steps toward global understanding. *New Directions for Child and Adolescent Development, 2020*(174), 185–192. https://doi.org/10.1002/cad.20389

Sydow, K. von, Retzlaff, R., Beher, S., Haun, M. W., & Schweitzer, J. (2013). The efficacy of systemic therapy for childhood and adolescent externalizing disorders: A systematic review of 47 RCT. *Family Process, 52*(4), 576–618. https://doi.org/10.1111/famp.12047

Szczygieł, D., Sekulowicz, M., Kwiatkowski, P., Roskam, I., & Mikolajczak, M. (2020). Validation of the Polish version of the Parental Burnout Assessment (PBA). *New Directions for Child and Adolescent Development, 2020*(174), 137–158. https://doi.org/10.1002/cad.20385

Taylor, S., Landry, C. A., Paluszek, M. M., Rachor, G. S., & Asmundson, G. J. G. (2020). Worry, avoidance, and coping during the COVID-19 pandemic: A comprehensive network analysis. *Journal of Anxiety Disorders,* 102327. https://doi.org/10.1016/j.janxdis.2020.102327

Teoh, Y. Y., Cunningham, W. A., & Hutcherson, C. A. (2023). Framing subjective emotion reports as dynamic affective decisions. *Affective Science, 4*(3), 522–528. https://doi.org/10.1007/s42761-023-00197-y

Terluin, B., De Boer, M. R., & De Vet, H. C. W. (2016). Differences in connection strength between mental symptoms might be explained by differences in variance: Reanalysis of network data did not confirm staging. *PLoS ONE, 11*(11), 1–12. https://doi.org/10.1371/journal.pone.0155205

Traag, V. A., & Bruggeman, J. (2009). Community detection in networks with positive and negative links. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *80*(3), 1–6. https://doi.org/10.1103/PhysRevE.80.036115

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., ... Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine, 169*(7), 467–473. https://doi.org/10.7326/M18-0850

Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment, 21*(4), 457–462. https://doi.org/10.1037/a0017653

Trull, T. J., & Ebner-Priemer, U. W. (2020). Ambulatory assessment in psychopathology research: A review of recommended reporting guidelines and current practices. *Journal of Abnormal Psychology, 129*(1), 56–63. http://dx.doi.org/10.1037/abn0000473

Ulferts, H. (2020). *Why parenting matters for children in the 21st century: An evidence-based framework for understanding parenting and its impact on child development*. Organisation for Economic Cooperation and Development (OECD). https://doi.org/10.1787/129a1a59-en

Vachon, H., Viechtbauer, W., Rintala, A., & Myin-Germeys, I. (2019). Compliance and retention with the experience sampling method over the continuum of severe mental disorders: Meta-analysis and recommendations. *Journal of Medical Internet Research, 21*(12), e14475. https://doi.org/10.2196/14475

Van Borkulo, C., Boschloo, L., Borsboom, D., Penninx, B. W. J. H., Lourens, J. W., & Schoevers, R. A. (2015). Association of symptom network structure with the course of longitudinal depression. *JAMA Psychiatry, 72*(12), 1219–1226. https://doi.org/10.1001/jamapsychiatry.2015.2079

van Borkulo, C. D., Boschloo, L., Kossakowski, J. J., Tio, P., Schoevers, R. A., Borsboom, D., & Waldorp, L. J. (2023). Comparing network structures on three aspects: A permutation test. *Psychological Methods, 28*(6), 1273–1285.

Van Dam, A. (2021). A clinical perspective on burnout: Diagnosis, classification, and treatment of clinical burnout. *European Journal of Work and Organizational Psychology, 30*(5), 732–741. https://doi.org/10.1080/1359432X.2021.1948400

van de Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., Van Nes, E. H., Viechtbauer, W., Giltay, E. J., Aggen, S. H., Derom, C., Jacobs, N., Kendler, K. S., Van Der Maas, H. L. J., Neale, M. C., Peeters, F., Thiery, E., Zachar, P., & Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences* *of the United States of America*, *111*(1), 87–92. https://doi.org/10.1073/pnas.1312114110

Van Der Maas, H. L. J., Kan, K.-J., Marsman, M., & Stevenson, C. E. (2017). Network models for cognitive development and intelligence. *Journal of Intelligence*, *5*(2), Article 2. https://doi.org/10.3390/jintelligence5020016

Van Lissa, C. J., Hawk, S. T., Koot, H. M., Branje, S., & Meeus, W. H. J. (2017). The cost of empathy: Parent–adolescent conflict predicts emotion dysregulation for highly empathic youth. *Developmental Psychology*, *53*(9), 1722–1737. https://doi.org/10.1037/dev0000361

van Loo, H. M., & Romeijn, J.-W. (2019). What's in a model? Network models as tools instead of representations of what psychiatric disorders really are. *Behavioral and Brain Sciences; New York*, *42*. http://dx.doi.org/10.1017/S0140525X18001206

van Roekel, E., Heininga, V. E., Vrijen, C., Snippe, E., & Oldehinkel, A. J. (2019). Reciprocal associations between positive emotions and motivation in daily life: Network analyses in anhedonic individuals and healthy controls. *Emotion*, *19*(2), 292–300. http://dx.doi.org/10.1037/emo0000424

Vannucci, A., Finan, L., Ohannessian, C. M., Tennen, H., De Los Reyes, A., & Liu, S. (2019). Protective factors associated with daily affective reactivity and instability during adolescence. *Journal of Youth and Adolescence*, *48*(4), 771–787. https://doi.org/10.1007/s10964-018-0943-1

Varese, F., Haddock, G., & Palmier-Claus, J. (2019). Designing and conducting an experience sampling study. In J. Palmier-Claus, G. Haddock, & F. Varese (Eds.), *Experience Sampling in Mental Health Research* (pp. 8–17). Routledge.

Vasconcelos e Sa, D., Hartley, S., & Barrowclough, C. (2019). Experience sampling in the study of dyads and family dynamics. In J. Palmier-Claus, G. Haddock, & F. Varese (Eds.), *Experience Sampling in Mental Health Research* (pp. 81–96). Routledge.

Veenman, M., Janssen, L., Houtum, L. van, Wever, M., Verkuil, B., Epskamp, S., Fried, E. I., & Elzinga, B. M. (2022). *A network study of family affect systems in daily life*. PsyArXiv. https://doi.org/10.31234/osf.io/tyw47

Vishkin, A., & Tamir, M. (2023). Emotion norms are unique. *Affective Science*, *4*(3), 453–457. https://doi.org/10.1007/s42761-023-00188-z

von Bertalanffy, L. (1950). An outline of general system theory. *The British Journal for the Philosophy of Science*, *1*(2), 134–165.

von Bertalanffy, L. (1967). General theory of systems: Application to psychology. *Social Science Information*, *6*(6), 125–136. https://doi.org/10.1177/053901846700600610

Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures? *Journal of Applied Psychology*, *82*(2), 247.

Waszczuk, M. A., Zimmerman, M., Ruggero, C., Li, K., MacNamara, A., Weinberg, A., Hajcak, G., Watson, D., & Kotov, R. (2017). What do clinicians treat: Diagnoses or symptoms? The incremental validity of a symptom-based, dimensional characterization of emotional disorders in predicting medication prescription patterns. *Comprehensive Psychiatry, 79,* 80–88. https://doi.org/10.1016/j.comppsych.2017.04.004

Watzlawick, P., Bavelas, J. B., & Jackson, D. D. (1967). *Pragmatics of human communication: A study of interactional patterns, pathologies and paradoxes*. Norton.

Webster-Stratton, C. (1997). From parent training to community building. *Families in Society: The Journal of Contemporary Social Services, 78*(2), 156–171. https://doi.org/10.1606/1044-3894.755

Weiss, M. J. (2002). Hardiness and social support as predictors of stress in mothers of typical children, children with autism, and children with mental retardation. *Autism, 6*(1), 115–130. https://doi.org/10.1177/1362361302006001009

Weiss, N. H., Contractor, A. A., Raudales, A. M., Greene, T., & Short, N. A. (2020). Extending our understanding of the association between posttraumatic stress disorder and positive emotion dysregulation: A network analysis approach. *Journal of Anxiety Disorders, 71,* 102198. https://doi.org/10.1016/j.janxdis.2020.102198

Weston, S. J., Ritchie, S. J., Rohrer, J. M., & Przybylski, A. K. (2019). Recommendations for increasing the transparency of analysis of preexisting data sets. *Advances in Methods and Practices in Psychological Science, 2*(3), 214–227. https://doi.org/10.1177/2515245919848684

Wichers, M. (2014). The dynamic nature of depression: A new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine, 44*(1349–1360), 13.

Wichers, M., Groot, P. C., & Psychosystems, E. G. (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics, 85*(2), 114–116. https://doi.org/10.1159/000441458

Wichers, M., Schreuder, M. J., Goekoop, R., & Groen, R. N. (2019). Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine, 49*(3), 380–387. https://doi.org/10.1017/S0033291718002064

Wichers, M., Wigman, J. T. W., Bringmann, L. F., & de Jonge, P. (2017). Mental disorders as networks: Some cautionary reflections on a promising approach. *Social Psychiatry and Psychiatric Epidemiology, 52*(2), 143–145. https://doi.org/10.1007/s00127-016-1335-z

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* [Computer software]. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Wright, A. G. C., & Zimmermann, J. (2019). Applied ambulatory assessment: Integrating idiographic and nomothetic principles of measurement. *Psychological Assessment, 31*(12), 1467–1480. https://doi.org/10.1037/pas0000685

Yang, B., Chen, B.-B., Qu, Y., & Zhu, Y. (2021). Impacts of parental burnout on Chinese youth's mental health: The role of parents' autonomy support and emotion regulation. *Journal of Youth and Adolescence, 50*(8), 1679–1692. https://doi.org/10.1007/s10964-021-01450-y

Yang, Z., Algesheimer, R., & Tessone, C. J. (2016). A comparative analysis of community detection algorithms on artificial networks. *Scientific Reports, 6,* 30750. https://doi.org/10.1038/srep30750

Zacchilli, T. L., Hendrick, C., & Hendrick, S. S. (2009). The romantic partner conflict scale: A new scale to measure relationship conflict. *Journal of Social and Personal Relationships, 26*(8), 1073–1096. https://doi.org/10.1177/0265407509347936

Zarin, W., Veroniki, A. A., Nincic, V., Vafaei, A., Reynen, E., Motiwala, S. S., Antony, J., Sullivan, S. M., Rios, P., Daly, C., Ewusie, J., Petropoulou, M., Nikolakopoulou, A., Chaimani, A., Salanti, G., Straus, S. E., & Tricco, A. C. (2017). Characteristics and knowledge synthesis approach for 456 network meta-analyses: A scoping review. *BMC Medicine, 15*(1), 3. https://doi.org/10.1186/s12916-016-0764-6

Zimmermann, J., Woods, W. C., Ritter, S., Happel, M., Masuhr, O., Jaeger, U., Spitzer, C., & Wright, A. G. C. (2019). Integrating structure and dynamics in personality assessment: First steps toward the development and validation of a personality dynamics diary. *Psychological Assessment, 31*(4), 516–531. https://doi.org/10.1037/pas0000625

Appendices

Appendix A:
Kalkan et al. (2023)

Emotional Exhaustion and Feeling Fed Up as the Driving Forces of Parental Burnout and its Consequences on Children: Insights from A Network Approach

Parental burnout results from chronic stress in parenting, and it can be accompanied by harmful behaviors such as parental neglect and violence (Mikolajczak & Roskam, 2018). Network analysis examines psychological phenomena within a system of its constituents, and thus it is promising for understanding the distinct features of parental burnout and behaviors related to it. Recently, Blanchard et al. (2021) conducted the first network analysis of parental burnout and related harmful behaviors in the family context, but did so using an outdated measure and conceptualization of parental burnout. In the present study, in a sample of French-speaking parents (N = 3144, from five different previous studies), we aimed to investigate how each of the four features in the new conceptualization of parental burnout (i.e., emotional exhaustion, feeling fed up, emotional distance, and contrast with the previous parental self) interact with one another and with parental neglect and violence in a network system. In this preregistered reanalysis, we generated two network models commonly used with cross-sectional data: a Graphical Gaussian Model and a Directed Acyclic Graph. Our results point to emotional exhaustion and feeling fed up as key driving forces of the network structure, while emotional distance appears as a critical feature tying parental burnout with parental neglect and violence.

Reference:

Kalkan, R.B.*, **Blanchard, M.A.***, Mikolajczak, M., & Heeren, A. (2023). Emotional exhaustion and feeling fed up as the driving forces of parental burnout and its consequences on children: Insights from a network approach. *Current Psychology, 42,* 22278-22289. https://doi.org/10.1007/s12144-022-03311-8

**Co-first authors*

A.1 Introduction

Parental burnout is a clinically significant condition arising from a long-term imbalance between parental stressors and parental resources (Mikolajczak & Roskam, 2018; Roskam et al., 2021). Parents experiencing parental burnout feel emotionally exhausted, distant from their children, fed up with being a parent, and a sense of contrast with the parent they were beforehand. Its prevalence estimates are considerably high worldwide, rising to 7 to 8% in countries with the highest rates (e.g., Belgium, Poland, the USA; Roskam et al., 2021)

Investigating parental burnout is especially important as it can be damaging for all individuals in the family system. Most crucially, parental burnout strongly increases parental neglect and violence towards children. This effect appears causal, supported by cross-lagged results from two independent samples (both reported in Mikolajczak et al., 2019) and a recent experimental intervention study (Brianda et al., 2020). When parental burnout is treated, child neglect and violence decrease proportionally to the decrease in parental burnout symptoms. The effect of parental burnout on parental behaviors requires special attention, as neglect and violence have widespread, long-term outcomes on children's development (Delima & Vimpani, 2021; National Scientific Council on the Developing Child, 2010). However, most of these previous parental burnout studies relied on a unitary approach by tallying the different constitutive features of parental burnout into one sum-score, thus ignoring any possibility that these distinct features could interact differently with child neglect and violence.

Recently, however, an approach focusing specifically on the interactions between variables has arisen: a network approach to psychopathology that conceptualizes mental disorders as network systems of interacting symptoms (Borsboom, 2017). From this perspective, instead of investigating mental disorders as reflecting a single, unitary construct, one can examine the structure of, and associations between, the symptoms themselves. Studying a mental disorder's network structure can thus grant new insight into its symptom-to-symptom associations and topology. Though only recently pioneered by Borsboom and colleagues (Borsboom, 2017; Borsboom et al., 2011; Borsboom & Cramer, 2013), this approach has quickly grown prevalent in contemporary clinical psychological sciences. Many studies are now using this framework to investigate the interrelations between systems of symptoms and to speculate as to the clinical implications (for a general overview, see Blanchard & Heeren, 2022; for systematic reviews, see Contreras et al., 2019; Robinaugh et al., 2020)

Moreover, using a network approach to model psychopathology allows the adoption of network tools and graph theory concepts to understand mental disorders. One of these network concepts is node centrality, a tenet which posits that some nodes (here, symptoms) in the network are more important to the network structure than others (e.g., Blanchard & Heeren, 2022; McNally, 2016). Nodes are viewed as more central if they are especially connected to the rest of the network (i.e., sharing many strong connections with other nodes), because they can then influence the entire network. In network models, highly central nodes are thought to maintain the network structure since, once activated, they can quickly spread that activation to the entire system (e.g., Barabási & Albert, 1999; Borgatti, 2005). Within the network approach to psychopathology, central nodes have accordingly been theoretically (Borsboom & Cramer, 2013) and empirically (e.g., Elliott et al., 2019; Papini et al., 2019) linked to the prognosis of a disorder, although statisticians also advise caution when interpreting what exactly central nodes signify (Bringmann et al., 2019).

Analyzing network structures can also give insight into whether nodes form into communities (i.e., clusters of nodes that are more connected to each other than other nodes) and whether any specific nodes act as bridges to connect these communities. According to the network theory, these 'bridge' nodes could, for example, trigger a transition from a cluster of specific psychopathology features to related harmful behaviors (Jones et al., 2019). Identifying communities and the nodes that connect them thus helps researchers understand how specific symptoms influence the network system.

A network approach is especially suitable for investigating parental burnout and its relationship to parental neglect and violence (Blanchard & Heeren, 2020). A network approach views a psychological phenomenon as a dynamic system of interacting components (Borsboom & Cramer, 2013), and parental burnout does not exist in isolation but instead arises through the interactions of many variables relating to the entire family. Indeed, parental burnout has been conceptualized as an imbalance of parental stressors (e.g., conflict with the spouse, family disorganization, inconsistent childrearing practices) as compared to parental resources (e.g., support from the spouse, consistent childrearing practices; Mikolajczak & Roskam, 2018), highlighting that parental burnout is embedded within the family system. A network approach thus enables us to investigate the features of parental burnout and their associations with maladaptive behaviors in the family by visualizing these variables as a system of interacting nodes.

Recently, in a large cross-sectional network study ($n = 1551$), Blanchard et al. (2021) conducted the first network analysis to examine how the hallmark features of parental burnout interact with one another and with parental behaviors related to the partner and the child(ren). They used two distinct computational network approaches: a Graphical Gaussian Model (GGM), which is an undirected network in which nodes represent variables and edges signify the partial correlations between each pair of nodes (Epskamp & Fried, 2018), and a Bayesian network method to estimate a directed acyclic graph (DAG), which encodes the conditional independence relationships between the variables of interest and wherein each edge has an arrow tip on one end, indicating the direction of probabilistic dependence between them (Heeren et al., 2020; McNally, 2016; McNally et al., 2017; Scutari, 2010). In the GGM, emotional distance was highly central, and strongly connected with the other parental burnout features and with neglectful and violent behaviors toward children. Moreover, a community detection analysis (implemented via the Spinglass algorithm) identified emotional distance forming a subcommunity with parental neglect and violence. These observations all proposed emotional distance as a key variable that may maintain parental burnout within the family system, and that specifically relates to parental neglect and violence. And these observations were echoed in the DAG, wherein emotional distance emerged as a probabilistic driving force of the entire network structure.

Yet, although Blanchard et al. (2021) offered insight into the interactions of parental burnout features and related harmful behaviors, extensive research has recently suggested a new conceptualization of parental burnout, which was not considered in Blanchard et al. (2021), nor in other recent network approaches to parental burnout that did not focus on behaviors toward children (i.e., Le Vigouroux et al., 2022). The original conceptualization of parental burnout arose from the job burnout literature, and focused on exhaustion, distance, and inefficacy as a parent. However, a new theoretical approach has focused on building a conceptualization of parental burnout directly from the qualitative experiences of burned-out parents (Hubert & Aujoulat, 2018). This new framework pivoted from inefficacy as a parent (relating to the output of work) toward feeling fed up (relating more directly to parents' emotional experiences). Crucially, this new framework also adds a new feature to parental burnout: parents report feeling a sense of contrast with their previous parental selves; they previously felt happy and invested in parenting, and now experience shame and guilt that they have lost that investment in their parenting role (Roskam et al., 2018). Many studies have provided support for this new framework of parental burnout (Aunola et al., 2020; Furutani et al., 2020; Matias et al., 2020; Sodi et al., 2020). In line with this new conceptualization, a new measurement (i.e., the Parental Burnout Assessment, PBA) has been developed and validated in many languages and cultures (e.g., Arikan et al., 2020; Gannagé et al., 2020; Szczygieł et al., 2020). Thus, understanding this new conceptualization and its relationship with parental neglect and violence through the lens of a network approach might set the scene for novel conceptual and translational avenues.

Considering that this new conceptualization has significant conceptual and translational value, we aimed to examine in an exploratory and data-driven way how these four features of parental burnout (i.e., exhaustion, emotional distance, feeling fed up, and contrast with the previous parental self) relate to neglectful and violent behavior towards the child(ren). With the new conceptualization of parental burnout, which pulls from parents' qualitative experiences with burnout and which has been validated in many cultures, does emotional distance still play a central role in maintaining the network, as it did in Blanchard et al. (2021)? Does emotional distance still connect strongly with parental neglect and violence? How do contrast with the previous parental self (an additional feature) and feeling fed up (an updated feature focusing on parents' emotional experience vs. their concrete parenting ability), both of which refer more to affective aspects of parental burnout, interact with the other parental burnout features? Crucially, how do all of these features interact with parental neglect and violence? We use two different network methods to answer all of these questions, so as to build a better understanding of this new conceptualization of parental burnout and how it relates to child maltreatment.

A.2 Method

This study was an exploratory reanalysis of secondary data. We preregistered the analysis plan, methodology, and prior knowledge of the dataset following the guidelines of Weston et al.

(2019) at https://osf.io/vqmzb. Anonymized data and the R code can be accessed at https://osf.io/a89sd/. There were no deviations from the preregistration.

Participants

The data came from five different previous studies (Brianda et al., 2020) wherein four features of parental burnout (i.e., exhaustion, feeling fed up, emotional distance, and contrast with previous parental self) were assessed along with neglect and violence towards children (see Table S1). One of these studies is further detailed in Brianda et al. (2020), another in Bayot et al. (under review), and the remaining three are unpublished master's theses. The total dataset included 3144 French-speaking parents (86.7% women) who had at least one child living with them. The participants were recruited through internet advertisements, social media, online forums, and word of mouth. Further information about the data collection of the unpublished master's theses can be found in the online Supplementary Materials (https://osf.io/n74ah). Readers can refer to Brianda et al. (2020) and Bayot et al. (under review) for detailed data collection procedures for those specific studies.

Table S1. Data Collection: Data Sources, Sample Size, and Percentage of Female Participants

Study	Source	N	% Women	Mean Age *(SD)*
1	Data collected for master's thesis (2021), not published	1003	90.42	38.37 (8.64)
2	Data collected for master's thesis (2021), not published	1538	84	42.99 (8.31)
3	Brianda et al. (2020)	148	87.83	*
4	Data collected for master's thesis (2021), not published	381	86.87	38.90 (7.48)
5	Bayot et al. (under review)	74	87.84	*

Note. * = Not able to calculate mean ages or standard deviations; since these two studies were intervention studies with smaller samples, the authors decided to ensure the anonymity of participants by not collecting their exact age information, but instead age ranges. We therefore were unable to report the exact mean age and standard deviation for participants in these two studies, but interested readers can find the reported age ranges in the original manuscripts.

Measurements

Parental Burnout

The four features of parental burnout were assessed using the Parental Burnout Assessment (PBA; Roskam et al., 2018). The PBA has four subscales, each referring to one of the four features of parental burnout: emotional exhaustion (9 items; e.g. *I find it exhausting just thinking of everything I have to do for my children*), emotional distance towards child(ren) (3 items; e.g. *I do what I'm supposed to do for my child(ren), but nothing more*), feelings of being fed up (5 items; e.g. *I don't enjoy being with my children*), and contrast with the previous parental self (6 items; e.g. *I tell myself that I'm no longer the parent I used to be*). Participants rated items on a 7-point Likert frequency scale, from never (0) to always (6). In this dataset, reliability estimates of the global score and each subscale were good: the Cronbach's alpha for the global score was .97 (.95 for emotional exhaustion, .81 for emotional distance, .93 for feelings of being fed up, and .93 for contrast with the previous parental self). For our analyses, we calculated the mean[1] of each subscale separately as we were interested in the four features of parental burnout. This meant that the scores for each subscale ranged from 0 to 6, with higher scores indicating higher levels of parental burnout.

Parental Neglect

Parental neglect was measured with the Parental Neglect Scale (17 items; Mikolajczak, Brianda, et al., 2018), which measures physical, educational, and emotional neglect. All items assessing parental neglect were rated on an 8-point frequency scale, from never (0) to several times a day (7). Two of the studies in this dataset (studies #2 and #4) assessed parental neglect with a three-item short form of this questionnaire which involved one representative item from each category of neglect. The other three studies (studies #1, #3, and #5) assessed the full versions of the scale; we used only items corresponding to the short form from these studies (for more details, see the online supplementary materials at https://osf.io/n74ah). The Cronbach's alpha for the neglect scale of three items was .62.

Parental Violence

Parental violence was measured with the Parental Violence Scale (15 items; Mikolajczak, Brianda, et al., 2018) encompassing verbal, physical, and psychological violence. Items were rated on an 8-point frequency scale, never (0) to several times a day (7). As in parental neglect, two studies (studies #2 and #4) used the three-item questionnaire of parental violence and three studies (studies #1, #3, and #5) used the full version, although we only used the items corresponding to the short scale from these. The Cronbach's alpha value for the violence scale of three items was .57. For further information on how the short and long form questionnaires of parental neglect and violence were combined to form the nodes for "Neglect" and "Violence", see the supplementary materials.

Network Analysis

Data Preparation

Our analysis included six nodes, each representing a different variable: the four main features of parental burnout (i.e., emotional exhaustion, emotional distance, feeling fed-up, and contrast), as well as parental violence and parental neglect. We nonparanormally

[1] When administering the PBA, a global score is calculated by summing all items. However, we were interested in the subscale scores, and for the purposes of our network analyses we wanted all subscale scores to be on the scale, and so we calculated the mean of each subscale. In doing so, we followed our preregistration.

transformed all variables using the R package *huge* (Jiang et al., 2019). This ensured that the data met the assumption of multivariate normality required by the GGM, as well as ensured that each node had the same distribution, since different nodes were assessed with different scales. Because non-positive definite matrices can result in densely connected networks that may include false positives and unexpected negative edges (Epskamp & Fried, 2018), we verified that our correlation matrix was positive definite, thereby confirming that nodes were not a linear combination of one another. Finally, we investigated whether any of the nodes were redundant or highly overlapping in content: we sought whether any nodes had high intercollinearity (i.e., $r > .50$) and similar correlations with other variables (i.e., only 25% of correlations significantly differing for a specific pair of nodes). To this end, we followed previous research (e.g., Bernstein, Heeren, et al., 2019) and implemented the Hittner method for comparing dependent correlations (Hittner et al., 2003) via the *goldbricker* function from the *networktools* package (Jones, 2018). We did not identify any redundant nodes.

Graphical Gaussian Model

We used the Graphical Gaussian Model (GGM) to investigate how the features of parental burnout and parental violence and neglect interact in a network system. We regularized the GGM via the LASSO (Least Absolute Shrinkage and Selection Operator) algorithm. The LASSO algorithm computes regularized partial correlations between each pair of nodes and thus limits the number of spurious edges due to associations with other nodes. It also shrinks small trivial associations (edges) to zero, thereby minimizing false positive edges and leaving a network with only strong edges (Friedman, Hastie, & Tibshirani, 2014).

We used the *qgraph* package to implement LASSO regularization together with the Extended Bayesian Information Criterion (EBIC) model selection (Foygel & Drton, 2011), which together have been shown to result in networks with high specificity (Epskamp & Fried, 2018). This procedure generates 100 models with varying degrees of sparsity, and the model with the lowest EBIC value is selected as the final model, given a specific hyperparameter (γ). This hyperparameter γ determines the balance between including false-positive edges and excluding true edges. The hyperparameter γ is typically set between 0 and 0.5: a hyperparameter closer to 0.5 results in the EBIC favoring a simpler model with fewer edges, while setting the value for γ closer to 0 leads the EBIC to favor models with more edges. To ensure that the edges in our network are true, we set this hyperparameter (γ) to 0.5, as recommended by simulation (Epskamp & Fried, 2018). We estimated the stability of the edges using bootstrapped confidence intervals for each edge by generating 1,000 bootstrapped samples via the *bootnet* package (Epskamp & Fried, 2018).

To assess the importance of each node to the network, we computed the expected influence centrality of each node. Expected influence is the absolute sum of negative and positive edge weights connected to a particular node. Higher expected influence values signify greater centrality and importance in the network. We used the case-dropping subset bootstrap framework (Epskamp, Borsboom, et al., 2018) to evaluate the stability of centrality estimates. In this framework, the centrality estimates are assumed to be stable if they remain the same or highly similar even after dropping a certain number of cases or nodes (Costenbader & Valente, 2003). We measured the central stability coefficient (CS-coefficient) to quantify the centrality stability (Epskamp, Borsboom, et al., 2018).

Lastly, we investigated whether the features of parental burnout and parental neglect and violence cohere as a network system, or whether they cluster within subnetworks (e.g., one community with parental burnout features and another community with parental neglect and violence). Following previous research, we implemented the spinglass algorithm (Reichardt & Bornholdt, 2006), a community detection procedure suited for relatively small networks with positive and negative edge-weight values (e.g., Traag & Bruggeman, 2009). We used the *spinglass.community* function of *igraph* to run the algorithm (Csardi & Nepusz, 2006). We also investigated the presence of bridge nodes, or nodes that connect one community to another. We did so by computing the bridge expected influence of each node, that is, the absolute sum of positive and negative edge weights connecting a given node to all nodes in the other community or communities. These 'bridge' nodes are theorized to have clinical significance (Jones et al., 2019), since they could, for example, trigger an activation from one cluster of nodes (e.g., of parental features) to another cluster of nodes (e.g., of parental neglect and violence).

Directed Acyclic Graph

Finally, we computed the DAG by implementing the Bayesian hill-climbing algorithm via the *bnlearn* package (Scutari, 2010). This algorithm computes different combinations of edges (i.e., adding, removing, and reversing their direction) and finds the best fitting model according to a goodness-of-fit target score (i.e., the Bayesian Information Criterion, or BIC). This iterative procedure is randomly restarted with different possible edges and node pairs, random restarts (to avoid local maxima), and system perturbations (i.e., adding, removing, or reversing an edge), and eventually yields the best fitting network according to this iterative process. We set the parameters to 50 random restarts and 100 perturbations, following previous research (Bernstein, Heeren, et al., 2019; McNally, Heeren, et al., 2017).

To check the stability of the DAG, we bootstrapped 10,000 samples (with replacement) and averaged across the networks resulting from these bootstrapped samples. In a first step, represented in one figure, we determined the frequency of a particular edge's presence in the bootstrapped networks. Then we set a criterion according to the optimal cut-point method of Scutari and Nagajaran (2013) for edge retention. In the second step, represented in a second figure, we ensured that the directions of each remaining edge were reliable in the 10,000 bootstrapped networks. Following the recommendations of Scutari and Nagajaran (2013), if an edge pointed from one node to another in 51% of the bootstrapped networks at minimum, we included that directed edge in the final network. Of note, very thin edges (closer to 50%) were nearly equally likely to point in either direction, and so they might represent bidirectional processes (McNally, Heeren, et al., 2017).

A.3 Results

Descriptive statistics (i.e., mean, standard deviation, skewness, and kurtosis) of the four features of parental burnout and parental neglect and violence (before the nonparanormal transformation) can be found in the supplementary materials. Pearson product-moment correlations between parental burnout and parental neglect and violence variables after the variables were nonparanormally transformed were comparable in size to the previous studies by Blanchard et al. (2021), varying between .37 and .47 depending on the parental burnout subscale considered (see Figure S1 in the online supplementary materials: https://osf.io/n74ah).

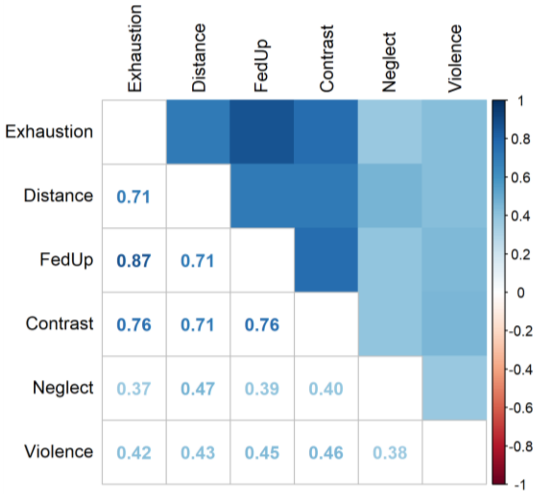


Figure S1. Pearson Product-moment Correlation Plot for All Variables after Nonparanormal Transformation

Note: Exhaustion = Emotional Exhaustion; FedUp = Feelings of being fed up; Distance = Emotional Distance; Contrast = Contrast with the previous parental self; Neglect = Parental neglect toward child(ren); Violence = Parental violence toward child(ren).

Graphical Gaussian Model

The resulting GGM network (regularized with graphical LASSO) is represented in Figure S2. Each node represents a variable, and the edges between nodes represent the regularized partial correlations between variables. There were no negative correlations, as all edges are colored in green. Several aspects of the network stand out. First, the largest edge weight connects exhaustion and feeling fed up, with an edge weight of $r = .62$. Second, all other edges are much smaller, with the second-largest edge weight, connecting contrast and emotional distance, having an edge weight of only $r = .27$. We conducted nonparametric bootstrapping and edge-weight difference tests to assess the stability and accuracy of each edge (see the online supplementary materials: https://osf.io/n74ah) and we concluded that the edge weights are stable: confidence intervals around each edge were small, and large edges significantly differed from smaller ones.

Figure S2. Graphical Gaussian Model structured with Graphical LASSO Algorithm

Note: The edge thickness represents the magnitude of regularized partial correlations, thicker edge implying higher partial correlation. Green edges denote positive regularized partial correlations. The different shades/colors represent different communities, as determined by the Spinglass algorithm. Exhaustion = Emotional Exhaustion; FedUp = Feelings of being fed up; Distance = Emotional Distance; Contrast = Contrast with the previous parental self; Neglect = Parental neglect toward child(ren); Violence = Parental violence toward child(ren)

Next, we computed expected influence centrality. As can be seen in Figure S3A, feeling fed up had the highest expected influence, followed by exhaustion, and then by emotional distance and contrast with the previous parental self.

To determine the subcommunities within the network, we used the Spinglass algorithm. The algorithm detected two subcommunities: a first community including the four parental burnout features (i.e., exhaustion, feeling fed up, emotional distance, and contrast with the previous parental self), and a second community encompassing parental neglect and parental violence toward the child(ren). Nodes with the highest bridge expected influence were parental violence and parental neglect as well as emotional distance (see Figure S3B). Both expected influence and bridge expected influence values were considerably stable, according to the case-dropping bootstrapping method (see online supplementary materials: https://osf.io/n74ah).

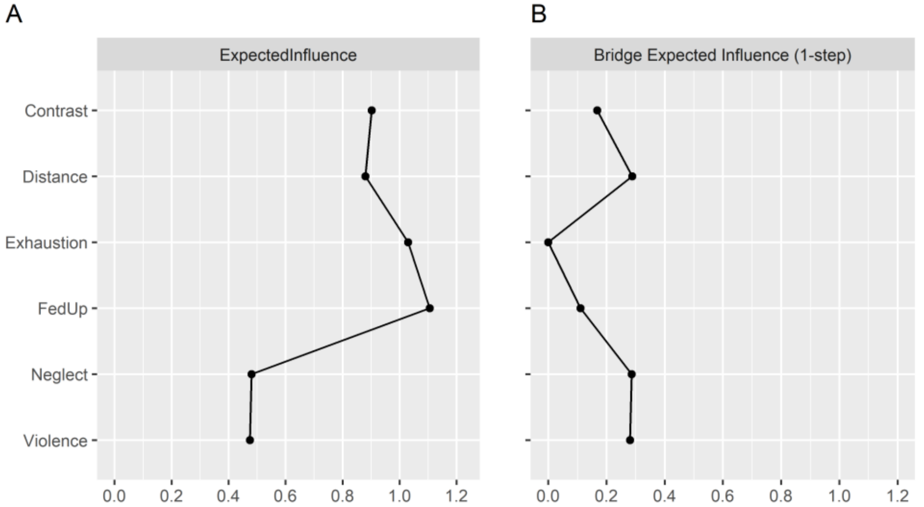


Figure S3. Expected Influence and Bridge Expected Influence of Each Node

Note: Exhaustion = Emotional Exhaustion; FedUp = Feelings of being fed up; Distance = Emotional Distance; Contrast = Contrast with the previous parental self; Neglect = Parental neglect toward child(ren); Violence = Parental violence toward child(ren)

Directed Acyclic Graph

The DAGs that result from the average of 10,000 bootstrapped networks can be seen in Figure S4A and S4B, each representing one step in generating the DAG (first, in Figure S4A, the presence of edges; second, in Figure S4B, their direction). Figure S4A thus reveals the importance of each edge to the network system, illustrating which edges are most likely to connect the different nodes. Thicker edges are more important for model fit (McNally, 2016), as edge thickness specifically represents the change in Bayesian Information Criterion (BIC, measuring a model's relative goodness-of-fit to the data) when that edge is taken out of the model. The edge connecting exhaustion and feeling fed up is the thickest, and thus the most important edge. The edges that connect exhaustion to contrast as well as feeling fed up to contrast also appear crucial for the network structure.

The edges in Figure S4B illustrate directional probabilities: a thicker edge represents an edge that was directed from node X to node Y in a greater proportion of the 10,000 bootstrapped networks. All of the edges in Figure S4B have similar low directional probabilities (see online supplementary materials: https://osf.io/n74ah). The two thickest edges are pointed from exhaustion to distance with a directional probability of .58 (i.e., in 5,800 of the 10,000 bootstrapped networks, the edge pointed in this direction) and from feeling fed up to violence.

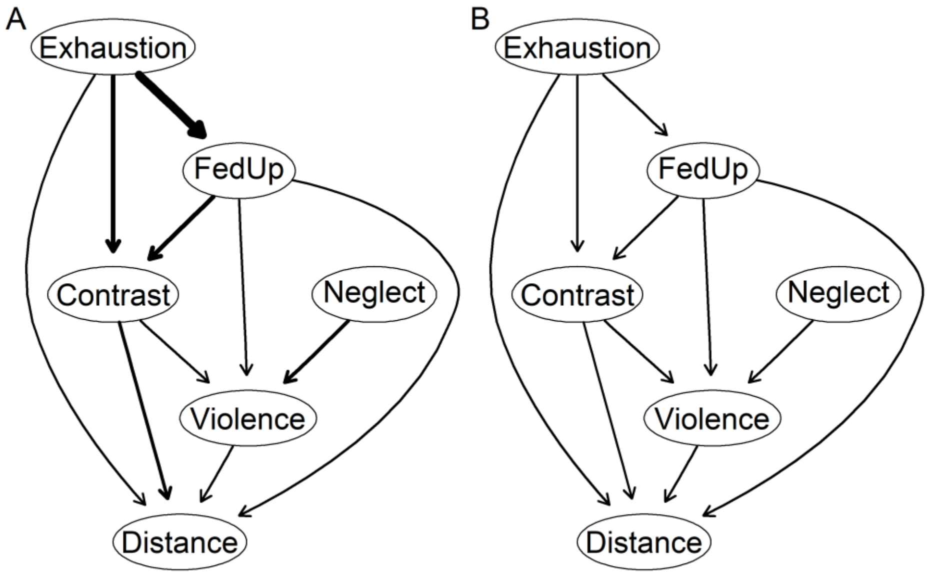


Figure S4. Directed Acyclic Graphs

Note: Panel A (presence of edges): the thickness of the edges denotes the importance of the edge for the model fit, specifically, high edge thickness implies bigger change in the Bayesian Information Criterion (a model's relative goodness-of-fit for the data) when the edge is excluded from the model. Panel B (direction of edges): demonstrates directional probabilities of the edges, a thicker edge implying that it was directed in the represented fashion in a greater proportion of the 10,000 bootstrapped networks. Exhaustion = Emotional Exhaustion; FedUp = Feelings of being fed up; Distance = Emotional Distance; Contrast = Contrast with the previous parental self; Neglect = Parental neglect toward child(ren); Violence = Parental violence toward child(ren)

As for the DAG's structure, exhaustion appears at the top of the network and directly influences the other features of parental burnout. Together with exhaustion, feelings of being fed up contributes to contrast with the previous parental self; and contrast directly influences emotional distance. Regarding parents' detrimental behaviors, violence is influenced by contrast, feeling fed up, and neglect. Interestingly, the R package *bnlearn* prompted a warning that there is likely an edge pointing from distance to parental neglect; however, it was not included in the graph as it would introduce a feedback loop.

Complementary Analyses

We wanted to see whether there were differences between mothers and fathers in terms of how parental burnout features interact with one another and with parental neglect and violence. To do so, we compared the GGM network structure of mothers (N = 2725) and fathers (N = 416), using the Network Comparison Test (van Borkulo et al., 2017). We found no significant differences in the connectivity, structure, or edge weights of these two networks.

In addition, since commentators have argued that conclusions about node centrality might be distorted when variables have drastically different variances (i.e., differential variability; Terluin et al., 2016), we computed the correlations between the standard deviation and the centrality estimates of each variable. Standard deviations were correlated with both expected influence, $r = .91$, $t(4) = 4.54$, $p = 0.01$, as well as with bridge expected influence, $r = -.89$, $t(4) = -5.06$, $p = 0.015$.

A.4 Discussion

In this study, we aimed to investigate how the features of parental burnout in its new conceptualization interact with each other and with parental neglect and violence through the lens of network analysis. To this end, we used undirected and directed network models: a GGM and a DAG.

Considering the results of both network models, exhaustion and feeling fed up and their relationship appear crucial for propagating activation in the network structure since 1) in both the GGM and the DAG, the strongest link is between these two variables, 2) feeling fed up has the strongest expected influence estimate (i.e., is most strongly connected to other nodes in the network), followed by exhaustion. The exhaustion component refers to being overwhelmed and fatigued by parenting-related tasks (e.g., *I find it exhausting just thinking of everything I have to do for my child(ren)*), while feeling fed up can be conceptualized as a reflection of high levels of exhaustion on parents' general affective experiences when parenting (e.g., *I can't take being a parent anymore*). Thus, it theoretically makes sense that being exhausted is often accompanied by feeling fed up.

In line with their theoretical distinction, parental burnout features and parents' harmful behaviors formed two distinct communities. The nodes that potentially connect these two communities, as indicated by high bridge expected influence values, are emotional distance and parental neglect and violence; these nodes, therefore, are likely to shift activation from one community (i.e., parental burnout) to another (i.e., parental maltreatment). This finding also aligns with the fact that emotional distance is the only node that has direct connections to all other nodes in the GGM.

Regarding the DAG representing the directional probabilities, the thickest edges pointed from exhaustion to distance and from feeling fed up to parental violence. However, the directional probabilities were all quite low (ranging between .59 to .51; see online supplementary materials: https://osf.io/n74ah), thus likely indicating bidirectionality—that is, the edges could be pointing from either node to the other. For instance, the directional probability of the edge directed from contrast to distance was .52, suggesting that emotional distance towards one's child(ren) can lead to a sense of contrast with the previous parental self as well as *vice versa*. Moreover, when computing the DAG, the R package *bnlearn* generated a message that an edge from emotional distance to parental neglect was also likely but not included in the graph, as it would introduce feedback loops to the DAG, which is constrained (as its name indicates) to be acyclic. This implies that parental neglect and emotional distancing might be reinforcing each other. These results together appear to be in line with family systems theory, which holds that an individual's cognitions and behaviors are embedded within a family context with all individuals influencing one another (Bavelas & Segal, 1982). If a parent feels emotionally distant from their children, this can feed into a vicious circle that involves neglect or violence towards children. Indeed, between the likely bidirectionality of most edges in the DAG and the likely presence of feedback loops, it appears that for this specific dataset, a DAG is not optimal to model the relationship between variables. Instead, to be able to model and disentangle the bidirectional relationships between variables, other methods, such as measuring the variables over time and visualizing their connections in a temporal network, are necessary next steps.

Interestingly, these findings do not overall align with the results of the previous network analysis (Blanchard et al., 2021). In Blanchard et al. (2021), the driving force of the network was emotional distance in both the GGM and DAG models; here, however, feeling fed up and exhaustion seem to be the most important nodes for driving the system. This could be due to the differences in the conceptualization and measurement of parental burnout, as Blanchard et al. (2021) used the previous conceptualization of parental burnout, the Parental Burnout Inventory (Roskam et al., 2017). In the new conceptualization of parental burnout, the PBA (which was developed through an inductive qualitative approach from the experiences of parents), "feeling fed up" emphasizes the emotional component as opposed to the functional aspects of parental burnout (i.e., inefficacy in parenting) in the previous conceptualization. Contrast with the previous parental self is also a new addition focusing on the affective experiences of parents – that is, feelings of guilt and shame that arise from parents' discrepancy between previous and current parental self. Our results, which highlight the key role of emotional exhaustion, support the Risks and Resources Theory of parental burnout which considers exhaustion as the starting point and a very important feature of parental burnout (Mikolajczak & Roskam, 2018). According to this framework, a chronic imbalance between parenting risks (e.g., lack of partner support, poor coping strategies) and parenting resources gives rise to emotional and physical exhaustion which, in turn, initializes a cascade of negative thoughts and feelings towards one's self as a parent and one's child(ren). Both the results from this article and the theoretical framework for parental burnout therefore pinpoint emotional exhaustion as key to the development of parental burnout. Moreover, our results draw attention to the strong connection between exhaustion and lack of joy and fulfillment in parenting (i.e., feeling fed up) as central to the burnout experience. When viewed from a network perspective, which assumes that turning off a central node via targeted treatments can downgrade activation throughout the network (McNally, 2016), a critical next step should be to examine whether efforts to lessen parents' emotional exhaustion, perhaps via identifying and eliminating parenting risks and increasing internal and external resources, could lead to improvements in the entire parenting network system.

Although emotional exhaustion was central to the network system in the present article, a divergence from the only previous network study of parental burnout (Blanchard et al., 2021), there are also several aspects that overlap with this previous network study. Of note, emotional distance still seems to be a key feature that connects parental burnout with parental neglect and violence. Indeed, emotional distance had the highest bridge expected influence estimate among parental burnout features. Nodes with high bridge centrality are thought to shift activation from one community (here: parental burnout features) to another (parental maltreatment), increasing the likelihood of the latter occurring (Jones et al., 2019). Better understanding the role emotional distance plays within the parental burnout system, and potentially targeting it in interventions, therefore seems critical for preventing or eliminating parental neglect and violence.

Despite its strengths, this study had several limitations. First, our sample consisted of only French-speaking parents, which restricts the generalizability of the results, especially as there are different parenting styles typical to different cultures (Bornstein, 2012). For example, parents in Euro-American countries have intensifying parenting norms (Roskam et al., 2021), with representations of a "good parent" involving parents who are always warm and positive (with less focus placed on parenting discipline, firmness, or rules; Dupont et al., 2022). These high societal standards for parenting can pressure parents to reach for difficult or unattainable goals, which can lead to parental exhaustion (Kawamoto et al., 2018; Sorkkila & Aunola, 2020). Other cultures that have different representations of a "good parent," or indeed different goals when raising a child, might exert less pressure on parents, and thereby less exhaustion (Bornstein, 2012). This could, for example, lead to a different relationship between parenting exhaustion and feeling fed up (currently the thickest edge in both networks). In addition, recent studies indicate that the rates of parental burnout vary significantly between cultures, particularly relating to individualism (Roskam et al., 2021). In addition, recent studies indicate that parenting and parental burnout vary significantly between cultures, particularly relating to individualism (Roskam et al., 2021). However, the effect-size of parental burnout on parental violence and neglect appears quite consistent across countries (see Mikolajczak et al., 2019 [Study 1, Study 2], and Szczygieł et al., 2020 for correlations in Belgium, the UK, and Poland, respectively). Nonetheless, we encourage future studies to investigate the interactions between parental burnout features and related harmful behaviors in different cultures. Second, as in most research on parental burnout, our sample mainly included mothers (86.7%). Although we did not find any differences between the network structure of mothers versus fathers, it would nonetheless be useful for future research to also investigate these questions with samples focusing on fathers. Third, the Cronbach's alphas of the short forms of the parental neglect and parental violence questionnaires were low (0.62 and 0.57, respectively). Thus, future research should consider using other measures with higher reliability for parental neglect and violence, or at least the longer forms of these questionnaires. Fourth, we found that both the expected influence and bridge expected influence of nodes were correlated with the standard deviations of nodes. This implies that these estimates of node centrality could be distorted by the variables having differential variability (although the variables do not have drastically different standard deviations; see online supplementary materials: https://osf.io/n74ah). Of note, since parental neglect and parental violence had high kurtosis values (see online supplementary materials: https://osf.io/n74ah), these two variables might be driving this differential variability issue. On the other hand, it is not surprising that these variables are highly skewed in a large, unselected sample. The interpretation of node centrality is generally under scrutiny with the psychological network literature (Bringmann et al., 2019) and as such the discussion of centrality for the GGM presented in this manuscript must therefore also be understood with caution.

Conclusion

The findings of this cross-sectional study pinpoint emotional exhaustion as the driving force of the network structure of parental burnout and parental neglect and violence, with feeling fed up also playing an influential role. This result aligns with the Risks and Resources theory of parental burnout (Mikolajczak & Roskam, 2018), which considers exhaustion as the feature that first arises as the risks outweigh the resources of the parent. Emotional distance, on the other hand, emerges as a critical feature of parental burnout which is closely linked to parental neglect and violence, thus requiring special attention for preventing and intervening on these detrimental behaviors. As our results suggest, parental burnout and related harmful behaviors seem to occur in a vicious circle with potential bidirectional associations between many variables. Thus, future research on the temporal dynamics of these relationships is going to be a critical next step.

Appendix B:
Supplementary Materials from Chapter 4

These supplementary materials can also be found at: https://osf.io/q56sn.

Further participant information

In the initial study (Mikolajczak, Brianda, et al., 2018), parents were informed about the questionnaire through social media, websites, schools, pediatricians, and word of mouth. To avoid (self-)selection bias, the questionnaire was presented as a study regarding "parental well-being and exhaustion," and not parental burnout. The initial study (i.e., Mikolajczak, Brianda, et al., 2018) concerned how individual parents perceived their own parenting, their relationship with their co-parent, and their behavior toward their children and was not designed as a dyadic study.

Sample size, power, and network analyses

Methods to estimate power a priori for a network analysis are largely unknown (especially when there is no knowledge regarding the likely structure of the network; Epskamp, Borsboom, et al., 2017). However, a general rule of thumb adopted from structural equation modeling is to include at least 10 participants per free parameter (Schreiber et al., 2006). Our eight-node network includes 36 parameters estimating (Epskamp, Kruis, et al., 2017), yielding a minimum sample size of 360, which our study far surpasses with a sample size of 1551. In addition, post-hoc estimations of stability are possible, including building bootstrapped confidence intervals around edge-weights and case-dropping bootstraps to assess the stability of centrality estimates. As we discuss below, these post-hoc stability assessments support that our edge-weights and centrality estimates are stable.

Use of nonparanormal transformed variables (vs. raw data) for correlation matrix

We present the correlation table of nonparanormal-transformed variables and not of the original raw data, since the raw data is non-normal, skewed and kurtotic, making it inappropriate for use with Pearson correlations. Since we present the correlation table to complement the Gaussian graphical model (based on linear relationships), we also did not want to use Spearman correlations with the raw data, since even though they are appropriate with non-normal data, they are based on monotonic relationships which are not relevant to network analyses. However, the Spearman correlation matrix of raw data is very similar to the one presented here.

Edge weights accuracy

To estimate the accuracy of the edges in the Graphical Gaussian Model (GGM), we used the R *bootnet* package (Epskamp, Borsboom, et al., 2017) to compute the bootstrapped confidence intervals for each edge. This was done using a non-parametric approach, with 1000 bootstrapped samples sampled with replacement. Figure S5 shows that the edges appear stable, with relatively small confidence intervals. Figure S6 portrays the results of the bootstrapped difference test, which suggests that the strongest edges are significantly different from weaker edges.

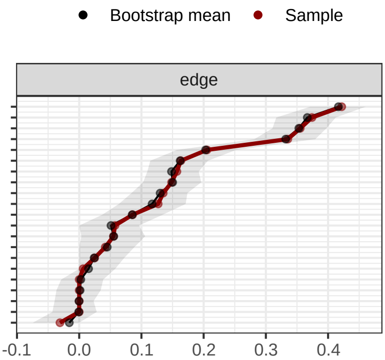


Figure S5. Bootstrapped Confidence Intervals of Estimated Weights for the Graphical Gaussian Model

Note. The red line indicates the values from the original sample, while the dark line indicates the mean bootstrapped values. The gray area indicates the 95% confidence intervals.

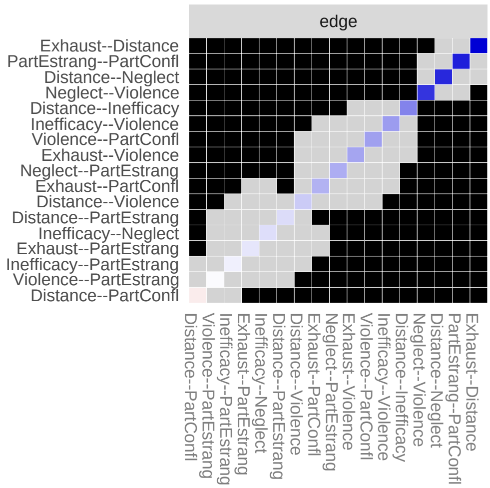


Figure S6. Bootstrapped Difference Tests (α = 0.05) Between Non-zero Edge Weights in the Estimated Network

Note. Black boxes denote edges that significantly differ from one another, while gray boxes indicate edges that do not. Colored boxes indicate the weight of that edge, with a darker color indicating a larger weight. Exhaustion = Emotional exhaustion; Distance = Emotional distance; Inefficacy = Loss of parental accomplishment and efficacy; Neglect = Neglectful behaviors toward children; Violence = Violent behaviors toward children; PartEstrang = Partner estrangement; PartConfl = Conflicts with partner.

Stability of centrality and bridge centrality metrics

To assess the stability of the centrality and bridge centrality metrics, we implemented a person-dropping subset bootstrap procedure (Costenbader & Valente, 2003). This involves computing centrality metrics for the original dataset and correlating these to centrality indices from bootstrapped subset datasets; if dropping a few cases leads to much lower correlations, then the centrality metrics are likely not stable (Epskamp, Borsboom, et al., 2017). Figure S7 reveals that the centrality metrics for the present study are stable.

We also quantified the stability of centrality metrics using the centrality stability correlation coefficient (CS-coefficient), which represents the maximum proportion of participants that can be dropped while maintaining 95% probability that the correlation between centrality metrics from the original dataset and the subset data set equal at least .70. CS-coefficients above .25, and preferably above .50, are recommended, based on simulations (Epskamp, Borsboom, et al., 2017). For the present study, the CS-coefficient for expected influence was 0.75, while the CS-coefficient for bridge expected influence was 0.67, meaning that we can be confident in our interpretations of these centrality metrics.

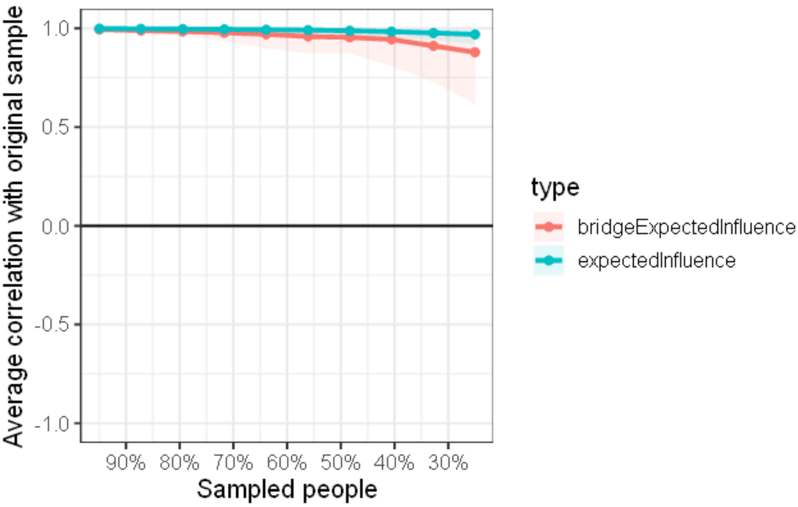


Figure S7. Average Correlations Between Centrality Indices (i.e., Expected Influence and Bridge Expected Influence) of the Original Network and Network with Persons Dropped

We also conducted a bootstrap difference test to investigate whether nodes significantly differ from one another between one another in terms of expected influence (see Figure S8) and bridge expected influence values (see Figure S9).

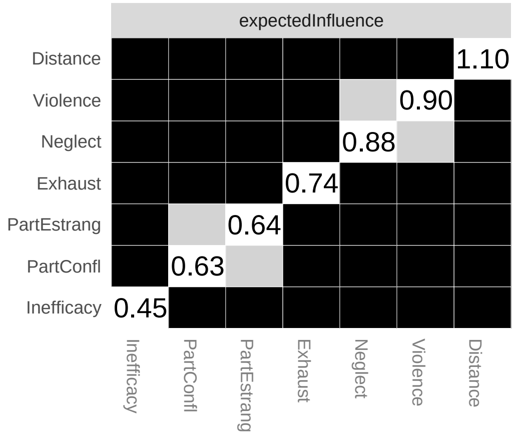


Figure S8. Bootstrapped Difference Tests Between the Expected Influence of Nodes

Note. Black boxes denote edges that significantly differ from one another, while gray boxes indicate edges that do not. Exhaustion = Emotional exhaustion; Distance = Emotional distance; Inefficacy = Loss of parental accomplishment and efficacy; Neglect = Neglectful behaviors toward children; Violence = Violent behaviors toward children; PartEstrang = Partner estrangement; PartConfl = Conflicts with partner.

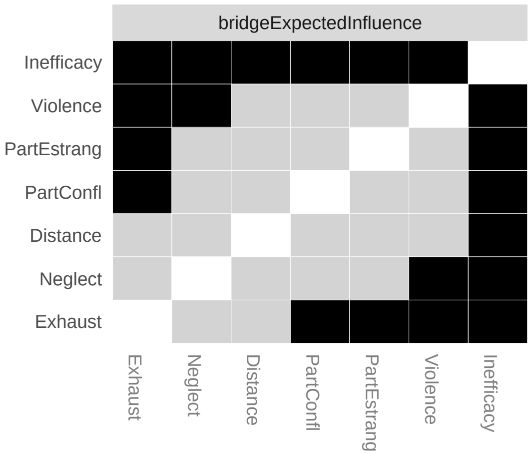


Figure S9. Bootstrapped Difference Tests Between the Bridge Expected Influence of Nodes

Note. Black boxes denote edges that significantly differ from one another, while gray boxes indicate edges that do not. Exhaustion = Emotional exhaustion; Distance = Emotional distance; Inefficacy = Loss of parental accomplishment and efficacy; Neglect = Neglectful behaviors toward children; Violence = Violent behaviors toward children; PartEstrang = Partner estrangement; PartConfl = Conflicts with partner.

Differential variability

Recent commentators have argued that differential variability—that is, the phenomenon that variables have drastically different variances—may distort conclusions about node centrality (e.g., Fried, 2016; Terluin et al., 2016). We therefore computed the correlations between the standard deviation and the centrality estimate for each variable, to test whether there is an association between a node's importance within the network and its variability. Both the two-tailed Pearson correlations between standard deviation and expected influence, $r(5) = .35$, $p = .45$, as well as with bridge expected influence, $r(5) = .13$, $p = .77$, were non-significant. This suggests that in the present study, differential variability does not influence the interpretation of expected influence and bridge expected influence.

Exact edge weight values for the DAGs in Figure 5

The exact values that determine edge thickness in Figure 8 (in Chapter 4) can be found in Table S2. This include change-in-BIC (i.e., how the model fit criterion changes if the specified edge is removed) for Figure 8A and directional probability (i.e. the frequency the edge was present in its current direction in the 10,000 bootstrapped networks) in Figure 8B.

Table S2. Edge Weight Values in the Directed Acyclic Graphs

Edge in Directed Acyclic Graph		Value determining edge thickness	
From	**To**	**BIC**	**Directional Probability**
Exhaustion	Violence	-29.57	0.81
Exhaustion	Partner Conflict	-24.10	0.70
Distance	Exhaustion	-338.29	0.53
Distance	Inefficacy	-140.54	0.69
Distance	Neglect	-167.37	0.74
Distance	Violence	-44.45	0.82
Inefficacy	Violence	-31.50	0.507
Neglect	Partner Estrangement	-47.85	0.67
Violence	Neglect	-120.07	0.54
Violence	Partner Conflict	-38.21	0.54
Partner Conflict	Partner Estrangement	-140.15	0.70

Note. BIC = change in Bayesian Information Criterion when that edge is removed from the network. BIC values determine edge thickness in Figure 8A (reflecting the importance of that edge to the network structure). For the BIC values, negative values correspond to decreases in the network score that would be caused by the edge's removal, whereas positive values correspond to increases in the network score that would be caused by the edge's removal. In other words, negative scores mean that model fit improves with the presence of that edge. Directional probability values determine edge thickness in Figure 8B (reflecting the frequency that edge was present in that direction in the 10,000 bootstrapped networks).

Comparing networks based on parents' gender and number of children

Theoretically, it is possible that network structure and strength differ based on the gender of the parent or family composition. We thus followed previous studies in the field and tested whether our results depended on the gender of the parent (father vs. mother) or the number of children (one, two, or three or more children). To do so, we used the Network Comparison Test (van Borkulo et al., 2023)—i.e., a permutation test assessing the difference between networks based on three invariance measures (i.e., network structure invariance, global strength invariance, edge invariance). For a more detailed explanation and example of how the network comparison test operates, see van Borkulo et al. (2023). When testing for differences between edges, we corrected for multiple comparisons within each network comparison test using Bonferroni corrections. When comparing the global strength and network structure having one, two or three and up children, we corrected for multiple comparisons using Bonferroni corrections. We only used network comparison tests to compare GGMs, as this test is suitable for GGMs but not for Bayesian networks.

There was no evidence for a difference in the global strength ($p = .77$), network structure ($p = .47$) or edges (see Table S2) of the networks of mothers compared to fathers. There was also no difference in global strength or network structure when comparing one vs. two children (strength: $p = .84$; structure: $p = .051$), two vs. three and up children (strength: $p = 1.0$; structure: $p = .23$), or one vs. three and up children (strength: $p = .33$; structure: $p = 1.0$). There was no evidence of a significant difference between edges of compared networks, with the exception of the node connecting neglect and partner estrangement for parents with one child vs. parents with two children (see Table S3). The network for parents with one child does not include an edge between neglect and partner estrangement, whereas the network for parents with two children, as well as the network for parents with three or more children, but do include this edge. One potential explanation for this could have to do with the length of time the parents have been together (with neglect more likely to be associated with partner estrangement the longer the parents had been together), but the original dataset did not include data on this variable so we cannot investigate further.

Table S3. Multiple-Comparison-Adjusted P-values of the Tests for the Presence of Every Edge in the Networks Compared Between Sub-samples of Mothers vs. Fathers and Being a Parent of One vs. Two vs. Three and More Children

Edges		*p*			
Node 1	**Node 2**	**Mothers vs Fathers**	**1 vs 2 children**	**2 vs 3+ children**	**1 vs 3+ children**
Exhaustion	Distance	1	1	1	1
Exhaustion	Inefficacy	1	1	1	1
Distance	Inefficacy	1	1	1	1
Exhaustion	Neglect	1	1	1	1
Distance	Neglect	1	1	1	1
Inefficacy	Neglect	1	0.82	0.58	1
Exhaustion	Violence	1	1	1	1
Distance	Violence	1	1	1	1
Inefficacy	Violence	1	1	0.60	1
Neglect	Violence	1	1	0.16	1
Exhaustion	Partner Estrangement	1	1	1	1
Distance	Partner Estrangement	1	1	1	1
Inefficacy	Partner Estrangement	1	1	1	1
Neglect	Partner Estrangement	1	0.02	1	1
Violence	Partner Estrangement	1	1	1	0.61
Exhaustion	Partner Conflict	1	1	1	1
Distance	Partner Conflict	1	1	1	1
Inefficacy	Partner Conflict	1	1	1	1
Neglect	Partner Conflict	1	1	1	1
Violence	Partner Conflict	0.68	1	1	1
Partner Estrangement	Partner Conflict	1	1	1	1

Appendix C:
Supplementary Materials from Chapter 5

These supplementary materials can also be found at: https://osf.io/gswq7.

C.1 Rationale for Technical Details of ESM

Single-Item Measure

ESM questionnaires need to be exceptionally brief to reduce the burden on participants as much as possible and limit their fatigue (Myin-Germeys et al., 2018; Robins et al., 2001; Varese et al., 2019). Single-item measures are one way to help limit the length of ESM questionnaires, while also reducing redundancy between items, which often leads to boredom and annoyance among participants (Robins et al., 2001; Wanous et al., 1997). Measuring a construct with multiple items is meant to help minimize measurement error, but it will only reduce random measurement error, while asking iterations of the same item will likely compound any systematic errors (Robins et al., 2001). Although it is more difficult to estimate the reliability of single-item measures, they have been recommended for use when studying a straightforward unidimensional construct, such as job satisfaction (Wanous et al., 1997), life satisfaction (Jovanović & Lazić, 2020), relationship satisfaction (Fülöp et al., 2022) subjective well-being, self-esteem, and relationships (Robins et al., 2001). They are particularly beneficial when situational constraints or practical considerations limit the use of longer scales, as with ESM measures (Fülöp et al., 2022; Gabriel et al., 2019; Wanous et al., 1997). This is especially true when asking busy parents to fill out ESM measures over a long period of time, and even more so for exhausted parents experiencing parental burnout.

Daily Interval

We chose to assess ESM items daily, as one previous study found that parenting exhaustion fluctuated on a daily basis (Gillis & Roskam, 2020). Moreover, a daily time frame seems correct to assess fluctuations in parenting behavior and family context. Parenting-related variables likely follow a daily routine, and assessing behaviors more commonly would be difficult for parents (especially as many parents, due to schooling or work obligations, do not see their children throughout the day, but only in the morning and evening). On the other hand, a more infrequent sampling period (e.g., on a two-day or weekly basis) would likely miss essential fluctuations. Moreover, a daily sampling frame has been reported as highly intuitive and easy to implement for participants (Zimmermann et al., 2019).

Continuous Scale

Following previous research (e.g., Aalbers et al., 2019), and to allow participant responses to have sufficient variability, we assessed each ESM item with a slider scale ranging from 0 (not at all) to 100 (absolutely).

C.2 ESM question development

Examples of item changes

First, when we asked parents for their comments on item phrasing during the development of the items, parents gave useful feedback. For example, to assess children's behavior in a manner that was generally applicable to any child, one item read: "Today, my child(ren) were noisy." A parent noted that children could engage in behaviors that were difficult for the parent but not necessarily noisy, or that children could be mute or quiet and not often make noise. We adapted this item to read: "Today, my children were wound up." After further discussion, we later further changed this item to be more general: "My children were agitated, noisy, or needed a lot of attention." Based on feedback from parental burnout experts, we then changed it to be more applicable to all children (e.g., teenagers can be withdrawn and quiet in a way that is difficult for parents): "My children were difficult to manage."

Second, based on parents' feedback during the pilot, we made a few small wording changes to items. For example, the item measuring time spent with kids was originally phrased: "Around how many hours did you spend with your children today." Parents asked whether any time spent together counted, or just time spent doing activities, etc. We altered the item to be clearer, specifying the hours the parent was close by their child, not counting sleeping hours: this would be relatively easy to estimate (unlike hours spent doing activities together), and we would also be able to know if, for example, parents and children were both at home all day or not.

Thirdly, the English translations of a few phrases used more precise language than their original French formulations, and so we adapted the French items to reflect the clearer phrasing. For example, we used the word "entourage"—referring to those around or close to you— in the French 'Social Support' item, but in English, the closest commonly used translation of that term is "family and friends." Since we wanted to ensure that parents did not include support they received from hired help (e.g., housekeepers, nannies, paid tutors) in their response, and since parents had commented that they were unsure of the exact definition of 'entourage' in the second phase of the data collection, we decided to use the phrase "family and friends" in both English and French versions of the Social Support item. We also considered that if parents received help from neighbors or others around them in taking care of their children, then these people would likely fall within the category "family and friends."

These are a few small examples of changes we made to items during the development process. To see the full list of changes, see Table online: https://osf.io/46z8x/.

Additional Information about the Back-translation Process

Two bilingual translators (who were both native American English speakers) independently translated from French to English and then discussed the differences to come to an agreed-upon English version (with a focus on clarity, unambiguous language and terminology used by the targeted population; Myin-Germeys et al., 2018; Varese et al., 2019). A third independent translator (native Belgian French speaker) translated the English items back to French. The original and back-translated French items were compared, and were found to have no substantive differences in content or meaning.

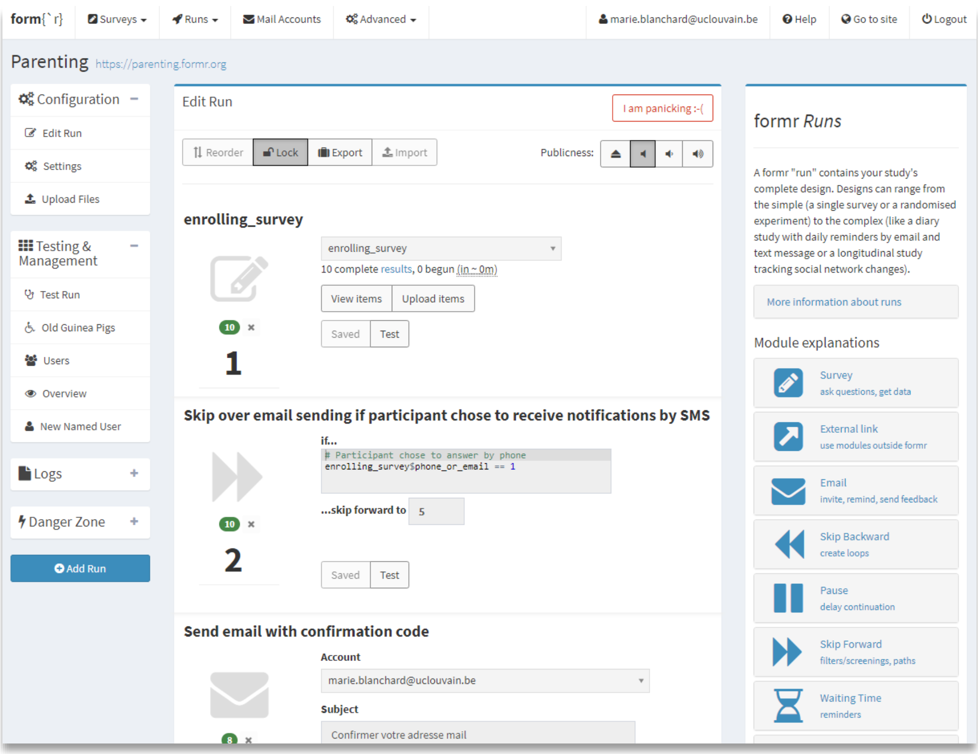
C.3 Software: formr

The researcher can schedule notifications containing the participant's link to the questionnaire via email within formr, or via text messages by connecting to an outside SMS gateway or server (formr provides instructions with an example SMS provider here: https://github.com/rubenarslan/formr.org/wiki/How-to-send-text-messages-(SMS). The formr software views a study as a 'run' containing specific modules with different functions (sending notifications, administering a questionnaire, skipping forward or backward to a different module; for an overview on how formr operates, see https://formr.org/documentation). An example of how the (beginning of the) run page looks from a formr administrator account can be seen in Figure S10. Questionnaires (or "surveys" in formr terminology) are embedded within runs, and they are created and managed through online spreadsheets, rendering them easy to share and fostering collaborative work. Surveys and runs in formr are also extensively customizable through R code. Although there is a small learning curve to understanding the logic of formr (and a larger learning curve when learning to incorporate R code), we found that its open-access format and customizability render it worth the effort. We share formr templates for our ESM studies (French-language version: https://osf.io/c492m/, and English-language version: https://osf.io/p95zs/). These are slightly cleaned up versions (code more straightforward, etc.); for interested readers, the template for the exact version of the run used in the second phase of data collection can be found here: https://osf.io/mjb29/.

C.4 Additional Information about Data Collections & Samples

There was no overlap between any of the groups of participants (i.e., the group that only provided feedback on the phrasing and content of items and the four groups that participated in data collection). There were two sets of couples within the seven participants that provided phrasing feedback, and one couple that participated in the French-language two-week data collection. However, since our focus was not on statistical analyses (other than to examine the variance and ICCs of items), but instead on participants' experience and feedback of the ESM study, we do not expect this to affect any of our results or conclusions.

Figure S10. Example of formr study ('run') page (Administrator view)



Note. This is the editing screen for a formr run. Further information can be found on formr's documentation page: https://formr.org/documentation

Pilot sample (n=5, 23 questions, 7 days)

- This was a first phase of data collection to assess which items yielded insufficient variation in participant responses, as well as how parents felt answering the items on a daily basis and whether they judged the items as clear, unambiguous, and easy to answer.
- Five parents (three women) were recruited from among the authors' acquaintances to participate; these parents had a mean age of 42.2 (SD = 5.64) and a mean of 1.8 children (ages ranging from 1 year to 18 years old).
- The pilot included 23 questions (3-4 items for each category) because we were testing multiple questions per category to narrow them down and see which parents thought were good daily questions and which weren't (see Table S1 for the full list of questions: https://osf.io/46z8x/).
- Administered via Qualtrics, with parents receiving the notifications at 6 PM every evening

- Did not receive any compensation

French-language two-week data collection (n=9)

- These parents had children ranging from 2 to 19, and 11% of parents had at least one child under the age of 5 living with them, 56% had at least one child under the age of 10
- We set the ESM survey to expire after five hours; this was done in the settings of that specific survey in formr (this was also the case for the later two data collections, which were also conducted through formr)
- Did not receive any compensation

English-language two-week data collection (n = 23)

- All participants live in Belgium but their nationality varies (Belgian, American, Belgian-American, Canadian, Mexican, etc)
- 18 of the parents work full-time, 4 work part-time, 2 are on parental leave
- Everyone has at least a B2 level of English fluency (1 person had a B2 level, everyone else has a native or fluent level)
- Received compensation (15 euros)

French-language eight-week data collection (n = 50): Briefing

- With each parent, a researcher (MAB) had a video call briefing session to introduce the study, explain how participants would be notified, show how the surveys would appear on their phone, etc. Most importantly, the researcher when over the ESM item themselves one by one, emphasizing that all responses are subjective and pertain specifically to the participant's daily life; if the parent had difficulties answering an item, the researcher would go over their current day (or yesterday, if early int the morning) together so that the parent could have an in-depth understanding of each item. To this end, we created a *formr* survey with just the ESM questions, to introduce participants during the introductory briefing session to how the software operates, as well as for interested readers to take a look: https://exempleparent.formr.org/.
- Received compensation (10 euros if they participated in the study, and an additional 25 euros if they completed at least 80% of ESM prompts)

C.5 French-language ESM items

Table S4. French-language items

Category	Item (version française)
Emotional Exhaustion	Je me suis senti épuisé lorsque je me suis occupé de mes enfants.
Emotional Distance	Je me suis senti proche de mes enfants dans les bonnes comme les mauvaises circonstances.
Feeling Fed Up	J'ai été dépassé par la gestion de mes enfants.
Partner Support	J'ai reçu de l'aide de mon conjoint dans la gestion de mes enfants.
Partner Conflict	Il y a eu des mésententes, des tensions, des disputes dans mon couple.
Kids' Behavior	Mes enfants ont été difficiles à gérer.
Behavior towards Kids (+)	J'ai partagé des moments positifs avec mes enfants.
Behavior towards Kids (-)	J'ai été en colère envers mes enfants.
Resources	J'ai manqué de moyens (par exemple, temps, énergie, ressources matérielles) pour m'occuper de mes enfants.
Social Support	J'ai reçu de l'aide de mes amis et de ma famille (autre que mon conjoint) pour m'occuper de mes enfants.
Time with Kids	Environ combien d'heures avez-vous passées (en dehors des heures du sommeil) à proximité avec vos enfants aujourd'hui ?

C.6 More information on Qualitative Participant Feedback

Initial Piloting during Item Development

Several parents found it difficult to answer the questionnaire between 6 PM and 8 PM, since this was the busiest time of their day with their children. Parents reported that the questionnaire was not at all a burden, especially if they would have been able to respond when they were getting ready for bed instead of earlier in the evening. They found no difficulties balancing their family or work life with participating in the study, and they appreciated the notifications to send out the questionnaires. Notifications for this pilot were sent out by email, and half the parents preferred email, while the other half said they would have preferred text message notifications. Parents reported that they appreciated participating in the study, especially as it allowed them to evaluate their emotional state and their day with the children each evening; one parent specifically wrote that the items allowed her to give more attention to the quality of time they spent with their children. Parents found it straightforward to mentally relive their day in order to respond to the questions. Multiple parents reported that the questions were redundant and repetitive (this data collection contained 3-4 items per category).

French-language Two-week Data Collection

Three parents mentioned that some of the questions might not be optimally suited for parents with older children who were more independent, since the parents would do less active caretaking. Parents did not report any emotional difficulties, annoyance, or burden in answering the daily questions; the one exception was a parent with a son in university, who felt the questionnaire was not applicable to her life. Two parents asked for definitions for specific terms (e.g., “difficult to manage” or “getting angry”), as they said these terms could vastly differ from family to family (we later included explanations in the briefing; either a recorded video briefing as in the English-language two-week data collection, or a face-to-face video call briefing in the French-language eight-week data collection). Some parents suggested slight changes to the phrasing of a few items to improve their clarity; we incorporated some of these to arrive at the final iteration of the items as reported in Table 5 (for details on changes, see https://osf.io/46z8x/).

English-language Two-week Data Collection

One parent mentioned they had older children and so the question about outside help was less relevant. Another mentioned the questions were quite focused on negative experiences.

C.7 Mean daily response and standard deviation for each sample (figures)

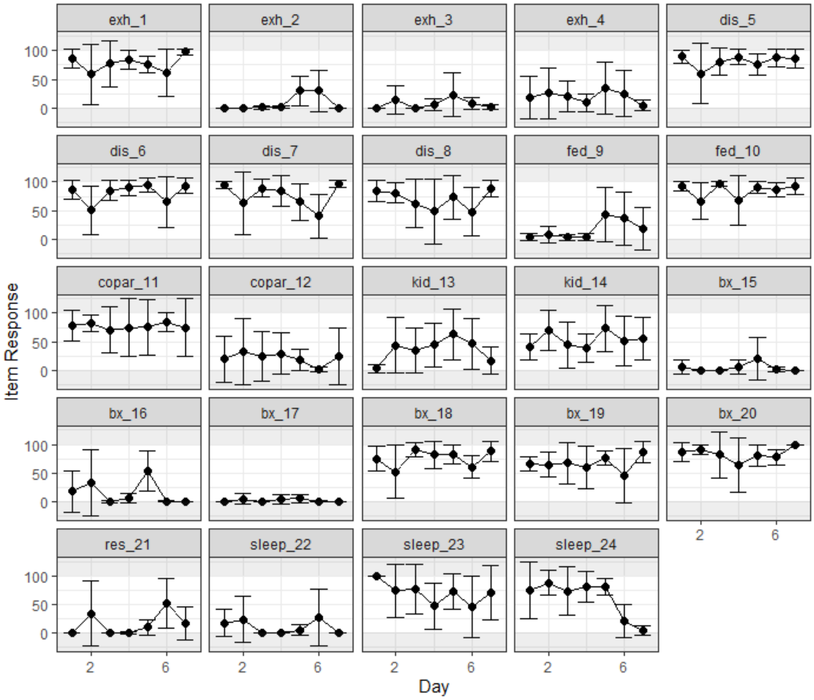


Figure S11. Mean Daily Responses for Each ESM Item (Pilot, with 23 questions) with SD Error Bars

Note. The ESM items were assessed with a slider from 0 to 100; the light gray space in each figure represents numbers not included in the responses but included to show the full range of each response's standard deviation. The items assessed can be found here: https://osf.io/46z8x/. ESM = experience sampling methodology; exh = emotional exhaustion; dis = distance; Fed = feeling fed up; copar = partner relationship; kid = children's behavior; bx = behavior toward children; res = resources.

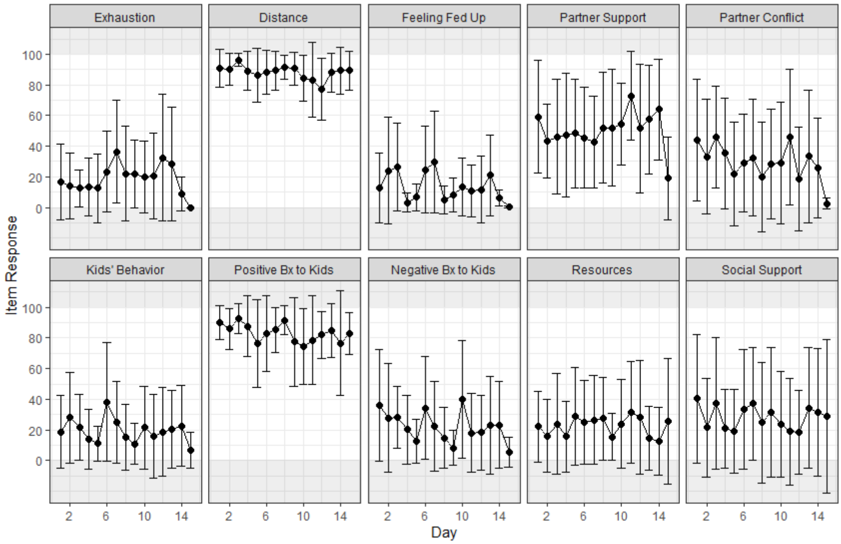


Figure S12. Mean Daily Responses for Each ESM Item (French-language Two-week Data Collection) with SD Error Bars

Note. The ESM items were assessed with a slider from 0 to 100; the light gray space in each figure represents numbers not included in the responses but included to show the full range of each response's standard deviation. The X-axis includes 15 days because two participants started the study on Day 2, so finished on Day 15. ESM = experience sampling methodology; Bx = Behavior.

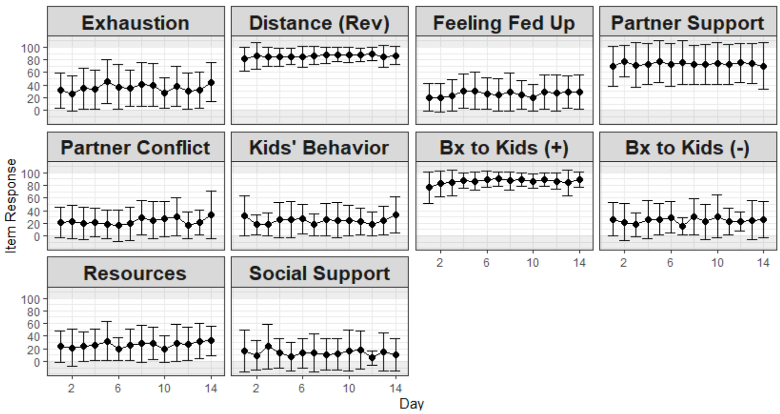


Figure S13. Mean Daily Responses for Each ESM Item (English-language Two-week Data Collection) with SD Error Bars

Note. The ESM items were assessed with a slider from 0 to 100; the light gray space in each figure represents numbers not included in the responses but included to show the full range of each response's standard deviation. ESM = experience sampling methodology; Bx = Behavior.

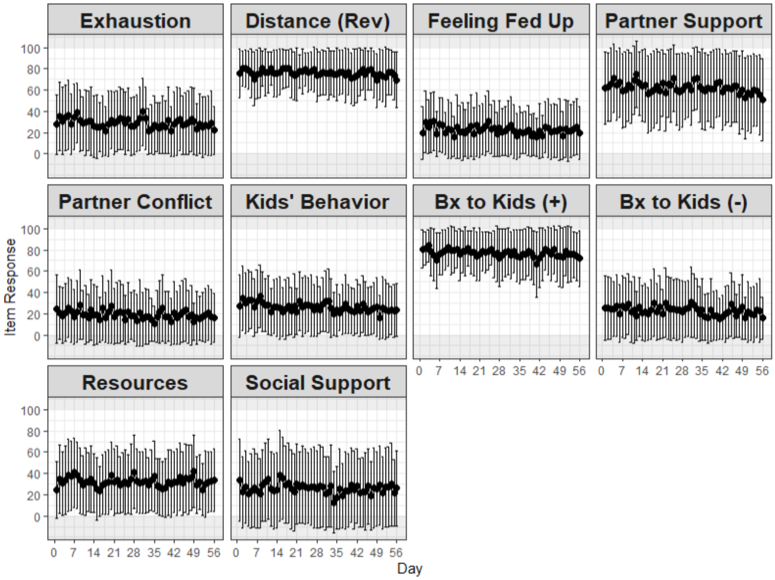


Figure S14. Mean Daily Responses for Each ESM Item (French-language Eight-week Data Collection) with SD Error Bars

Note. The ESM items were assessed with a slider from 0 to 100; the light gray space in each figure represents numbers not included in the responses but included to show the full range of each response's standard deviation. ESM = experience sampling methodology; Bx = Behavior.

C.8 Sensitivity analyses for ICC values (removing participants with low variation in English-language sample)

ESM item	ICC (original)	ICC (no SD = 0)[a]	ICC (SD < 4.19)[b]
Emotional Exhaustion	0.51	0.48	0.35
Emotional Distance	0.62	0.59	0.44
Feeling Fed Up	0.40	0.41	0.32
Partner Support	0.57	0.55	0.51
Partner Conflict	0.31	0.32	0.29
Kids' Behavior	0.24	0.25	0.15
Behavior toward Kids (+)	0.36	0.37	0.31
Behavior toward Kids (-)	0.25	0.28	0.22
Resources	0.33	0.32	0.16
Social Support	0.46	0.51	0.54
Time Spent with Kids	0.39	0.37	0.37

Note. When removing participants for having no or low variation in their responses, we did not include low variation for the following items, since these might objectively not change over a two-week period: partner support (e.g., if partner is gone or does not help at all for two weeks), social support (e.g., parents might not have anyone other than themselves and/or their partner helping with childcare) and time with kids.

[a] When we removed participants who had responses over the 14 days with a SD = 0 (e.g., they answered exactly the same every day), this affected two participants only (from Belgium and Canada, both with high levels of English).

[b] When we removed participants who had responses with a SD in the lowest 5% (e.g., they showed very little variance in their responses, this specifically affected Belgian, Belgian-American, American and Canadian participants. These participants had a high level of English (one participant had a B-2 level, all other were fluent or native).

Appendix D:
Supplementary Materials from Chapter 6

These supplementary materials can also be found at: https://osf.io/tua9f.

D.1 Article Selection

Our inclusion criteria did lead to some articles that were ambiguous on whether or not they should be included in the review. For example, here are some examples of articles we did not include, although we did debate their inclusion:

- Savelieva et al. (2021) and Komulainen et al. (2020) specifically modeled change scores/improvement,
- Spanakis and colleagues (2016) did not publish their temporal network in a peer-reviewed journal, and
- Wichers (2014) focused on illustrating the temporal network approach, without the goal of providing an empirical contribution to the literature.

D.2 Data Charting Process

Following the PRISMA guidelines (Tricco et al., 2018), we report specific data items collected for our scoping that involve interpretation and how we defined these items. Here is the full list of questions used to extract data: https://osf.io/mpj89/.

Items that involved interpretation and further specification:

- Q15 Sample size rationale: Do the authors provide a sample size rationale specifically for the temporal network analyses (not for analyses for a larger project/etc)
- Q59 Other networks: this question specifically involves if authors estimate a different type of network, not if they divided their sample into groups/etc (which is assessed by Q65)
- Q75 Preregistration: has to specifically be preregistration for the temporal network analysis (not for a previous project)
- Q76 Code Available: has to specifically include the temporal network estimation code

Iterative Charting Process:

We also added items during the scoping review (listed in our updated preregistration in blue, https://osf.io/hz9ev/):

- Q14a. If the authors compute multiple versions of the same network with different samples, what is the sample size for each network (i.e., if they compute one network for clinical participants and another for control participants)?
    - *Justification: Many authors computed multiple networks, and we wanted to know the sample size per network, not the sample size overall*
- Q38a. If the authors compute multiple networks from the same dataset and change the number of nodes (not to assess sensitivity but for theoretical reasons), specify how the authors changed the number of nodes (i.e., removed or added one node, etc)
    - *Justification: Some authors computed multiple networks but the networks had different numbers of nodes (took one away, added one, etc). Since we were also recording number of nodes, we needed to report if this number changed.*
- Q51a. How did they estimate networks? Check all that apply:
    - Sequential univariate estimation of a multilevel VAR model
    - Multivariate estimation of a multilevel VAR model
    - GIMME model
    - Other type of VAR model (specify)
    - Other: _________
    - *Justification: We wanted a standardized question about the method authors used to estimate networks so that there would be consistency in reviewer responses (e.g., not just VAR models, but what type of VAR model)*
- Q65a. Do the authors perform any other comparison (including visual) between networks (i.e., to compare subgroups, or different networks from the same dataset with the inclusion/removal of one node)? Check all that apply: Yes / No / Other: _________
    - *Justification: We wanted a question to categorically address whether there was any type of comparison between networks, include visual comparisons*
- Q72b. What do the authors do regarding missing data?
    - *Justification: Since a number of authors reported this, it was worth reporting in its own column (for just under Q72 Additional analyses)*

Organizing Charted Data:

We also organized the final spreadsheet, such as categorizing data and or combining columns. Here are further details:

- Combined questions 9 (larger project) and 10 (reanalysis) into one question (answered 'yes' if the project is part of either a larger project or reanalysis)

  - *Justification: Often no distinction between the two*
- Added column Q51b (Model specifications), to put all information about model specifications
  - *Justification: reviewers reported this information but in various columns; this reunited the information in one central place*
- Categorized information from free response columns (specifically, Q64b and 66b, to categorize information about comparing centrality, topology, etc)
  - *Justification: since there were so many comparisons, it was useful to categorize the methods authors used to conduct them*

D.3 Rater Agreement for Data Extraction

The unweighted Cronbach's kappa values for all categorical variables (independent from previous responses) are in Table S5.

Table S5. Interrater reliability: Cronbach's Kappa Values for All Categorical Variables

Question	Kappa	p
Q39.VariableType	0.082	0.52
Q10.Reanalysis	0.291	0.033
Q65a.ComparisonOther	0.3	0.047
Q24.DataCollectionDesign	0.318	0
Q40.VariableDefined	0.34	0
Q32.Exclusions	0.367	0.006
Q73.Robustness_Temporal	0.38	0.012
Q41.RedundancyCheck	0.404	0.005
Q63.ComparisonCentrality	0.484	0.001
Q15.SampleRationale	0.534	0
Q65.ComparisonTopology	0.551	0
Q75.Preregistration	0.557	0
Q51b.NetworkType_BS	0.575	0
Q73.Robustness_Contemp	0.645	0
Q12.Replication	0.656	0
Q13.Hypothesis	0.658	0
Q9.LargerProject	0.658	0
Q30.Compensation	0.672	0
Q77.SharedData	0.719	0
Q43.NormalityCheck	0.739	0
Q46.StationarityCheck	0.766	0
Q11.LinktoProject	0.769	0
Q70.AndOrRule	0.775	0
Q20.SampleType	0.78	0

Q52.CentralityTemporal	0.796	0
Q51b.NetworkType_Contemp	0.83	0
Q57.CentralityBetweenSubj	0.941	0
Q76.SharedCode	0.946	0
Q55.CentralityContemp	0.947	0
Q51.BetweenSubjNetwork	0.951	0
Q51.ContempNetwork	0.953	0
Q51.TemporalNetwork	1	NA
Q54.Lag	1	NA
Q7.Published	1	NA
Q73.Robustness_BetweenSubj	1	0

Some low kappa values are due to lack of clarity in how articles reported the relevant information. For example, Question 34 in the data extraction form (link here: https://osf.io/mpj89/), 'variable definition' (k = .34), involved how variables were combined into nodes; this was often not directly reported in the text but could be pieced together based on the number of items reported to assess a specific variable, and the number of nodes reported to assess that same variable. Another example: the variable 'hypothesis' (k = .66; Question 13) assessed whether articles included a clearly stated hypothesis; a number of articles stated that their analyses were fully exploratory but also included clear and specific predictions. This led to disagreement between the two raters, and upon discussion with the third reviewer, we decided to rate as 'yes' any article that stated clear and specific predictions, even if they also said their analyses were exploratory. A third example: the variable with the second smallest kappa (k = .29) was whether articles reanalyzed previously collected data (Question 10). Many articles seemed to analyze data that was not collected for the purpose of a temporal network analysis (either specified in the article itself or apparent from details in their methods procedures, such as if they included very strict inclusion criteria that were not relevant to the temporal network analysis) but did not clearly specify whether they were reanalyzing other data, part of a larger project, or something else.

Other low kappa values reflected insufficient specificity in the data extraction form, despite extensive discussion and testing of this form beforehand. For example, "Variable type" (k =.08; Question 39) assessed whether the items assessed in the temporal network articles were continuous, ordinal, binary, etc. The two reviewers had different definitions of ordinal and continuous (one based on the number of points in the scale; the other based on whether the type of scale was defined as Likert or not), which led to many inconsistencies for Likert scales (the most common assessment method among the 43 included articles). Upon discussion with the third reviewer, the reviewers changed an answer option to Ordinal (Likert scale), which solved all disagreements.

D.4 Data Collection

The assessment schedule of each article is reported in Table S6.

Table S6. Assessment Schedule

First Author	Number of Assessments/Time
Aalbers	7 per day/ 14 (days)
Bos	10 per day/5 (days)
Bringmann (2016)	10 per day/ 7-14 (days)
Bringmann (2015)	1 per week/ 3-20 (weeks)
Bringmann (2013)	10-11 per day/ 7-12 (days)
Curtiss	1 per day/ 56 (days)
de Vos	3 per day / 30 (days)
Ellison	6 per day/ 21 (days)
Greene (2019)	2 per day/ 30 (days)
Greene (2018)	2 per day/ 30 (days)
Groen (2019)	3 per week/6 (weeks)
Hasmi	10 per day/ 6 (days)
Hoorelbeke	6 per day/7 (days)
Johnson	1 per week/ 9,4 (weeks)
Jongeneel	10 per day/ 6 (days)
Kaiser	1 per day/ 24-63 (days)
Klippel	10 per day/ 5-6 (days)
Levinson (2018)	4 per day/7 (days)
Lutz	4 per day/14 (days)
Oreel	9 per day/ 7 days
Pavani	5 per day/14 days
Pe	8 per day/ 7 (days)
Rath	10 per day/ 6 (days)
Snippe	10 per day/ 6 (days)
van Roekel	3 per day / 30 (days)
Wigman	10 per day/ 5-6 (days)
Contreras	10 per day/ 7 (days)
Drukker	10 per day/ 7(days)
Elovainio	3 per day (7-20 (days)
Faelens	6 per day/ 14 (days)
Geraets	10 per day/ 6-10 (days)
Groen (2020)	5 per day/ 14 (days)
Hoffart	1 per week / 10 (weeks)
Huckins	1 per week/ 40 (weeks)
Kreiter	10 per day/ 7 (days) at three times: baseline, 3 and 6 months
Kuranova	10 per day/ 6 (days)
Lazarus	4 per day/ 15 (days)
Levinson (2020)	1 per month/48 (months)
Martín-Brufau	1 per day/14-20 (days)
Meng	5 per day/13 (days)
Pannicke	4 per day /14 (days)
Thonon	5 per day/14 (days)
McCuish	1 per year/7 (years)

Device & Platform

Regarding the device and platform used to collect the intensive data, we found that 33 articles relied on an electronic device (i.e., 19 used smartphones; 1 used a wristwatch/actigraph; 9 used Psymate/personal pocket computer; 2 used an iPod and 2 relied on a computer), 4 research used a combination of both, an electronic device and paper and pencil, while 6 did not specified the device used to collect the data. Among those investigations using an electronic device, 26 articles reported the platform used to send the questionnaire notification to the electronic device (68.4%), 2 sent notifications via text messages (5.3%), while the remaining (26.3%) did not specify the method used to send the assessments.

Data Collection Design

Our results also yield information related to the data collection design, categorized following the time-series data collection descriptions provided by Myin-Germeys et al. (2018) to better describe the wide range of designs. Accordingly, our results indicated that 42 out of the 43 researchers used time as a sampling unit, while only one did not specify the design used. Among those 42, we found differences in the sampling schedule. For instance, 14 used a fixed sampling (where questionnaires are completed at equal time intervals); 3 used a random sampling schedule (where questionnaires are completed at random intervals); 19 followed a pseudo-random sampling schedule (where questionnaires are completed randomly within a specific time interval), and 6 did not specify the details.

Compensation

Almost half of the articles compensated participants for their time. In particular, 15 articles (34.9%) used financial compensation or gift cards, 2 (4.7%) used course credit, 1(2.3%) reward the compliance by providing feedback or personalized report about participant's performances and 3 (7%) provided a combination of the aforementioned types of compensation (i.e., financial compensation and personalized feedback).

D.5 Data Preprocessing

Node Redundancy

Among articles that assessed for redundancy through statistical methods (n = 10), six articles did so by looking for multicollinearity or high correlations between variables, two did so via factor analysis, and two did so by testing whether the matrix used as input for the network was positive definite and then used the *goldbricker* function in R from the *networktools* package to identify nodes that likely measured the same construct.

Statistical Software and Packages

Almost all authors report using R for network estimation or visualization, but some authors also used other software types (mainly Stata and MPlus; see Figure S15). The specific packages authors reported using are listed in Table S7.

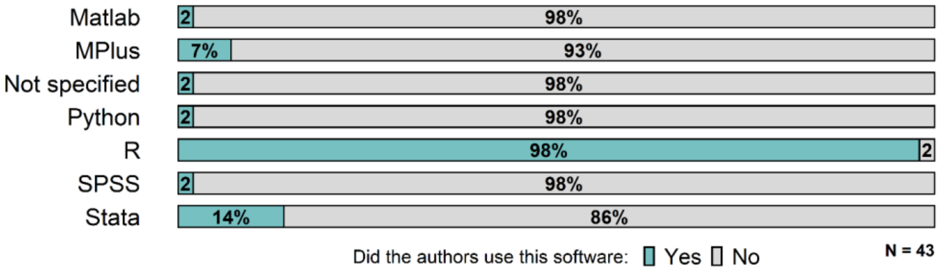


Figure S15. Software used (for network estimation or visualization)

Table S7. Reported Packages Used for Temporal Network Analyses

Software	Temporal network	Contemporaneous network	Between-subjects network
R	mlVAR (n = 21) lm4 (n = 4) nlme (n = 4) graphicalVAR (n =2) qgraph (n = 2) Gimme (n = 1) ImerTEST (n = 1) Bootnet (n =1) Networktools (n = 1)	mlVAR (n = 18) graphicalVAR (n = 2) networktools (n = 2) qgraph (n = 1) nlme (n = 1) Bootnet (n = 1) lme4 (n = 1)	mlVAR (n = 14) graphicalVAR (n = 1) Bootnet (n = 1) lme4 (n = 1) qgraph (n = 1) networktools (n = 1)
Mplus	DSEM (Dynamic Structural Equation Modeling Package) (n = 1)		
Python	Tigramite (n = 1)	Tigramite (n = 1)	

Variable Type and Node Construction

Thirty-four articles report measuring ordinal (i.e., Likert scale variables), while 29 articles report measuring continuous variables (see Figure S16). Only 1 article reported assessing a binary variable; they report multiple ordinal variables and one binary variable, and used univariate estimation of mlVAR models to estimate their network; they used logistic mixed-effect regression models when the binary variable was the outcome variable, and standard linear mixed-effect models for the other variables.

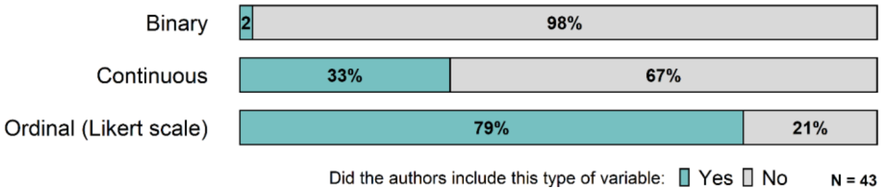


Figure S16. Variable Type

Almost all articles report constructing a node using only 1 item, while only 16% of articles report averaging multiple items to form a node, and a few articles report using multiple items to form a node, but not how they combine these items (see Figure S17).

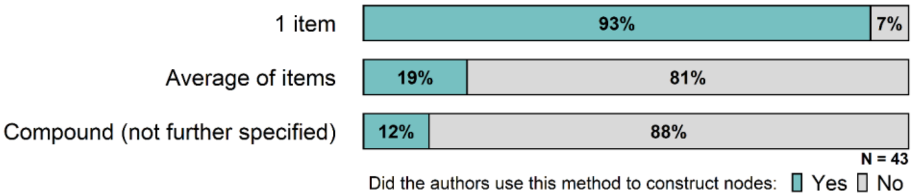


Figure S17. Node Construction

Among articles that calculated centrality indices for any of their estimated networks (n = 30 articles), 87% (n = 26) estimated strength centrality, 33% (n = 10) estimated betweenness centrality, 23% (n = 7) estimated closeness centrality, and 10% (n = 3) estimated expected influence centrality (see Figure S18).

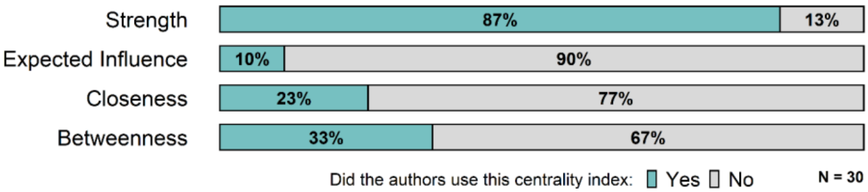


Figure S18. Centrality Indices (For Any Network Type)

Missing Data

To see all details about how authors performed specific actions or justified a lack of action regarding missing data, as well as specific references, see here: https://airtable.com/shrR0Fk8PA7YbGffL.

D.6 Assumptions: Additional Details and Links to References

Here we list all methods used to assess and correct assumptions, when the full list of methods was not reported in the results section of the manuscript (due to space constraints).

Normality (Assessing and Correcting)

The full list of articles that assessed and corrected normality (as well as their detailed methods) can be found here: https://airtable.com/shrxWbMJdPhbPMjAF.

Stationarity (Assessing and Correcting)

The full list of articles that assessed and corrected stationarity (as well as their detailed methods) can be found here: https://airtable.com/shrqq9nnRO5O98ez2.

In addition, here are more details on the methods used to correct for stationarity:

Eighteen articles corrected for stationarity (this was sometimes related to if they had found a violation of the assumption of stationarity, but often times without first assessing stationarity as well) using the following methods:

- Detrend data (n = 10)
    - Some do not specify their methods (n = 2)
    - One article specifies they do this by fitting a nonparametric smoothing spline to the univariate time series of each item & then subtracting this curve from the original time series (n = 1)
    - Most articles do some version of this: subtract both the person mean & linear trend from the data by regressing each symptom separately on a linear function of time and using the residuals as the predictor variable (this is the procedure of Curran & Bauer, 2011) (n = 8)
- Include a linear trend (n = 8)

Redundancy

The full list of articles that assessed node redundancy, along with details about their methods, can be found here: https://airtable.com/shrzyUXhkhdRKNBYP.

D.7 Data Analysis

Centrality Estimates for Different Network Types

Seventy percent of articles (n = 30 out of 43) estimated centrality indices for their temporal networks, while 59% (n = 13 of 21) did so for contemporaneous networks and 71% (n = 5 of 17) for between subject networks; information on which centrality indices are used can be found in the supplementary materials.

Additional Info About Comparisons

Twenty-four articles compared at least two networks, using the following methods:

- Compare centrality:

  - Correlate centrality indices of two networks (to compare similarity between networks, here to compare to a validation dataset) (n = 1)
  - Permutation tests (n = 4)
  - Bayesian model so can use point estimates & credible estimates (e.g., if 95% credible interval includes 0 or not) (n = 1)
- Formal test to compare topology:
  - Correlating edge strengths (to compare similarity between networks, here to compare to a validation dataset) (n = 1)
  - Permutation tests (to compare network density and/or connectivity, or to compare specific edges) (n = 10)
  - *t*-tests (to compare regression weights/edges or overall connectivity) (n = 3)
  - Compare models using a fit index (e.g., Akaike information criterion) (n = 1)
- Compare visually (n=11)

The full list of articles that statistically comparing centrality indices can be found here: https://airtable.com/shrLRzHxgSm7rusqo while those that statistically compared topology can be found here: https://airtable.com/shrmxgJN8ZonB6GcV.

Visualization Adjustments

When computing networks, one can implement visualization adjustments for interpretation purposes. Of the articles reviewed, 25 (58%) made visualization adjustments based on the significance of edges, with only significant edges are represented on the network. Six (14%) articles made adjustments based on the significance of edges as well as adjustments to allow comparisons between networks (e.g., scaling the edges of all models in relation to the highest value detected for each type of model).[1] One article (2%) made adjustments based on the strength of nodes, with the size of the nodes representing the strength of each variable. Fifteen articles (35%) did not make any visualization adjustments.

Visualization Algorithms

The placement of nodes in the network can be in a predetermined shape (e.g., circle) or computed by using visualization algorithms. Ten (23.3%) articles used the *Fruchterman-Reingold* algorithm where the centrality index determines the place of the node (i.e., the higher the centrality index, the more central the node is in the network graph). Twenty-nine (67.4%) articles did not specify the algorithms used.

[1] Note: some articles combined multiple methods

Additional Information on Robustness

Table S8 provides information on the six articles that assessed the robustness and/or stability of the centrality or edge-weights estimates (and for which networks). This information can also be found here: https://airtable.com/shr033LmROeB3keiD.

Table S8. Methods Used to Assess Robustness

Article	Method to assess robustness	Model type
Bringmann (2013)	Bootstrap CIs of centrality estimates - See correction (https://doi.org/10.1371/journal.pone.0096588) - to take the time dependency into account when calculating bootstrap confidence intervals for their betweenness centrality estimates, they use nonparametric bootstrapping directly on the multilevel model - they report this is very computationally intense & the R code cannot be run on a standard computer. They share their updated code as Appendix S1 in the correction.	Sequential univariate estimation of a multilevel VAR model
Geraets (2020)	Stability (dropping one of five predictors) - They performed post-hoc checks for coefficient stability by omitting one of the five predictors (results showed similar coefficients, and so stable coefficients). They do not provide further information on their methods or results and do not share their R code.	Sequential univariate estimation of a multilevel VAR model
Groen (2020)	Bayesian credible intervals (e.g., uncertainty estimate for each parameter); (the authors do not explicitly call this assessing network stability or robustness and they mostly use it to compare groups, but it nonetheless gives information on the stability/robustness of the network estimates)	Multivariate estimation of a multilevel VAR model
Jongeneel (2020)	Stability (dropping participants) - They estimated 100 networks that included a randomly selected 80% of participants, then they calculated Spearman rank correlation between the centrality measures of the main model and the 100 new networks, then plotted the 100 correlations in a histogram (correlations above .7 were interpreted as similar to the main model). They share this R code in the online supplementary materials (included in the MainAnalysis R code file)	Sequential univariate estimation of a multilevel VAR model
Lazarus (2020)	Accuracy (non-parametric bootstrap CI around edges) - They check the accuracy of the networks by performing non-parametric bootstrap confidence intervals around the edge-weights of all networks, using the *bootnet* package and following the Epskamp et al. (2018) method; the authors' full explanations are here: https://osf.io/ju5yk/. They also perform case-dropping bootstraps to assess the stability of the centrality indices, also using the *bootnet* package. They do share their code for their main analyses but not for these bootstrap networks.	Pooled VAR
McCuish (2020)	Stability of centrality (bootnet package) - They used the bootnet package in R to measure the correlation stability (CS) coefficient for each of the centrality indices they used (although they only report doing so for the between-subjects network, not the contemporaneous or temporal networks estimated), citing Epskamp et	Pooled VAR

al. (2018). Although they do not link to their code in their article, they do share their code for these analyses on the first author's Github page: https://github.com/EvanMcCuish/YPI-Network-Analysis/blob/master/PPD%20L%20Net%20-%20Bootnet%20(Subscales).R.

Additional Information on Sensitivity Analyses

The full list of articles that conducted sensitivity analyses, as well as the information on what they focused on in their sensitivity analyses, can be found here: https://airtable.com/shrEY8xEAcXCRgpLg.

Appendix E:
Supplementary Materials from Chapter 7

These supplementary materials can also be found at: https://osf.io/g5t7h.

E.1 Additional Information on Data Collection & ESM Items

Sample Size & Participant Information

We preregistered recruiting a minimum of 40 participants with 80% compliance. However, since it took at least 45 days to determine if participants had answered at least 80% of the daily diary surveys (e.g., minimum 45 prompts), we would not know how many of our participants had completed 80% completion until near the end of our data collection period. We therefore decided to continue enrolling participants until we had reached 50 parents.

We had preregistered that we would exclude participants from analyses if they answered less than 30% of daily diary prompts (following Aalbers et al., 2019), since these participants would have very few timepoints (e.g., less than 17). However, all participants but two answered at least 85% of daily diary surveys. One participant answered 70% of surveys, and another participant dropped out of the study due to incompatibility with their work schedule, but they had already completed 34 daily diary surveys (61%) by then, so their data are still included in the analyses.

Regarding participants' income, 1 participant (2%) reported a net monthly income of €1500-€2500; 18 participants (38.3%) reported €2500-€4500; 26 participants (55.3%) reported more than €4500; and two participants preferred not to specify. The average Belgian net income for 2019 was of 19,105€ per person, or 1592€ per month (*Statbel: 2019 Average Belgian Income*, 2021). Therefore, the average monthly income for two people (as the 47 participants in the main sample were all co-parents) was of 3184€. Participants included in the study thus mostly reported average-to-high net monthly incomes for Belgian families.

Nineteen participants (47.5%) answered that they were currently following a psychological treatment, which fits the sample not only due to its clinical nature but also since recruitment took part largely through clinicians of various types. Among these participants, the length of treatment was mostly under a year but varied from 1 month to 15 years.

Further Information on ESM Procedure

Participants answered all questionnaires via a *formr* webpage and were prompted to answer daily diary surveys with either an email or a text message, depending on their preference. Parents received a reminder notification two hours later if they had not yet filled in the survey, and the daily diary survey expired after five hours. After the final daily diary survey, participants had the opportunity to answer a very brief feedback survey, and then a debriefing video-call with the experimenter if they wished.

Because of an issue with *formr* and their servers in Germany midway through the study (causing timing issues), some participants were able to answer a daily diary questionnaire twice in one evening, which normally would not have been possible. We therefore decided to only keep the first response anytime this occurred (as if the participant had only had one chance to respond, like in the normal procedure).

E.2 Additional Information on Preprocessing Decisions and Results

Missing Data

Since we had limited amounts of missing data (6% overall), we decided to not take any specific actions (e.g., imputation). The mlVAR package does not take any specific action with missing data and so automatically deals with missing data using listwise deletion (Jordan et al., 2020).

Descriptive Statistics: ESM items

For each participant, we calculated the within-person mean and standard deviation for each variable; we then averaged these values to produce group-level intra-individual means and standard deviations for each variable (see Table S9). We also calculated the intraclass coefficients of each item to ensure there was sufficient within-person variance (e.g., an ICC under 0.5 indicates for an item indicate greater within-person than between-person variability; see Table S9).

Table S9. Intra-individual Means, Standard Deviations, and ICCs for Each ESM Item

Daily diary item	iiMean	iiSD	ICC
Emotional Exhaustion	29.06	24.85	0.23
Emotional Distance	24.65	16.43	0.37
Feeling Fed Up	22.34	22.33	0.16
Partner Support	60.46	26.00	0.40
Partner Conflict	18.87	23.05	0.20
Kids' Behavior	26.12	23.00	0.18
Behavior toward Kids (+)	76.33	16.75	0.38
Behavior toward Kids (-)	23.21	24.00	0.17

Resources	67.78	25.47	0.25
Social Support	25.35	32.22	0.13
Time Spent with Kids	8.02	3.63	0.23

Note. ICC = intraclass correlations; iiMean = intraindividual Mean; iiSD = intraindividual standard deviation.

E.3 Centrality Indices

We calculated how connected each node is to the rest of the network by generating a centrality index for each node. We therefore computed the strength index for all networks, which measures the summed weight of a node's edges (McNally, 2016). Since the temporal network includes directed edges, we calculated in- and out-strength. A node with high in-strength is highly influenced by the other nodes in the network (i.e., has many edges pointing toward it); a node with high out-strength predicts many other nodes (i.e., has many edges pointing toward other nodes).

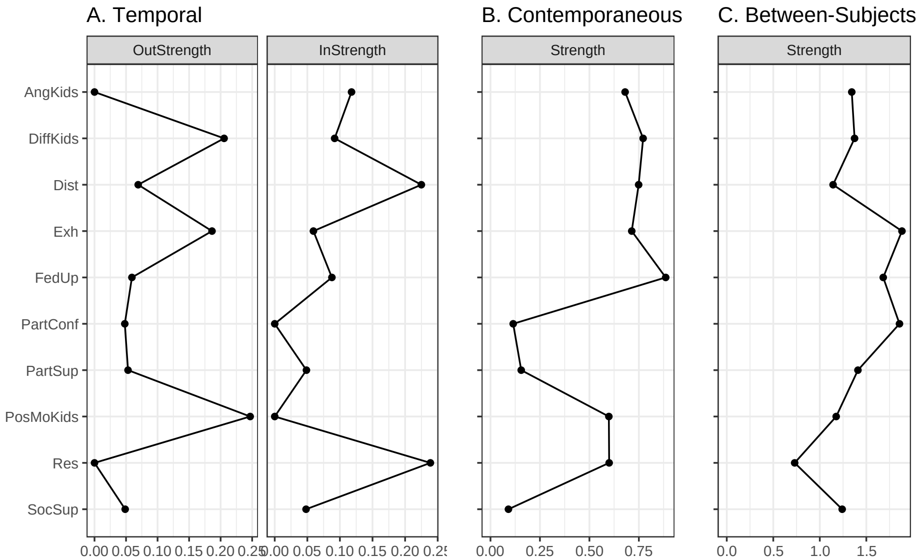


Figure S19. Raw Strength Centrality Indices for all Networks

Note. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PartnerConf = partner conflict; PartnerSupp = partner support; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support.

Stability Analyses for Centrality Indices

Although not preregistered, we also wanted to assess the stability of our centrality analyses, following the method and code of Jongeneel et al. (2020). We calculated 200 networks that included a randomly sampled 80% of participants (i.e., 38 participants) and computed centrality indices for these networks (strength for the contemporaneous and between-subject networks, and in- and out-strength for the temporal networks). We then correlated the centrality indices of the original model with the 100 randomly sampled models using Spearman rank correlations and plotted these correlations in a histogram (see Figure S20).

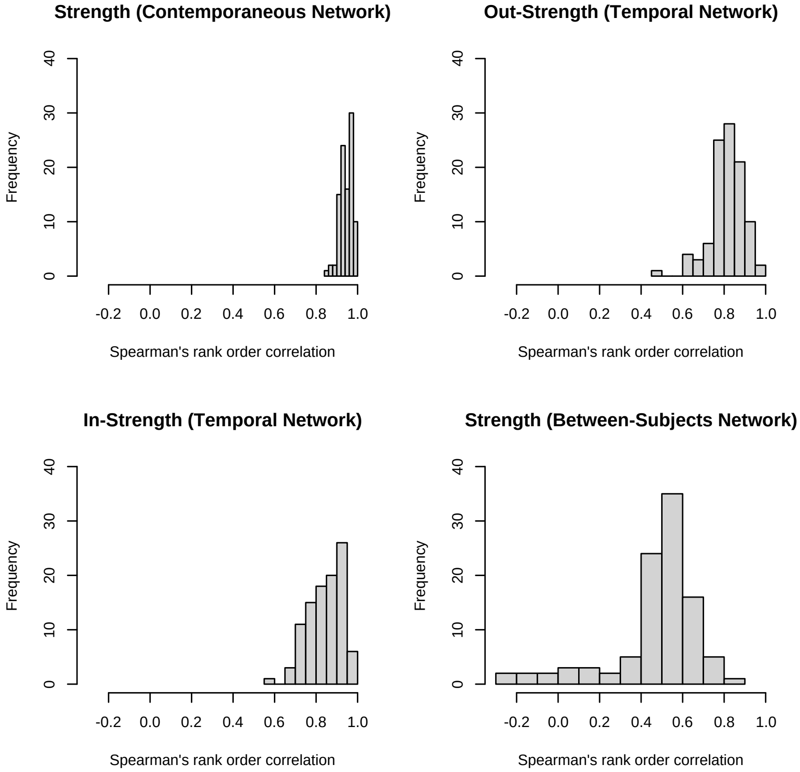


Figure S20. Correlations between case-dropped network and original network

E.4 Sensitivity Analyses

We generated four additional mlVAR models as sensitivity analyses, to examine if they altered the pattern of results compared to the model in the main manuscript (which included raw data from all parents with a coparent, not including the variable 'Time with Kids').

Model Estimated with Raw Data of All Parents with Coparents (n=47)

The first sensitivity analysis includes the sample in the main manuscript (raw data of mothers with coparents, n=43), including the four fathers (n=47). Since our sample was 91% mothers, a reviewer suggested that our results might not fully represent or generalize to all parents but be specific to mothers. We thus decided to only include mothers in the main analyses, but also examine a network with all parents with coparents, including the fathers, as a sensitivity analysis.

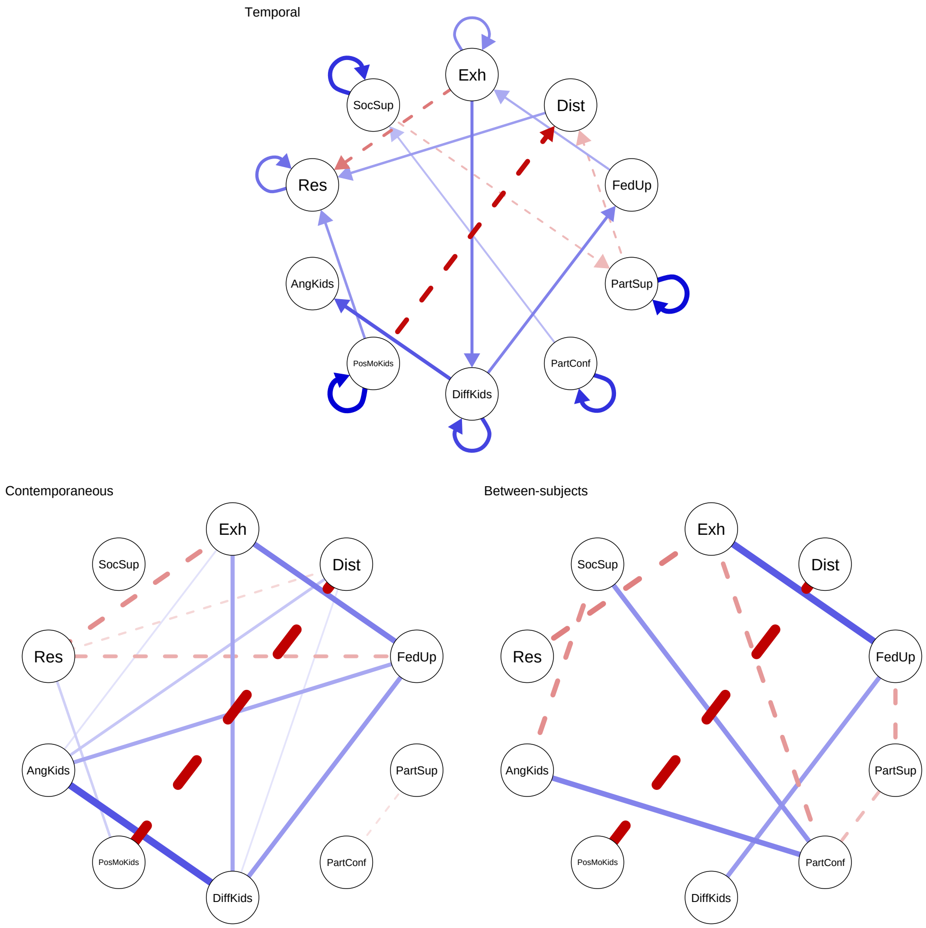


Figure S21. Network Analyses for Raw Data of All Parents with Coparent (n=47), including Fathers

When comparing this network (see Figure S21) with the network analysis in Chapter 7 with raw data (see Figure 18), the contemporaneous networks are very similar (just one faint additional negative edge in Figure 18, connecting Social Support and Difficult to Manage – Kids) while the between subject networks are identical. The temporal networks have some small differences in edges: Figure 18 has a faint additional positive edge pointing from Getting Angry toward Children toward Resources, while Figure S21 has two additional edges: one pointing from Sharing Positive Moments with Children toward Resources, and another connecting Feeling Fed Up with Emotional Exhaustion. All thick edges can be found in all networks, but the potential feedback loop between Exhaustion, Difficult to Manage – Kids, and Feeling Up is no longer present in the network with only mothers (Figure 18), as one edge is missing.

Model Estimated with Transformed Data (Approximating a Normal Distribution)

This sensitivity analysis includes transformed data since the raw data is not normally distributed (with all parents with coparents, n=47). We had preregistered that this transformation would be a log transformation, but this did not alter the variables' distribution to be closer to a normal distribution (e.g., skew and kurtosis had larger values). We therefore used a Gaussianize transformation from the R package *LambertW* which transforms heavy tailed and skewed data (Goerg, 2015).

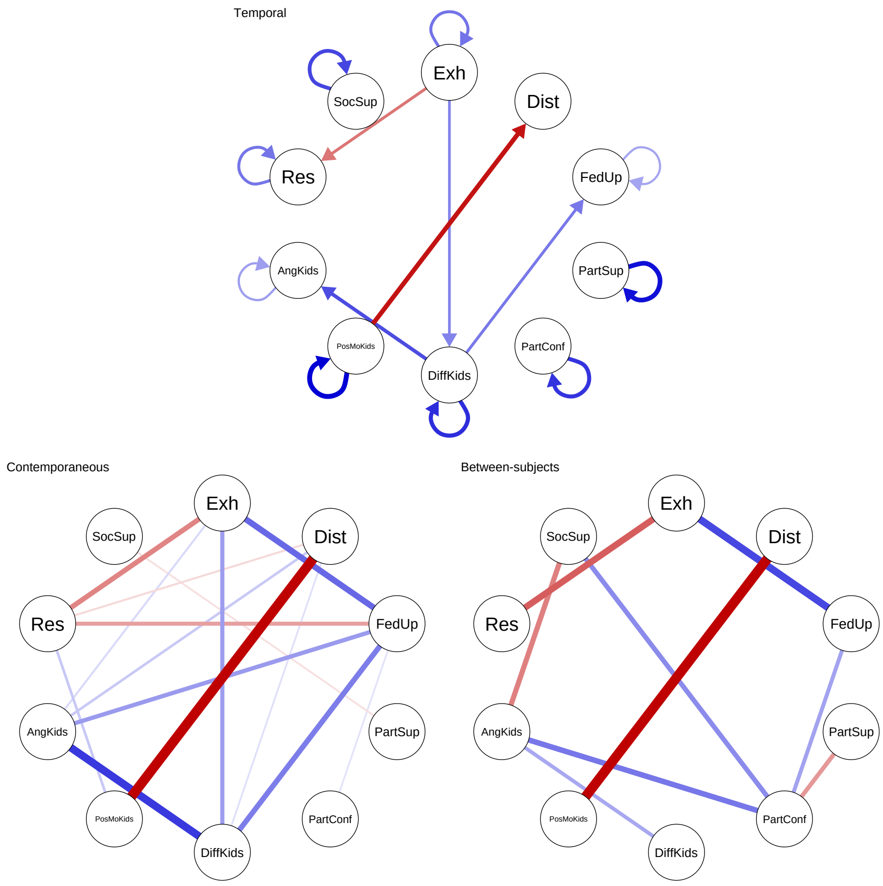


Figure S22. Network Analyses for Gaussianized Transformed Data.

Note. Blue edges represent positive associations, while red edges represent negative associations. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PartnerConf = partner conflict; PartnerSupp = partner support; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support.

Model Estimated with "Time with Kids" Variable

The second sensitivity analysis included the same 10 variables and raw data as in the network in the main manuscript (with all parents with coparents; n=47), with the addition of the variable "Time with Kids," since it also had large within-person variability.

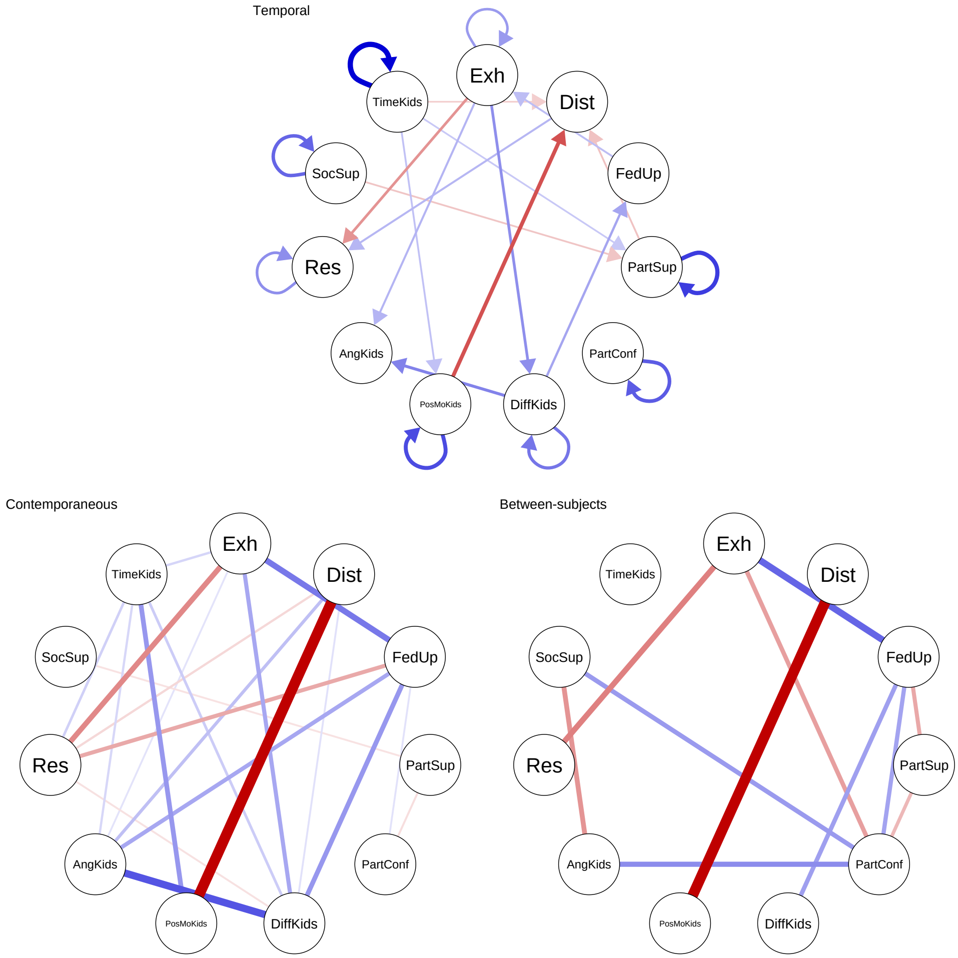


Figure S23. Network Analyses for Variables with 11 Variables (Including 'Time with Kids')

Note. Blue edges represent positive associations, while red edges represent negative associations. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PartnerConf = partner conflict; PartnerSupp = partner support; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support.

When comparing this network (see Figure 23) with the network analysis in the main manuscript with 10 variables (see Figure 18), the networks are very similar overall. For the temporal network, Time with Kids has a few thin outgoing edges but most of the thicker edges from Figure 18 are still present in the network (with the exception of the edge point from Positive Moments - Kids to Resources). The contemporaneous network similarly is comparable to that in Chapter 7: although Time with Kids has many thin edges connecting it to other nodes, all thick edges in Figure 18 are also present in Figure 23. The only additional thicker edge links Time with Kids to Positive Moments - Kids. The between-subject networks are almost identical, as interestingly enough Time with Kids is not connected to any other node. However, there is an additional edge in Figure 23 between Feeling Fed Up and Partner Conflict.

This model has interesting implications for time spent with kids: that although the time a parent spends with their children has a small impact on whether they feel distant with their children the next day, and is linked with exhaustion and sharing positive moments within the same day, it is not connected to any other node in the between-subjects network. This implies that when you zoom out from a daily perspective and look at the average-level, the time spent with children is not associated with the features of parental burnout or the interactions with the child, partner, or wider family context. This is, however, an exploratory sensitivity analysis, and so no firm conclusions can be drawn without substantial further research. However, these results do align with previous research suggesting that the overall quantity of time parents spend with kids is not as important as *how* they spend that time (Hsin, 2009).

Model Estimated with All 50 Parents (Including Single Parents), Without Partner Variables

The third sensitivity analysis includes all 50 parents who participated in the data collection process, including single parents (n=3); it therefore does not include the two partner-related variables: Partner Support and Partner Conflict.

Overall, these networks (see Figure S24) are relatively similar to the networks in Chapter 7 with 47 parents and the partner variables (see Figure 18). In the temporal network, most of the edges are the same, with one thin edge missing in Figure S24 (pointing from Positive Moments with Kids to Resources) and a few additional edges in Figure S24 (from Exhaustion toward Angry with Kids, as well as additional autocorrelations for Feeling Fed Up and Angry with Kids).

The contemporaneous networks were very similar, which is expected since the Partner Support and Partner Conflict variables in Figure 18 were only thinly connected to each other and not connected at all to any other nodes in the network. Lastly, the between-subjects networks are overall relatively similar. The main differences include a few additional edges in the network with all 50 parents in Figure S24 (specifically connecting Feeling Fed Up and

Social Support, Feeling Fed Up and Angry with Kids, and Difficult to Manage and Angry with Kids) and one edge missing (connecting Social Support and Angry with Kids).

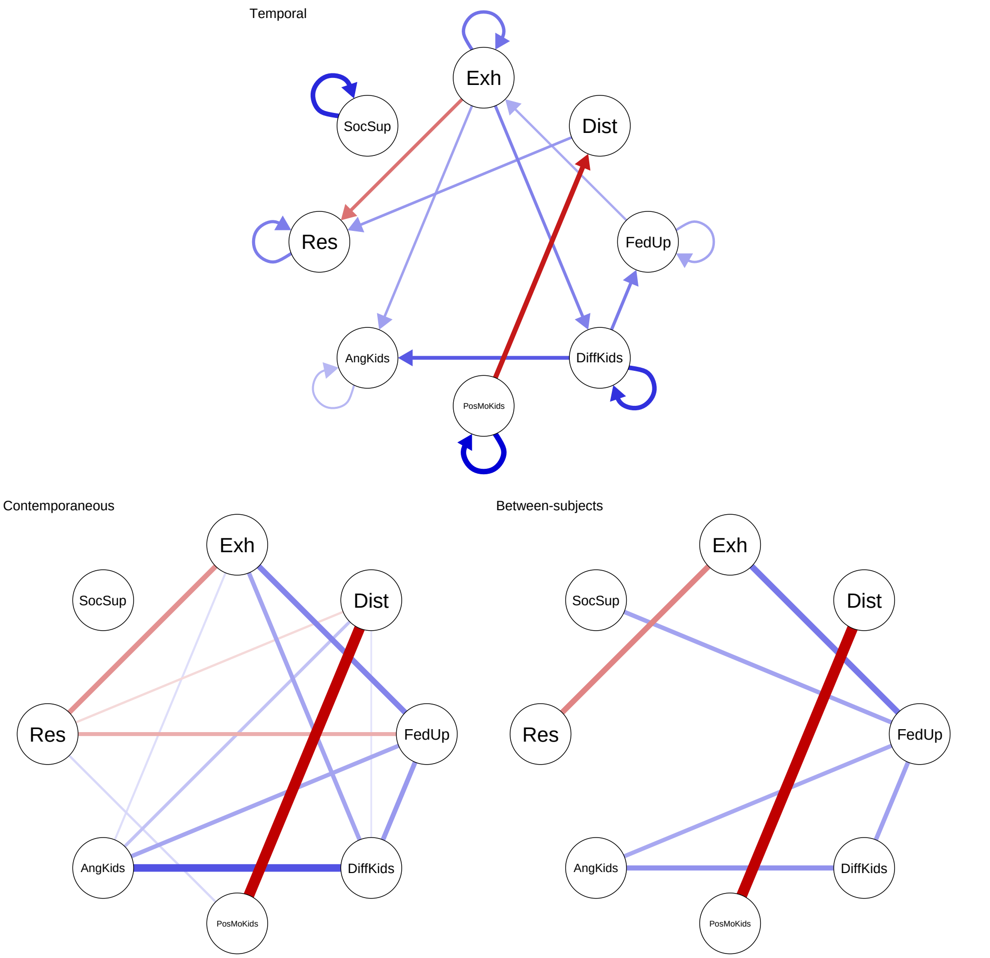


Figure S24. Network Analyses for Variables with 50 Participants (Including Single Parents), Without Partner-related Variables

Note. Blue edges represent positive associations, while red edges represent negative associations. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PartnerConf = partner conflict; PartnerSupp = partner support; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support.

E.5 Edge Values of All Networks in Manuscript (Mothers with Coparents, n=43)

Table S10. Edge values of the temporal network (Figure 18)

From	To	Edge value (fixed effects)	SE	*p*	SD from random effects
Exh	**Exh**	**0.085**	**0.028**	**0.002**	**0.036**
Exh	Dist	-0.04	0.026	0.117	0.051
Exh	FedUp	0.045	0.03	0.133	0.056
Exh	PartSup	-0.016	0.027	0.542	0.077
Exh	PartConf	0.031	0.029	0.288	0.056
Exh	**DiffKids**	**0.092**	**0.03**	**0.003**	**0.081**
Exh	PosMoKids	0.028	0.027	0.305	0.08
Exh	AngKids	0.056	0.033	0.092	0.107
Exh	**Res**	**-0.089**	**0.027**	**0.001**	**0.005**
Exh	SocSup	-0.006	0.034	0.851	0.116
Dist	Exh	-0.038	0.036	0.287	0.095
Dist	Dist	0.022	0.033	0.520	0.104
Dist	FedUp	-0.037	0.037	0.315	0.09
Dist	PartSup	-0.035	0.032	0.262	0.086
Dist	PartConf	0.000	0.033	0.990	0.063
Dist	DiffKids	-0.068	0.04	0.086	0.135
Dist	PosMoKids	-0.044	0.035	0.213	0.125
Dist	AngKids	-0.031	0.035	0.376	0.064
Dist	**Res**	**0.068**	**0.035**	**0.050**	**0.089**
Dist	SocSup	-0.033	0.039	0.399	0.12
FedUp	Exh	0.058	0.030	0.058	0.087
FedUp	Dist	0.023	0.026	0.376	0.07
FedUp	FedUp	0.046	0.031	0.143	0.084
FedUp	PartSup	0.012	0.029	0.678	0.103
FedUp	PartConf	0.027	0.032	0.410	0.107
FedUp	DiffKids	0.004	0.029	0.889	0.065
FedUp	PosMoKids	-0.021	0.028	0.454	0.091
FedUp	AngKids	-0.02	0.031	0.513	0.085
FedUp	Res	-0.01	0.027	0.717	0.047
FedUp	SocSup	0.034	0.031	0.274	0.086
PartSup	Exh	0.05	0.027	0.068	0.071
PartSup	**Dist**	**-0.055**	**0.026**	**0.030**	**0.078**
PartSup	FedUp	0.021	0.027	0.446	0.053
PartSup	**PartSup**	**0.185**	**0.03**	**0.000**	**0.132**
PartSup	PartConf	-0.023	0.028	0.416	0.08
PartSup	DiffKids	0.025	0.027	0.341	0.057
PartSup	PosMoKids	0.013	0.026	0.614	0.092
PartSup	AngKids	0.001	0.032	0.971	0.124
PartSup	Res	0.001	0.027	0.975	0.068
PartSup	SocSup	-0.027	0.034	0.431	0.142
PartConf	Exh	0.03	0.024	0.201	0.066
PartConf	Dist	-0.037	0.022	0.092	0.068
PartConf	FedUp	0.023	0.024	0.334	0.052
PartConf	PartSup	0.034	0.021	0.100	0.059

From	To	Edge value (fixed effects)	SE	*p*	SD from random effects
PartConf	**PartConf**	**0.154**	**0.032**	**0.000**	**0.147**
PartConf	DiffKids	0.017	0.028	0.546	0.105
PartConf	PosMoKids	0.010	0.020	0.630	0.052
PartConf	AngKids	-0.009	0.026	0.738	0.085
PartConf	Res	-0.01	0.022	0.669	0.049
PartConf	**SocSup**	**0.056**	**0.026**	**0.031**	**0.085**
DiffKids	Exh	0.049	0.033	0.133	0.111
DiffKids	Dist	0.027	0.028	0.329	0.081
DiffKids	**FedUp**	**0.092**	**0.034**	**0.006**	**0.108**
DiffKids	PartSup	0.018	0.024	0.448	0.027
DiffKids	PartConf	-0.011	0.030	0.707	0.072
DiffKids	**DiffKids**	**0.125**	**0.039**	**0.001**	**0.167**
DiffKids	PosMoKids	-0.008	0.026	0.753	0.064
DiffKids	**AngKids**	**0.123**	**0.032**	**0.000**	**0.091**
DiffKids	Res	-0.040	0.030	0.182	0.086
DiffKids	SocSup	-0.021	0.034	0.537	0.116
PosMoKids	Exh	-0.044	0.039	0.266	0.135
PosMoKids	**Dist**	**-0.174**	**0.038**	**0.000**	**0.151**
PosMoKids	FedUp	-0.008	0.034	0.804	0.054
PosMoKids	PartSup	-0.005	0.033	0.893	0.113
PosMoKids	PartConf	0.016	0.034	0.635	0.081
PosMoKids	DiffKids	-0.028	0.037	0.439	0.106
PosMoKids	**PosMoKids**	**0.177**	**0.039**	**0.000**	**0.161**
PosMoKids	AngKids	-0.031	0.037	0.400	0.109
PosMoKids	Res	0.065	0.040	0.104	0.147
PosMoKids	SocSup	-0.064	0.037	0.085	0.108
AngKids	Exh	0.005	0.029	0.869	0.077
AngKids	Dist	0.004	0.025	0.866	0.056
AngKids	FedUp	-0.007	0.028	0.796	0.037
AngKids	PartSup	0.029	0.029	0.307	0.108
AngKids	PartConf	-0.025	0.028	0.365	0.069
AngKids	DiffKids	0.031	0.029	0.292	0.076
AngKids	PosMoKids	0.004	0.025	0.873	0.073
AngKids	AngKids	0.041	0.028	0.141	0.048
AngKids	**Res**	**0.055**	**0.026**	**0.038**	**0.044**
AngKids	SocSup	-0.015	0.028	0.594	0.057
Res	Exh	-0.013	0.027	0.642	0.069
Res	Dist	-0.029	0.026	0.266	0.09
Res	FedUp	0.022	0.028	0.421	0.061
Res	PartSup	-0.006	0.022	0.771	0.029
Res	PartConf	0.043	0.027	0.117	0.067
Res	DiffKids	0.023	0.028	0.406	0.074
Res	PosMoKids	0.019	0.025	0.446	0.074
Res	AngKids	0.000	0.028	0.988	0.07
Res	**Res**	**0.096**	**0.03**	**0.001**	**0.108**
Res	SocSup	0.026	0.031	0.407	0.113
SocSup	Exh	0.004	0.023	0.869	0.056
SocSup	Dist	0.015	0.021	0.494	0.067
SocSup	FedUp	-0.004	0.024	0.881	0.06

From	To	Edge value (fixed effects)	SE	*p*	SD from random effects
SocSup	**PartSup**	**-0.058**	**0.025**	**0.020**	**0.109**
SocSup	PartConf	0.003	0.022	0.882	0.044
SocSup	DiffKids	-0.012	0.023	0.607	0.054
SocSup	PosMoKids	-0.003	0.022	0.885	0.083
SocSup	AngKids	-0.02	0.026	0.444	0.094
SocSup	Res	0.011	0.022	0.603	0.049
SocSup	**SocSup**	**0.142**	**0.032**	**0.000**	**0.146**

Note. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PartnerConf = partner conflict; PartnerSupp = partner support; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support; SE = standard error. Bold text denotes an edge present in the temporal network (as the p-value < .05).

Table S11. Edge values of the contemporaneous network (Figure 18)

Node 1	Node 2	*p* (Node 1>2)	*p* (Node 2>1)	Partial correlation	SD (random effects)
Dist	Exh	0.852	0.896	-0.001	0.081
FedUp	Exh	0.000	0.000	0.259	0.079
FedUp	Dist	0.037	0.068	0.045	0.031
PartSup	Exh	0.095	0.11**0**	-0.038	0.031
PartSup	Dist	0.41**0**	0.721	0.016	0.071
PartSup	FedUp	0.081	0.041	-0.042	0.022
PartConf	Exh	0.211	0.18	0.031	0.042
PartConf	Dist	0.089	0.097	0.049	0.083
PartConf	FedUp	0.108	0.041	0.048	0.059
PartConf	PartSup	0.036	0.015	-0.067	0.088
DiffKids	Exh	0.000	0.000	0.171	0.043
DiffKids	Dist	0.043	0.017	0.050	0.038
DiffKids	FedUp	0.000	0.000	0.208	0.133
DiffKids	PartSup	0.302	0.249	0.024	0.019
DiffKids	PartConf	0.249	0.158	0.03**0**	0.037
PosMoKids	Exh	0.261	0.141	-0.03**0**	0.035
PosMoKids	Dist	0.000	0.000	-0.506	0.194
PosMoKids	FedUp	0.995	0.979	0.**000**	0.044
PosMoKids	PartSup	0.264	0.478	0.022	0.039
PosMoKids	PartConf	0.406	0.293	0.024	0.052
PosMoKids	DiffKids	0.679	0.826	0.008	0.045
AngKids	Exh	0.024	0.038	0.059	0.069
AngKids	Dist	0.000	0.000	0.114	0.085
AngKids	FedUp	0.000	0.000	0.173	0.088
AngKids	PartSup	0.351	0.674	-0.017	0.06
AngKids	PartConf	0.272	0.194	0.028	0.031
AngKids	DiffKids	0.000	0.000	0.346	0.161
AngKids	PosMoKids	0.82**0**	0.639	0.01	0.083
Res	Exh	0.000	0.000	-0.231	0.138
Res	Dist	0.000	0.002	-0.087	0.057
Res	FedUp	0.000	0.000	-0.15	0.112
Res	PartSup	0.267	0.339	-0.028	0.063

Res	PartConf	0.709	0.869	-0.008	0.086
Res	DiffKids	0.285	0.231	-0.03	0.064
Res	PosMoKids	0.034	0.012	0.084	0.141
Res	AngKids	0.695	0.598	-0.012	0.051
SocSup	Exh	0.161	0.171	-0.034	0.046
SocSup	Dist	0.405	0.116	0.027	0.023
SocSup	FedUp	0.2**00**	0.344	0.028	0.043
SocSup	PartSup	0.096	0.025	-0.048	0.044
SocSup	PartConf	0.704	0.864	-0.009	0.102
SocSup	DiffKids	0.029	0.038	-0.055	0.053
SocSup	PosMoKids	0.97**0**	0.761	-0.004	0.063
SocSup	AngKids	0.704	0.444	-0.015	0.06**0**
SocSup	Res	0.136	0.04**0**	0.042	0.043

Note. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PartnerConf = partner conflict; PartnerSupp = partner support; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support; SE = standard error. Bold text denotes an edge present in the contemporaneous network, using the and rule (i.e., p-values from both directions of prediction must be significant at $p < .05$).

Table S12. Edge values of the between subjects network (Figure 18)

Node 1	Node 2	*p* (Node 1>2)	*p* (Node 2>1)	Partial Correlation
Dist	Exh	0.596	0.380	-0.092
FedUp	**Exh**	**0.000**	**0.000**	**0.623**
FedUp	Dist	0.218	0.895	0.073
PartSup	**Exh**	**0.367**	**0.001**	**0.222**
PartSup	Dist	0.653	0.593	-0.059
PartSup	**FedUp**	**0.026**	**0.026**	**-0.35**
PartConf	**Exh**	**0.016**	**0.013**	**-0.323**
PartConf	Dist	0.161	0.634	-0.121
PartConf	FedUp	0.124	0.014	0.276
PartConf	**PartSup**	**0.013**	**0.008**	**-0.301**
DiffKids	Exh	0.512	0.083	0.158
DiffKids	Dist	0.200	0.385	0.139
DiffKids	**FedUp**	**0.004**	**0.000**	**0.437**
DiffKids	PartSup	0.190	0.002	0.308
DiffKids	PartConf	0.627	0.844	0.018
PosMoKids	Exh	0.726	0.653	-0.054
PosMoKids	**Dist**	**0.000**	**0.000**	**-0.946**
PosMoKids	FedUp	0.845	0.131	0.045
PosMoKids	PartSup	0.352	0.738	-0.073
PosMoKids	PartConf	0.956	0.997	-0.004
PosMoKids	DiffKids	0.948	0.907	0.012
AngKids	Exh	0.782	0.948	0.016
AngKids	Dist	0.764	0.825	0.035
AngKids	FedUp	0.387	0.829	0.106
AngKids	PartSup	0.101	0.000	0.333
AngKids	**PartConf**	**0.000**	**0.000**	**0.533**
AngKids	DiffKids	0.018	0.053	0.279
AngKids	PosMoKids	0.595	0.793	-0.014
Res	**Exh**	**0.000**	**0.000**	**-0.459**

Res	Dist	0.707	0.107	0.128
Res	FedUp	0.881	0.373	-0.022
Res	PartSup	0.003	0.23	0.248
Res	PartConf	0.786	0.776	-0.002
Res	DiffKids	0.043	0.414	0.206
Res	PosMoKids	0.524	0.098	0.153
Res	AngKids	0.955	0.438	-0.044
SocSup	Exh	0.752	0.210	-0.037
SocSup	Dist	0.345	0.016	-0.197
SocSup	FedUp	0.157	0.167	0.227
SocSup	PartSup	0.793	0.232	0.084
SocSup	**PartConf**	**0.001**	**0.000**	**0.422**
SocSup	DiffKids	0.439	0.687	0.081
SocSup	PosMoKids	0.166	0.020	-0.229
SocSup	**AngKids**	**0.000**	**0.000**	**-0.440**
SocSup	Res	0.452	0.809	0.066

Note. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PartnerConf = partner conflict; PartnerSupp = partner support; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support; SE = standard error. Bold text denotes an edge present in the between subjects network, using the and rule (i.e., p-values from both directions of prediction must be significant at $p < .05$).

Appendix F:
Supplementary Materials from Chapter 8

These supplementary materials can also be found at: https://osf.io/5sgkr/.

F.1 Further sample information

Data Collection

Dropout

One participant dropped out after completing the briefing interview and five daily surveys; thus, their responses were under the 30% threshold for participants to be included in data analyses, and so their information was not included in the demographic information. Almost all participants answered at least 80% of ESM items; only 1 participant answered less (specifically, 77% of items, equivalent to 43 timepoints).

Compensation

Participants were compensated a total (if they answered at least 80% of daily surveys) of 50€. This is a larger maximum amount than the previous similar study (i.e., 35€; Blanchard et al., 2023), as we found out in the meantime that similar ESM studies in the area were compensating participants 50€ and we wanted to align with this amount. All participants in this study were paid at the same rate.

Further Demographic Information

Employment

In terms of employment, 21 participants (52.5%) worked full time, 10 parents (25%) worked part-time, 1 parent (2.5%) was unemployed, 2 parents (5%) were unable to work due to disability (although one parent noted this was actually a temporary pregnancy-related leave due to job requirements), 3 parents (7.5%) were on parental leave, and 3 parents (7.5%) were stay-at-home parents.

Income

In terms of income, participants reported their net household monthly income as follows: 2 participants (5%) between 1€-1500€, 9 participants (22.5%) between 1500€-2500€, 8 participants (20%) between 2500€-3500€, 11 participants (27.5%) between 3500€-4500€, 7 participants (17.5%) greater than 4500€, and 3 participants (7.5%) preferred not to say. The average net Belgian income for 2020 was of 19,671€ (*Statbel: 2020 Average Belgian Income*, 2022), which comes to 1639.25 monthly (and 3278.5 for dual-income households). More than half (52.5%) of the present sample made above both these thresholds, but a substantial percentage was therefore also underneath these thresholds. Therefore, there is a wide range of incomes across participants, but at least half report average-to-high net monthly incomes for Belgian families.

Education.

In terms of education, participants answered with the number of years starting with secondary school (e.g., schooling after sixth grade, when children are around 12 years of age typically). Answered ranged from 0 to 16, with a mean of 10.28 (SD = 2.61) and a median of 10, which corresponds in typical Belgian education to 6 years of secondary + 3 years of bachelors + 1 year of masters.

F.2 Baseline measures: Further information

Beck Depression Inventory-II (BDI-II)

We assessed depression symptoms using the BDI-II (Beck et al., 1996). Participants answered 21 questions, each time choosing the one statement out of four that most described how they felt during the past two weeks regarding specific symptoms (e.g., *I do not feel sad* to *I am so sad and unhappy that I can't stand it*) on a four-point scale from 0 to 3. Scores could range from 0 to 63. Within this sample, internal reliability was high, with a Cronbach's α of .99.

Generalized Anxiety Disorder Scale (GAD-7).

We examined trait anxiety using the GAD-7 (Spitzer et al., 2006). Participants indicated their anxiety over the past two weeks on seven items (e.g., *worrying too much about different things)* using a four-point scale from 0 (*Not at all*) to 3 (*Nearly every day)*. Scores could range from 0 to 21. Internal reliability was high within this sample, with a Cronbach's α of .89.

Balance of Risks and Resources (BR2) Questionnaire

We examined participants' parenting risks and resources using the BR2, including perfectionist personality traits, stress management capabilities, parenting practices, co-parenting, etc. (Mikolajczak & Roskam, 2018). Each of the 39 items took a bipolar form, with a risk factor on the left side (e.g., *I find it difficult to reconcile my family life and my professional life*) and the corresponding resource on the right side (e.g., *I can easily reconcile my family life and my professional life*). Parents answered each item from -5 to 5, with a negative number indicating the risk factor statement more closely mirrored their experience, and a positive number indicating the resource statement more closely reflected their experience. A rating closer to |5| signifies a stronger endorsement of the (negative or positive) statement, while 0 indicates that the parent possessed neither risk factor nor resource factor. Scores could range from -195 to 195, with a negative score indicating parents have more risks than resources (and vice versa for a positive score).

F.3 Preprocessing

Normality

When assessing normality of the residuals and of within-person centered raw data with Shapiro-Wilk tests, no variables were normally distributed. We attempted to correct data to align with a normal distribution. We had preregistered a first attempt to correct with log-transformation, but within-person centered data had many negative values or values close to 0. We then attempted our alternate preregistered transformation, using the automatic transformation "Gaussianize" from the LambertW package (Goerg, 2015), particularly useful for transforming data with heavy tails. However, the resulting corrected data also was not normally distributed according to Shapiro-Wilk tests, and the skew and kurtosis values were either very close to the raw (within-person-centered) data or worse (see Table S13). We therefore decided not to include a sensitivity analysis with corrected data, as we could not be confident the transformation was closer to a normal distribution then the raw (within-mean-centered) data.

Table S13. Skew and Kurtosis Values for Raw and Gaussianize-Corrected Data (Both Within-Person Mean-Centered)

Variable	Raw Data		Gaussianize-Corrected	
	Skew	Kurtosis	Skew	Kurtosis
Emotional Exhaustion	0.33	2.66	0.38	3.01
Emotional Distance	1.01	4.42	1.2	4.92
Feeling Fed Up	0.81	3.48	0.75	3.66
Partner Support	-0.24	2.54	-0.24	2.54
Partner Conflict	1.65	4.97	1.65	4.97
Difficulty Managing Kids	0.67	3.09	0.58	3.04
Positive Moments with Kids	-0.94	4.8	-1.18	5.86
Anger Toward Kids	0.95	3.24	0.96	3.1
Resources	-0.37	2.9	-0.44	2.99
Social Support	0.89	2.98	1.02	3.14

Note. The raw data is within-person centered, and the gaussianize-centered data is therefore also (as it transforms that data). The Gaussianize transformation is an automatic transformation from the LambertW package especially suited for corrected heavily tailed data (Goerg, 2015).

Stationarity

No subjects had non-stationarity. Some participants had trends for specific variables (e.g., a linear relationship between the ESM responses and time), but when checking with a KPSS test for trend stationarity, the test did not reject the null hypothesis of stationarity. For variables that did not have a linear trend, we checked for level stationarity with the KPSS test.

F.4 Further Information on Group Comparisons for Participant SDs

We performed one-tailed t-tests to compare the intraindividual standard deviation of daily parental burnout features between the sample in the main manuscript (i.e., parents who self-reported as exhausted and overwhelmed by parenting, hereafter termed the "Parental Burnout Sample") with a sample of unselected parents (from Blanchard et al., 2023). We used one-tailed tests as we had hypothesized that parental burnout symptoms would fluctuate less from day to day for the sample of parents in burnout, but we did not find support for this hypothesis (see Table S14).

Table S14. Group Comparisons of Daily Fluctuations of Parental Burnout Features

Variable (iiSD)	Unselected Sample		PB Sample		*t*	*df*	*p*
	M	*SD*	*M*	*SD*			
Emotional Exhaustion	24.31	6.00	28.45	5.88	-3.29	84.43	0.999
Emotional Distance	15.74	7.00	21.67	7.12	-3.95	83.12	0.999
Feeling Fed Up	21.89	6.78	25.91	5.64	-3.08	87.85	0.999

Note. iiSD = intraindividual standard deviation; PB = Parental Burnout.

To ensure that the two samples differed in daily mean levels of parental burnout features, we also did one-tailed t-tests to compare samples. Results support that these groups did not indeed report different mean levels of daily parental burnout (see Table S15).

Table S15. Group Comparisons of Daily Mean Levels of Parental Burnout Features

Variable (iiMean)	Unselected Sample		PB Sample		*t*	*df*	*p*
	M	*SD*	*M*	*SD*			
Emotional Exhaustion	29.20	14.2	41.61	19.2	-3.41	69.96	<0.001
Emotional Distance	23.89	13.5	31.00	18.0	-2.08	70.85	0.021
Feeling Fed Up	22.52	10.6	30.53	16.5	-2.66	63.49	0.005

Note. iiMean = intraindividual mean; PB = Parental Burnout.

F.5 Sensitivity Analyses

Sensitivity Analysis #1: Model Estimated with Data from All Parents with Parental Burnout (n=40)

The first sensitivity analysis included all mothers from analyses in main manuscript (n=36) as well as the four fathers who participated in the study (n=4). When comparing this network (Figure S25) with the main analysis (Figure 19), results are overall very similar—all thick edges are present in both figures.

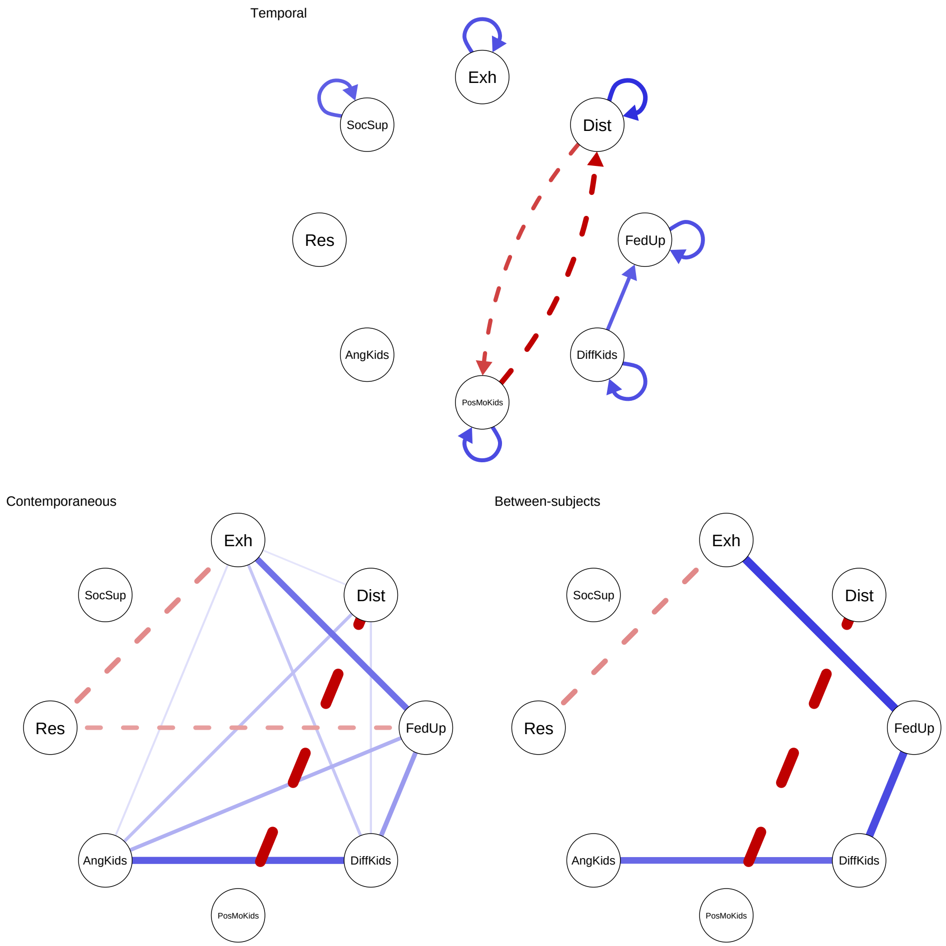


Figure S25. Network Analyses for All Parents (n=40), including Fathers (n=4)

Note. Solid blue edges represent positive associations, while dashed red lines represent negative associations. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support.

The only difference in the temporal network is one edge present in the only-mothers network (Figure 19) but not present in the network of 40 parents (Figure S25): a missing positive edge from Positive Moments with Kids toward Resources. In the contemporaneous network, there are two small edges that are no longer present in the network with 40 parents (Figure S25): a very thin edge connecting Exhaustion to Distance, and a thin edge connecting Positive Moments with Kids toward Resources. The between-subjects networks are identical.

Sensitivity Analysis #2: Adding coparenting variables (and thus removing single parents)

We wanted to include a network with the partner-related nodes (i.e., Partner Support and Partner Conflict), but to do this we had to remove single parents, reducing the sample size to 30 (4 fathers and 26 mothers). The number of nodes for this network was therefore 10. The restricted sample size is why this network is included as a sensitivity analysis and not as the main analysis as was the case in Blanchard et al. (2023), which was an unselected sample (and therefore recruitment was easier and the study included more participants). Note that there were estimation issues when generating this network with correlated random effects (which was specified for all other networks), possibly due to convergence issues with a network with only 30 participants and additional nodes. Because of this, we estimated this network with orthogonal (non-correlated) random effects.

For the temporal network, all edges are thinner in the network with partner-related effects (Figure S26) than in Figure 19; only the autoregressive effects for Partner Support and Partner Conflict are thick. This implies that for parents with coparents, the strongest effect is the self-prediction for partner interactions which subsumes all other variable predictions. Adding these variables into the network also impacted how other variables interacted with Positive Moments with Kids and Resources. In Figure 19, Positive Moments with Kids had a strong autoregressive association and predicted Resources, as well as had a dampening cycle with Distance (e.g., both negatively predicting the other). In Figure S26 with coparents, meanwhile, the only edge relating to Positive Moments with Kids goes toward Distance. Moreover, Resources is positively predicted by Partner Support and positively predicts Distance in turn. These changes imply that adding partner-related variables into the network (and thus limiting our sample to parents with coparents) impacts how parents sharing Positive Moments with Kids interacts with the rest of the network—this variable becomes much less connected. There are also two (thin) edges that are counter-intuitive in Figure S26: Partner Conflict predicts feeling less Distance, while Resources predicts increased Distance. It is possible that when parents are feeling conflict with their partner they feel closer with their children (although this goes against hypotheses such as the spillover hypothesis, where marriage problems might spill over into parenting; Coiro & Emery, 1998), but it is particularly curious that Resources predicts increased Distance. Note, however, that Figure S26 has limited participants for many nodes and because of this likely had convergence issues, and so all results are very preliminary and exploratory.

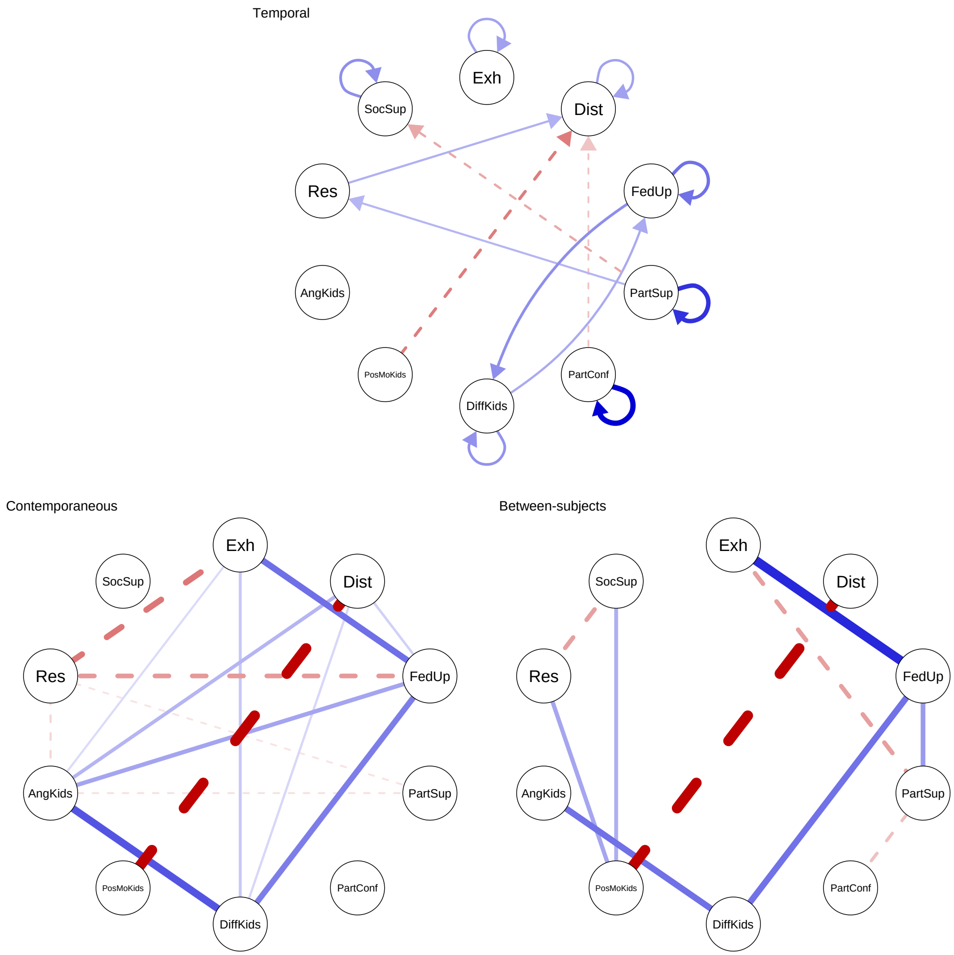


Figure S26. Network Analyses Including Partner-Related Nodes (n=30 Parents with Coparent)

Note. Solid blue edges represent positive associations, while dashed red lines represent negative associations. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PartConf = partner conflict, PartSup = partner support, PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support.

For the contemporaneous network, all thick edges are the same between Figure 19 (all mothers) and Figure S26 (all coparents). There are some very thin edges that are different between these two networks, however: Resources is no longer positively connected with Positive Moments with Kids, but is now (very weakly) negatively connected to Partner Support, which in turn is weakly and negatively connected to Anger toward Kids. For the between-subjects network, all thick edges from Figure 19 are present in Figure S26. There are new edges present, however, some of which are not directly connected to partner-related variables: Social Support is now positively connected to Positive Moments with Kids, which is positively connected to Resources, which is negatively connected to Social Support.

Overall, adding in partner-related variables into the network analyses impacts the temporal and the between-subjects network, and in particular seems to impact how other variables interact with Resources and with Positive Moments with Kids. These changes could be due to taking into account partner conflict and partner support (i.e., these two variables explain additional variance), but also could be due to a limited sample and limited statistical power (which also meant the model was estimated differently, with orthogonal, or non-correlated, random effects). Another explanation could also be related to the specificity of the coparenting relationships: each coparenting relationship will be distinct (potentially more so than each individual differences in how parents experience exhaustion, or how they interact with their children), and so perhaps 30 parents is too limited to capture group-level commonalities amongst how parents in burnout perceive support and conflict with their partner.

Sensitivity Analysis #3: Adding variable "Time with Kids" to full sample

We also wanted to examine how the network structure would change when including the variable "Time Spent with Kids" in the network (i.e., controlling for it). This network was constructed with all parents (n=40) and included 9 nodes (the 8 from the main network analysis in Figure 19 as well as Time with Kids; no partner-related variables were included as the sample for this network included single parents).

Across all three networks (see Figure S27), adding Time with Kids did not lead to any major changes in the structure of the networks. For the temporal network, all edges were thinner and the autoregressive association for Time with Kids was the only very thick edge (i.e., the self-predicting effect for Time with Kids was the strongest predictor in the network). The only changes for the structure involved Positive Moments with Kids: Time with Kids positively predicted Positive Moments with Kids, while the autoregressive association on Positive Moments with Kids was now missing. This suggests that the amount of hours spent with kids does predict the amount of shared positive moments with their children the next day, but not any other variables (e.g., getting angry with children, finding children difficult to manage, feeling exhausted, distant, fed up).

For the contemporaneous network, however, Time with Kids was associated with more variables in the network (while all thick edges from Figure 19 are still present in Figure S26): Time with Kids still strongly positive connected to Positive Moments with Kids, but also to Finding Kids Difficult to Manage, and thinly to Fed Up and Exhausted. This means that within one day, the more time parents spent with their children was associated with not only sharing positive moments with their children but also with finding children difficult to manage, feeling fed up, and feeling emotionally exhausted.

For the between-subjects network, the results in Figure S27 are almost identical to Figure 19 in the main analyses, with the exception of a small negative edge between Time with Kids and Social Support. This implies that on average across the eight weeks, the time parents spent with their children did not impact their affective experience as a parent, nor their interactions with their children.

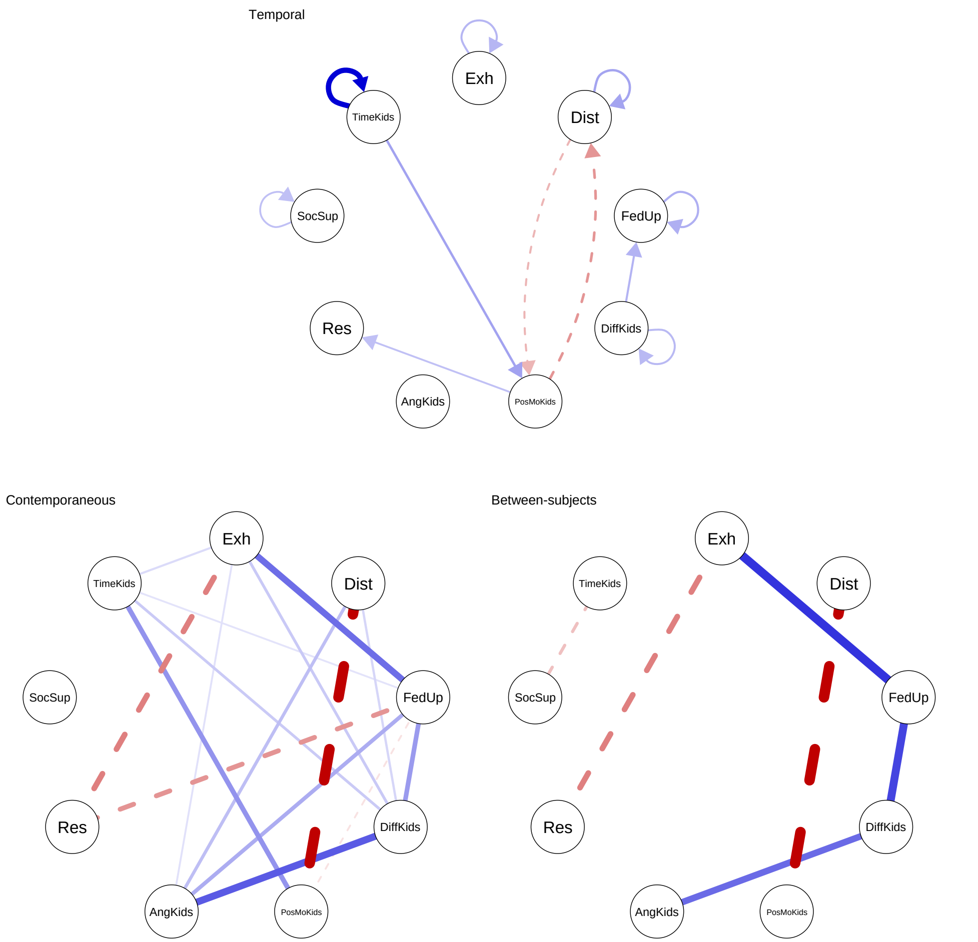


Figure S27. Network Analyses including "Time with Kids" Variables (with n=40 Parents)

Note. Solid blue edges represent positive associations, while dashed red lines represent negative associations. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support, TimeKids = time spent with kids (in hours).

F.6 Centrality Analyses

We estimated strength centrality (i.e., an index of how connected each node is to other variables). There are general concerns with interpreting centrality indices for psychology networks (Bringmann et al., 2019), but many researchers still report and interpret centrality indices for symptoms networks, as symptoms that are highly connected to other symptoms are typically seen as maintaining the network (McNally, 2016). The networks estimated in this study do not include only symptoms, however—indeed, they include mostly contextual variables, including some that are not always purely "negative" or "positive" but depend heavily on the context (e.g., getting angry at kids). As such, we estimated centrality indices just for the sake of completeness and only report them here in the supplementary materials.

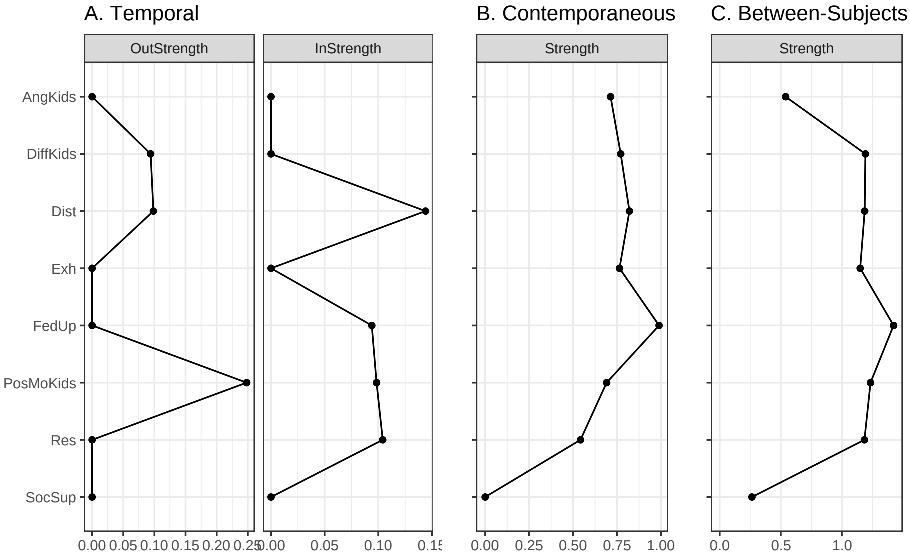


Figure S28. Raw Strength Centrality Indices for all Networks (n=36 mothers)

Note. AngKids = getting angry toward children; DiffKids = finding children difficult to manage; Dist = emotional distance; Exh = emotional exhaustion; FedUp = feeling fed up; PosMoKids = sharing positive moments with children; Res = resources; SocialSupp = social support.

Stability Analyses for Centrality Indices

As we estimated centrality indices, even if only in the supplementary materials (see Figure S28), we also wanted to assess the stability of these estimates (as we also did in Blanchard et al., 2023), following the method and code of Jongeneel et al. (2020). We calculated 200 networks that included a randomly sampled 80% of participants (i.e., 38 participants) and computed centrality indices for these networks (strength for the contemporaneous and between-subject networks, and in- and out-strength for the temporal networks). We then correlated the centrality indices of the original model with the 100 randomly sampled models using Spearman rank correlations and plotted these correlations in a histogram (see Figure S29).

Note that since this sample is already limited, we included all participants (including fathers and single parents, and so n=40). We estimated orthogonal random effects for these 200 networks (and not correlated random effects, as in the main manuscript), both for time and due to the limited sample size. Even so, not all models converged (i.e., 2 out of the 200), likely due to the limited sample size (e.g., 80% of 40 so 32), and likely other networks in the estimated 200 were not adequately powered.

When examining Figure S29, we can conclude that centrality estimates were very stable for the contemporaneous network, relatively stable for the between-subjects network, and not very stable for the temporal network (although better for out-strength than in-strength centrality).

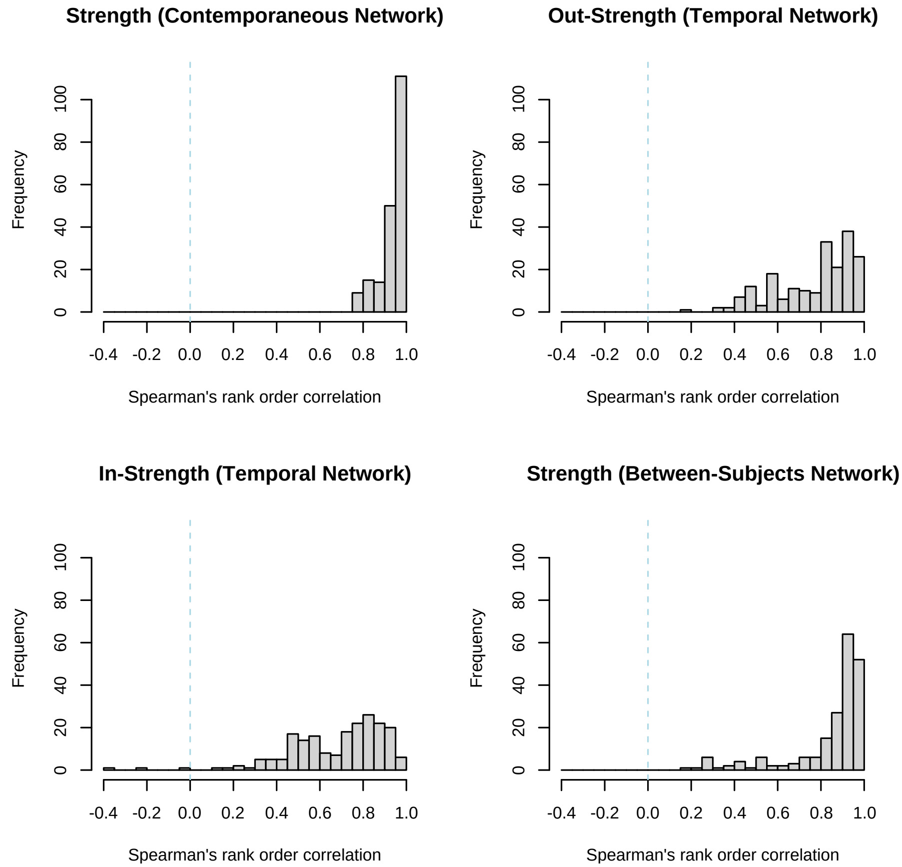


Figure S29. Correlations between case-dropped network and original (n=40) network

Appendix G:
Supplementary Materials from Chapter 9

These supplementary materials can also be found at: https://osf.io/xc5rd.

Note that all R code and materials in this document (e.g., tables and figures) are shared on the Open Science Framework: https://osf.io/5nfbu/.

G.1 Additional Information: Samples

US Sample: Demographic Information

All parents had at least some college education, and more particularly: 10 participants had some college education, 33 had a college degree, 47 a master's degree, and 14 a doctoral degree. Regarding employment, 14 parents had no job, 12 had one part-time job, 71 had a full-time job, two parents were on parental leave, and five parents answered "other." Regarding household income, 13 parents reported a household income below the US household median (i.e., 70,000 USD) while 88 reported a household income at or above the median; additionally, one parent's income was unknown and two preferred not to answer. Regarding public assistance, ten parents received public assistance currently or had in the past year, eight parents had received public assistance in the past (over one year ago), and 86 parents had never received public assistance. Regarding ethnicity, four parents chose Hispanic/Latino, 99 chose not Hispanic/Latino, and one parent preferred not to answer. Lastly, regarding race, 101 parents chose a single race, and three parents chose more than one race. Across all options, two parents selected American Indian/Alaskan Native, 12 selected Asian/Pacific Islander, 9 selected Black/African American, 83 White, and 1 other (self-described as Middle Eastern).

The two Belgian samples include demographic information relating to the age of parents and children in their associated articles (Blanchard, Hoebeke, et al., 2023b, 2023a), but so include similar information on the US sample included in this study here. Parents in the US sample had a mean age of 33.24 years (ranging from 24 to 45 years), while their children had a mean age of 16.79 months (ranging from 12 to 24 months). In addition, families in the study had an average of 1.62 children (ranging from one to five), and an average of 2.06 adults in the household (ranging from one to five, with a medium of two).

US Sample: Inclusion into Present Study

By the time data was downloaded for inclusion in this study, the team in the US had recruited 146 parents for their Baby EMU study, with data collection spanning from February 2023 to July 2023. However, 24 of these participants were excluded as they did not finish the study, did not complete measures required for this study (i.e., PBA or daily questions) or failed the attention checks embedded in the PBA. An additional 18 participants were excluded because they answered less than 50% of daily items (i.e., fewer than 11 out of 21), as preregistered; this yielded 104 participants.

G.2 Affective Indices: Distributions of Remaining Indices

The distribution of the affective indices not included in the main manuscript (i.e., ICC and residual SD) can be found in Figure S30 Figure S31. As can be seen, the distributions for all indices vary by sample (US vs Belgium).

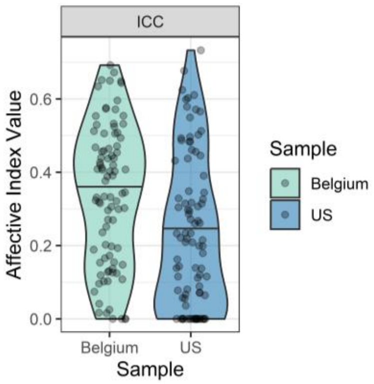


Figure S30. Distribution of ICC (i.e., Covariation Index) by Sample

Note. ICC = Intraclass correlation coefficient.

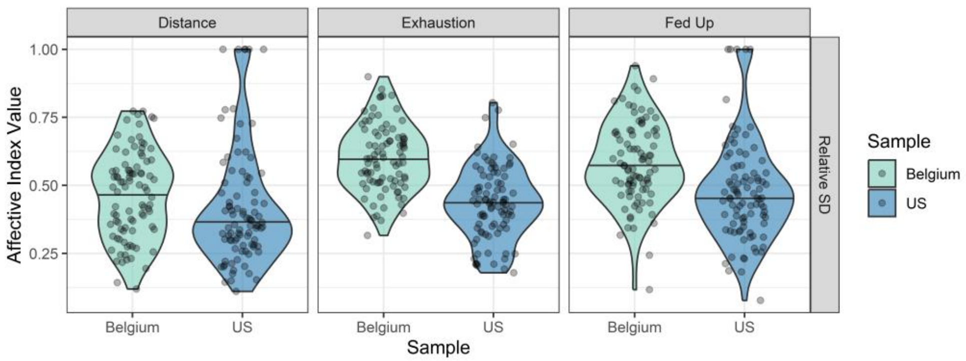


Figure S31. Distribution of Residual Standard Deviation (i.e., Variation Index) by Sample

G.3 Affective Indices: Correlation Plots with Parental Burnout Assessment

Correlation plots from all affective indices with parental burnout severity can be seen in Figure S32 - Figure S34. As can be seen in Figure S32, the mean levels and standard deviations of all three parental affect variables are significantly correlated with parental burnout severity (when examining with a simple bivariate correlation), while relative SD is not.

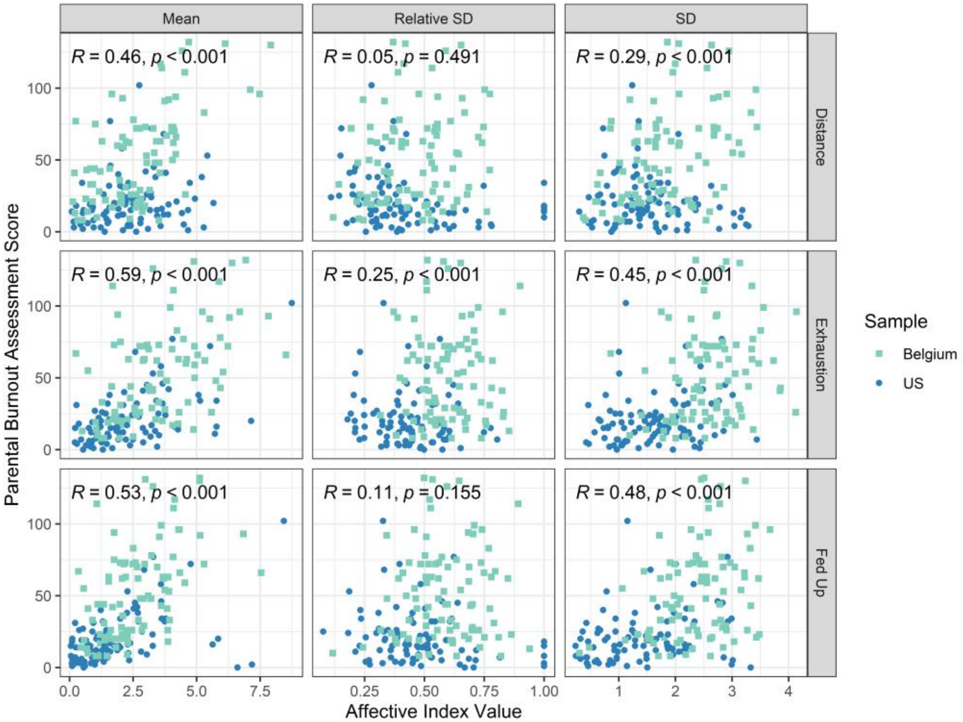


Figure S32. Correlation of Mean, SD and Relative SD with Parental Burnout Severity

When examining the correlation plots of all inertia (i.e., autoregressive) indices in Figure S33, all indices were significantly associated with Parental Burnout Assessment scores, apart from inertia estimates from the simple autoregressive model (specifically for Distance and Exhaustion). These plots also demonstrate how the distributions per sample are quite different (particularly for mlVAR indices). In Figure S34, ICC estimates are significantly correlated with parental burnout severity.

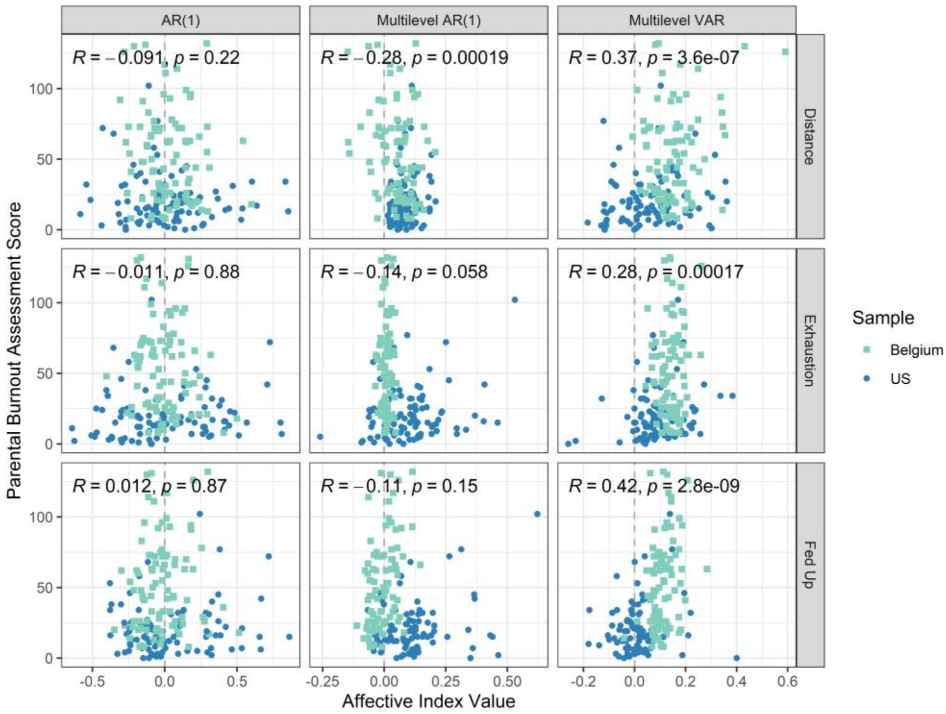


Figure S33. Correlation of all Inertia Indices with Parental Burnout Severity

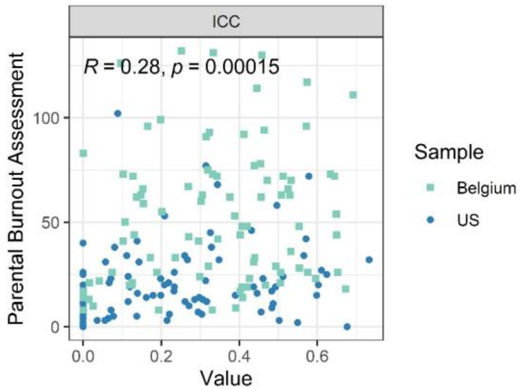


Figure S34.Correlation of ICC (Intraclass Correlation Coefficient) Estimates with Parental Burnout Severity

G.4 Sensitivity Analyses: Further Information

The full results tables for all sensitivity analyses—including the coefficient estimates and p-values for all models—can be found in the online supplementary materials (https://osf.io/5nfbu/).

A visual summary of all sensitivity analyses can be found in Figure S35.

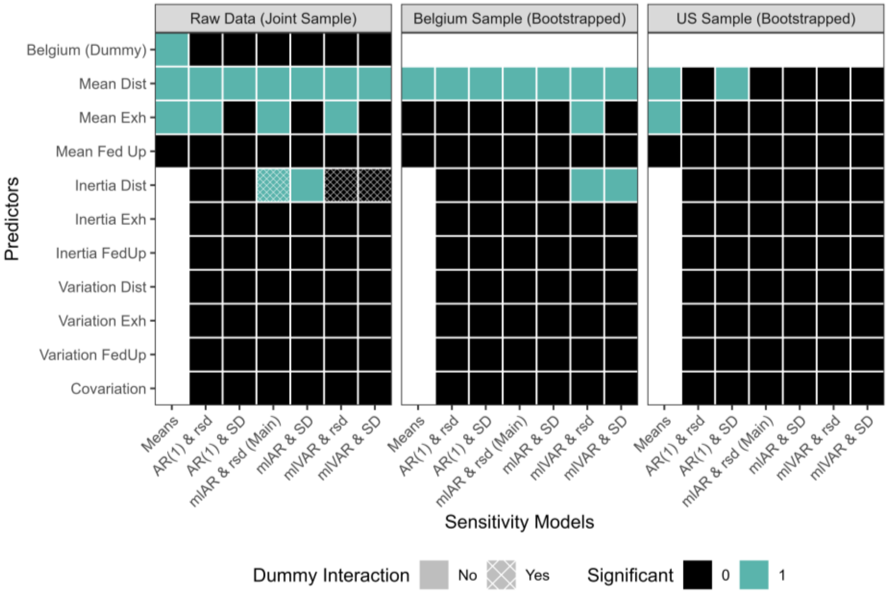


Figure S35. Summary of Significant Predictors Across Sensitivity Models with Raw Data and By Sample

Note. Belgium (Dummy) = dummy predictor indicating sample (Belgium or US); Dist = emotional distance; Exh = emotional exhaustion; FedUp =feeling fed up; AR(1) = autoregressive model; rsd = relative SD; mlAR = autoregressive parameter from a multilevel AR(1) model; mlVAR = multilevel vector autoregressive model.