



Trinity College Dublin

Coláiste na Tríonóide, Baile Átha Cliath

The University of Dublin

Parenting in Protracted Refugee Situations

Qusai Khraisha

Thesis submitted to the Trinity College Dublin, the University of Dublin, for the degree of

Doctor of Philosophy in Psychology

2025

Supervised by: Dr Kristin Hadfield

Interim supervisor: Dr Elizabeth Nixon

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).



Qusai Khraisha

3rd/07/2025

In loving memory of

*my grandmother, Fayzeh,
my aunt, Anaam,
and my childhood friend, Amro,*

who passed away during my PhD

List of publications for thesis

- Khraisha, Q.**, Abujaber, N., Carpenter, S., Crossen, R. J., Kappenberg, J., Kelly, R., Murphy, C., Norton, O., Put, S. M., Schnoebelen, K., Warraitch, A., Roney, S., & Hadfield, K. (2024). Parenting and mental health in protracted refugee situations: A systematic review. *Comprehensive Psychiatry*, 152536.
<https://doi.org/10.1016/j.comppsy.2024.152536>
- Khraisha, Q.**, Put, S., Kappenberg, J., Warraitch, A., & Hadfield, K. (2024). Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. *Research Synthesis Methods*, 15(4), 616–626.
<https://doi.org/10.1002/jrsm.1715>
- Khraisha, Q.**, Sawalha, L., Hadfield, K., Al-Soleiti, M., Dajani, R., & Panter-Brick, C. (2024). Coparenting, mental health, and the pursuit of dignity: A systems-level analysis of refugee father-mother narratives. *Social Science & Medicine*, 340, 116452. <https://doi.org/10.1016/j.socscimed.2023.116452>
- Khraisha, Q.**, Fan, M., Lee, M., Himsi, M., Raguragui, M., Chuhan, C., Put, S., Volta, B., Wadhwa, A., Dos Santos, E., Çakmak, M., Qasim, A., Sulowska, M., El Kharouf, A., von Stumm, S., Marechal, I., Dajani, R., Vining, R., Michalek, J., & Hadfield, K. (under review). OBSESS to measure interactions transcending word meaning. *Child Development*.

Dissemination efforts for thesis

Khraisha, Q. (Chair & Presenting Author), Skinner, A. (Presenting Author), Miller-Graff, L. (Presenting Author), & Betancourt, T. (Discussant). *War, stress, and socialisation: innovative approaches to studying young people and relationships in persistent crisis*. [Symposium session, May 2025]. SRCD Biennial Meeting, Salt Lake City, Utah, USA.

Khraisha, Q. (Presenting Author). *Can LLMs create more inclusive sciences?* [Invited oral presentation, March 2025]. Trinity College Dublin School of Psychology Conference, Ireland.

Khraisha, Q. (2024). How does trauma affect parenting? Parenting in Crisis Situations with Qusai Khraisha, Your Shrink In Bansko (Director). [Video recording]. <https://www.youtube.com/watch?v=III98P4wCCA>

Khraisha, Q. (Presenting Author). *Cascading effects that war trauma and subsequent displacement can have on a child's positive and prosocial behaviour through the intermediary of the mother's interactions*. [Invited oral presentation, Sep 2024]. University College Dublin, Ireland.

Khraisha, Q. (Presenting Author). *Can AI replace humans in systematic reviews?* [Symposium session, September 2024]. Global Evidence Summit Biennial Meeting, Prague, Czech Republic.

Khraisha, Q. (Chair & Presenting Author), Hiott, C. (Presenting Author), Panter-Brick, C. (Presenting Author), & Betancourt, T. (Presenting Author). *Intergenerational pathways to peace: using developmental science to promote prosociality through positive caregiver-child interactions after war* [Symposium session, June 2024]. ISSBD Biennial Meeting, Lisbon, Portugal.

- Khraisha, Q.** (Chair & Presenting Author), Hein, S. (Presenting Author), Begum, T. (Presenting Author), & Bosqui, T. (Presenting Author). *War and caregiving systems: Towards a holistic child development approach in refugee contexts* [Symposium session, June 2024]. ISSBD Biennial Meeting, Lisbon, Portugal.
- Khraisha, Q.** (Presenting Author). *Refugee mother-child dyads' sensitivity and synchrony: A longitudinal observational study*. [Individual oral presentation, May 2024]. Karolinska Institutet – UNICEF Joint Conference on Global Child and Adolescent Mental Health, Stockholm, Sweden.
- Khraisha, Q.** (Presenting Author). *Can large language models replace humans in the systematic review process?* [Invited oral presentation, April 2024]. University of Jordan, Amman, Jordan.
- Khraisha, Q.** (Presenting Author). *Coparenting, mental health, and the pursuit of dignity: A systems-level analysis of refugee father-mother narratives*. [Invited oral presentation, Dec 2023]. International Medical Corps Office and Save the Children, Amman, Jordan.
- Khraisha, Q.** *The bonds that keep us strong* [Ted-style talk, October 2023]. Irish Global Health Network Humanity and Health Conference, Dublin, Ireland. <https://www.youtube.com/watch?v=uLbRFvAPzRQ&t=488s>
- Khraisha, Q.** (Chair & Presenting Author), Eltanamly, H. (Presenting Author), Panter-Brick, C. (Presenting Author), & Miller, K. E. (Presenting Author). *What happens to family dynamics in a protracted refugee situation?* [Symposium session, March 2022]. SRCD Biennial Meeting, Salt Lake City, Utah, USA.
- Khraisha, Q.** (Chair & Discussant), Zolman, N. (Presenting Author), Jung, E. (Presenting Author), & González Ballesteros, L. M. (Presenting Author). (2022, March). *Exploration of caregiver experiences and community-based interventions among*

refugee, internally displaced, and war-impacted communities. [Symposium session, March 2022]. SRCD Biennial Meeting, Salt Lake City, Utah, USA.

Table of Contents

List of Tables	10
List of Figures	10
Acknowledgements.....	11
Summary	13
Chapter 1: Gaps in understanding ‘refugee parenting’	16
1. Narrow conceptualisation of ‘refugee parenting’	17
1.1. Viewing ‘refugee parenting’ solely as ‘refugee mothering’	17
1.2. Viewing ‘refugee parenting’ in a vacuum.....	18
2. Overlooking common displacement contexts.....	21
2.1. Late recognition of the importance of context.....	21
2.2. Bias towards framing displacement as a past event	23
3. Methodological and representational issues	26
3.1. Social desirability in ‘refugee parenting’ research.....	26
3.1. Cost barriers for Global South researchers	28
4. Addressing the three gaps in this thesis	30
Chapter 2: Parenting and mental health in protracted refugee situations: A systematic review.....	32
Abstract.....	33
1. Introduction	35
1. Methods.....	38
2.1. Eligibility Criteria.....	40
2.1.1. Population.....	40
2.1.2. Outcomes.....	42
2.1.3. Exposure	43
2.1.4. Study design	46
2.2. Search strategy	46
2.2.1. Peer-reviewed literature.....	46
2.2.2. Grey literature.....	49
2.2.3. Selection process	51
2.2.4. Use of large language model: GPT-4.....	51
2.2.5. Data extraction.....	52
2.2.6. Quality appraisal.....	52
2. Results.....	53
2.3. Past stressors.....	59
2.3.1. War and unspecified past trauma.....	60
2.3.2. En-route trauma	62
2.4. Ongoing stressors	62
2.4.1. Universal pathways	63
2.4.2. Cultural-specific pathways	65
3. Discussion.....	66
2.5. Limitations.....	70
2.6. Future research directions	72
2.7. Policy and intervention implications.....	74
4. Conclusion	77
Chapter 3: Coparenting, mental health, and the pursuit of dignity: A systems-level analysis of refugee father-mother narratives	79
Abstract.....	80
1. Introduction.....	81
2. Methods.....	85
2.1. Participants	85
2.2. Data collection.....	86
2.3. Thematic analysis.....	88
3. Results.....	92

3.1.	Demographics	92
3.2.	Themes	93
3.2.1.	Negotiation.....	96
3.2.2.	Mirroring.....	99
3.2.3.	Anchoring.....	102
3.2.4.	Transformation.....	105
4.	Discussion.....	109
5.	Conclusion	113
Chapter 4: OBSESS to measure interactions transcending word meaning in refugee mother-child dyads		114
	Abstract.....	115
1.	Introduction.....	116
2.	Methods	121
2.1.	Participants.....	122
2.2.	Study design.....	122
2.3.	Eligibility criteria	123
2.4.	Timeline and recruitment procedure	124
2.5.	Home-visit protocol	124
2.5.1.	Materials and activity structure.....	125
2.6.	Observational coding system	126
2.6.1.	Sub-domains and scoring.....	129
2.6.2.	Design and training	130
2.6.3.	Reliability and validity.....	134
3.	Results.....	136
4.	Discussion.....	142
5.	Conclusion	147
Chapter 5: Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages.....		148
	Abstract.....	149
1.	Introduction.....	150
2.	Methods	152
2.1.	Prompt engineering approach	154
2.2.	Screening and extraction.....	156
2.3.	Analysis metrics.....	157
3.	Results.....	160
4.	Discussion.....	165
5.	Conclusion	168
Chapter 6: Discussion.....		169
1.	Linking temporality with relationality in 'refugee parenting'	169
2.	Understanding inertia and friction in protracted refugee situations	171
3.	OBSESS and LLMs: Towards more inclusive methodological approaches.....	172
4.	Limitations	174
5.	Conclusion	175
References		177
Supplements		240
	Supplementary Table 1: List of protracted refugee situations.....	240
	Supplementary Table 2: Quality appraisal raw results	243
	Supplementary Table 3: Sociodemographic characteristics of the qualitative study	247
	Supplementary Table 4: Demographics of sample agreeing	248
	Supplementary Table 5: OBSESS and self-reported measures with low reliability.....	249
	Supplementary Table 6: Detail of the iterative process to engineer the prompt across the different stages	250
	Supplementary figure 1: Normal Range (a) and Extent of Differences in Coding (b).....	251

List of Tables

Table 1. Studies on Parental Mental Health and Parenting in Protracted Refugee Situations	53
Table 2. Data Exclusion Summary and Inter-rater Reliability Assessment	57
Table 3. Example of dyadic narratives of war-related experiences and responses after forced displacement	89
Table 4. OBSESS Compared to Other Systems Validated on Displaced Mother-Child Dyads	127
Table 5. Performance Evaluation of GPT-4 versus Human Reviewers in Screening and Extraction	161

List of Figures

Figure 1. Country's GDP and Authorship Diversity Among Systematic Reviews.	29
Figure 2. Stages of Systematic Literature Review Process on Parenting and Parental Mental Health in Protracted Refugee Situations.	39
Figure 3. Decision Tree for Inclusion in a Systematic Review on Parenting and Parental Mental Health in Protracted Refugee Situations.	40
Figure 3. Protracted Refugee Situation of Chadians in Sudan.	45
Figure 5. Boolean Search Query Keywords Framework.	47
Figure 6. Systematic Literature Review on Parenting and Parental Mental Health in Protracted Refugee Situations. PRISMA Flow Diagram.	54
Figure 7. Pathways Linking Protracted Refugee Situations with Parenting and Parental Mental Health: A Conceptual Model.	70
Figure 8. Data visualisation of topic overlap from dyadic mother-father narratives.	89
Figure 9. Data visualisation showing how different themes captured different patterns in the textual data.	89
Figure 10. Four dimensions of dyadic father-mother narratives and family caregiving systems, with representative quotes.	94
Figure 11. Examples of coparenting narratives: (a) stabilising family interactions between Faris (father) and Fatma (mother), (b) destabilising family interactions between Fawaz (father) and Fayrooz (mother).	95
Figure 12. Overview of the Initial and Follow-up Data Collection Phases.	121
Figure 13. The Observational Behavioural System Evaluating Sensitivity and Synchrony (OBSESS).	127
Figure 14. Decentralised Coding of Sensitivity and Synchrony in Refugee Mothers: A Visual Framework	133
Figure 15. Distribution of Sensitivity and Synchrony at Initial and Follow-up Data Collection Phases.	138
Figure 16. A Bland-Altman Plot Visually Representing Systematic Variability in Coding.	139
Figure 17. Sensitivity and Synchrony Correlations Across the Same (a) and Different Contexts (b).	141
Figure 17. Correlations of Sensitivity and Synchrony Averaged Across Contexts at Initial Data Collection and Follow-up	142
Figure 19. Flowchart of study steps.	153
Figure 20. Distribution documents by language and type.	154
Figure 21. Optimising Generative Pre-Trained Transformer (GPT)-4 prompts for systematic review tasks for different stages using an iterative approach.	155
Figure 22. Generative Pre-Trained Transformer (GPT)-4 performance in systematic review screening with balanced and imbalanced datasets.	164

Acknowledgements

It took a whole village to raise this ‘thesis baby.’ First and foremost, my heartfelt thanks go to my family, whose unwavering support gave me the courage to start this project and the strength to finish it. And to my childhood friends who shared their own experiences of war and conflict, thank you—this work was inspired by each of you.

Above all, to the refugee parents enduring the relentless challenges of protracted displacement—your resilience and unbreakable spirit inspire me. This work is for you, in honour of the dreams you nurture for a brighter future for your children.

A supervisor shapes the PhD journey, and I was fortunate to have an exceptional one; Dr Kristin Hadfield, thank you, for your responsiveness and the kind approach you consistently shown with your time and attention. Special thanks to my committee, Drs Jean Quigley and Olive Healy, who were steadfast through some challenging times, and to Martin McAndrew, who helped me navigate when I was not sure where to turn. I am also grateful for the support given to me by my interim supervisor, Dr Elizabeth Nixon.

To all the people who directly assist this research, thank you for being part of this journey. I am immensely grateful to the incredible research team who helped me comb through thousands of documents across the web: Dr Nadeen Abujaber, Laila Banerjee, Sadhbh Carpenter, Meltem Cakmak, Evan Carron-Kee, Abigail Coyne, Robert J. Crossen, Caitríona Finlay, Anna Flynn, Ferdia Geary, Jiayue Sun, Johanna Kappenberg, Ronan Kelly, Chuhan Mei, Cameron Murphy, Orla Norton, Anna O’Riain, Stella Roney, Kate Schnoebelen, Irewonuola Sanyaolu, Matthew Walker, Caroline Welch, and Aislinn Wilkins. Together, we took on a behemoth of literature.

To the dream team who became my ‘eyes’ for coding videos that I could not access myself—Simone Brady, Meltem Cakmak, Anisa Qasim, Erica Dos Santos Ribeiro, Katherine Farrington, Ella Catherine Flynn, Madeline Fan, Lucía Heckel Ruiz, Maria

Lee, Marwa Raguragui, Caoimhe O'Regan, Sophie Marleen Put, Fatma Selcen Serin, Sarah Sweeney, Matylda Sulowska, Beatrice Volta, Anisha Wadhwa, and Rebecca Vining—you were invaluable. I am also very grateful of the efforts by Lina Qtaishat for her support with the video data collection.

At the Trinity Centre for Global Health, where I spent most days (and evenings), my gratitude extends to Tooba Akhtar, Pengpeng Cai, Salam Jabbour, Azza Warraitch, Mel Ó Súird, Mohannad Ramadan, Nicole Maiorano, and Charles Zemp (who are all, fingers crossed, soon to be Drs). Thank you for all the wonderful staff who works there, including Fiona Clarke and Helen Farrelly, in addition to Drs Meg Ryan, Felicity Daly, and Frederique Vallieres.

My heartfelt appreciation goes to my fellow Jordanian scholars—Dr Rana Dajani, Lama Sawalha, Dr Amal El Kharouf and Dr Majd Al-Soleiti, and to the other incredible scholars from around the world I have had the chance to discuss my work with: Drs Theresa Betancourt, Tania Bosqui, Mark Cummings, Eric Dubow, Hend Eltanamly, Sophie von Stumm, Sascha Hein, Isabelle Mareschal, Laura Miller-Graff, Kenneth Miller, and Catherine Panter-Brick. Thank you for being such an inspiration along this path.

I would like to further acknowledge my funders for their support: the Society for Research on Child Development and the Provost of Trinity College Dublin. Finally, I am deeply grateful to the people of Jordan, who allowed the research to take place on the ground and to the people of Ireland, who hosted me with such warmth and kindness, you have truly been some of the most welcoming.

To each of you—and to those I may have unintentionally missed—thank you.

Summary

Millions of parents must raise their children in displacement. Though our knowledge of these experiences continues to grow, it is currently undermined by dated approaches relating to how and where ‘refugee parenting’ should be studied.

As outlined in *Chapter 1*, these limited approaches manifest in three significant ways: contextually, the research disproportionately focuses on the unusual 1% of refugees who reach resettlement destinations, systematically overlooking the vast majority living in protracted refugee situations. Conceptually, frameworks to study parenting typically apply individualistic perspectives, viewing parenting solely as ‘mothering’ and disregarding the relational networks central to many refugee families. Methodologically, the literature heavily relies on costly and labour-intensive approaches for data synthesis, while tools designed to gather empirical data on ‘refugee parenting’ are either highly prone to social desirability bias or lack rigorous and appropriate validation, having been mostly developed and established within contexts foreign to most displaced populations.

I take a threefold approach to addressing these gaps: (i) find and redirect attention to research conducted on the 78% of displaced people in protracted refugee situations, (ii) expand conceptual frameworks to include ‘refugee coparenting,’ and (iii) optimise current methodological approaches to data collection and synthesis. In *Chapter 2*, I conduct the first systematic review to explicitly focus on parenting and mental health in protracted refugee situations, analysing nearly 20,000 documents spanning over 70 countries. Studies reviewed indicate that stressors characteristic to populations in protracted refugee situations, such as movement restrictions and employment limitations, have both universal and culturally specific impacts on parents’ mental health, which in turn, shape ‘refugee parenting’ behaviours. In addition to highlighting several methodological

weaknesses, including the limited use of live or video observational data, some evidence suggested that the impacts of protracted refugee situations were mediated by the ‘family system’ in such a way that the response of other family members, including the child, played a role in the parenting behaviour of refugee parents.

What follows in *Chapter 3* is the first dyadic qualitative investigation of ‘refugee coparenting,’ where parenting is approached and enacted collaboratively among refugee mothers and fathers in a protracted refugee situation. I identify four approaches characterising how refugee couples navigated coparenting interactions: negotiation, mirroring, anchoring, and transformation. Specifically, Syrian couples negotiated how to balance responsibilities, sought verbal and non-verbal emotions and behaviours that reflected calm and respect, prioritised family togetherness over education or resettlement opportunities, and adopted gentler parenting approaches to transform intergenerational experiences. These findings illuminate key relational aspects underpinning the ‘family system’ in the context of protracted refugee situations.

Chapter 4 builds on findings from *Chapters 2-3*, particularly the lack of live and video observational studies in the field as confirmed by the systematic review and the emphasis on non-verbal ‘refugee parenting’ behaviours, as highlighted in the theme of mirroring by the qualitative study. I develop and validate the Observational Behavioural System Evaluating Sensitivity and Synchrony (OBSESS) in the first longitudinal study that uses behavioural observation with displaced families. I demonstrate OBSESS’s reliability across multiple tests and time points, and how it captures behaviours that parental self-reports did not fully reflect. Here I also discuss the issue of expensive proprietary observation systems as a significant reason for the limited use of video data in the ‘refugee parenting’ literature, which is why OBSESS is made free and open-access.

I continue addressing the issue of cost in *Chapter 5*, focusing this time on efficient approaches to synthesis of ‘refugee parenting’ research—which is done to address the methodological challenge of managing voluminous and heterogeneous literature that characterised my work in *Chapter 2*. To this end, *Chapter 5* showcases the first evaluation of Generative Pre-trained Transformer 4 (GPT-4), the biggest Large Language Model at the time, against human performance in conducting systematic reviews—based on the review material from the start of the thesis. Findings indicate that while substantial caution should be exercised if LLMs are being used to conduct systematic reviews, they also offer preliminary evidence that, for certain review tasks delivered under specific conditions, LLMs can rival human performance.

Chapter 6 discusses the implications of having more inclusive research tools and the added value of analysing the relational aspect of parenting in protracted refugee situations. Taken together, my thesis advances the field in multiple ways: by systematically and comprehensively identifying literature associated with the most prevalent displacement context; by applying a more expansive relational conceptual framework grounded in the everyday experiences of refugee families; and by enabling more rigorous, efficient, and inclusive research approaches.

Chapter 1: Gaps in understanding ‘refugee parenting’

As far back as four thousand years ago, various ancient scriptures tell of Hajar, an Egyptian woman in Abraham’s household (the common Hebrew patriarch of the Abrahamic religions of Judaism, Christianity, and Islam; Genesis 21:8–21; Quran 14:37). Unable to conceive, Sarah, Abraham’s wife, brought Hajar as his concubine. However, soon after giving birth to Ishmael, Hajar and her son were expelled into the desert with limited supplies. As Hajar and Ishmael wandered in the desert, Ishmael became distressed from thirst, so Hajar placed him under a bush to shield him from the sun and searched for water in the area. This account—whether read as historical record, cultural narrative, or religious scripture—offers one of civilisation’s earliest documented reflections on the challenges of parenting amid forced displacement.

From the time of the Pharaohs to the present day, millions of parents continue to face the challenge of caring for their children under displacement conditions (UNHCR, 2024; for brevity, I am using the word parent, though this term is meant to encompass all types of guardians, including those non-biologically related to the child). Consequently, increased attention has been directed toward understanding and addressing the needs and unique barriers to fostering positive parenting practices or reducing negative ones among displaced populations, particularly refugees (Gillespie et al., 2022). However, although the growth in research has expanded current knowledge, this progress is being undermined by narrow, inefficient and inaccessible approaches relating to how and where ‘refugee parenting’ should be studied. What follows is a discussion of the three areas where these limitations manifest.

1. Narrow conceptualisation of ‘refugee parenting’

1.1. Viewing ‘refugee parenting’ solely as ‘refugee mothering’

‘Refugee parenting’ is viewed and analysed through the mother-as-primary-caregiver model. The push to change this model in the wider parenting literature gained momentum in the 60s and 70s through various social movements (Deslippe, 1996), which also spurred an increase in research focused on fathers, notably spearheaded by Michael Lamb (e.g., Lamb, 1977). Responding to the various societal changes at the time, including increased maternal employment and evolving family structures (Oppenheimer, 1994), Lamb’s work challenged the prevailing ‘maternal assumption’ in developmental psychology (Lamb, 2010). In its beginnings, Lamb’s work demonstrated that fathers could form significant attachments with their children and play unique roles in child development beyond the traditional provider role (Lamb et al., 1985). Lamb’s research continued to follow the changing family dynamics of the era, including rising divorce rates and the growth of dual-income households, which prompted questions about father involvement, parental roles, and their impacts on child development (Lamb et al., 1997). Other notable scholars who expanded our understanding of fathering include Parke and Sawin (1976), who demonstrated that fathers could be equally effective in parenting as mothers, in addition to Pleck, (2010), who helped clarify the distinction between the concept of fatherhood as a parental status in contrast to fathering as a behaviour.

Despite significant progress in challenging maternal-centric research paradigms since the 1970s, research and interventions continued to demonstrate a disproportionate focus on maternal experiences and perspectives throughout the 21st century (Cabrera et al., 2000; M. E. Lamb, 2000; Panter-Brick et al., 2014). Mirroring these broader trends, a large proportion of studies and interventions on refugees predominantly focused on mothers, often overlooking the role of fathers in child-rearing (see van Ee et al., 2013,

who first highlighted this). This narrow focus neglected the contributions of refugee fathers, the dynamics of the parental partnership, and their impact on children's development, emotional security, and socialisation (Panter-Brick et al., 2014). Recent evidence, however, suggests a change in this trend: generally speaking, parenting research is becoming more inclusive to fathers (Schoppe-Sullivan et al., 2013; Yaffe, 2023), and recently more refugee fathers have started being included in the 'refugee parenting' research (Wissö & Bäck-Wiklund, 2021a). Recent data suggests that war-affected fathers may experience higher tension between their ideals of fatherhood (the father they hoped they would become) and their actual fathering, because of the structural circumstances brought by conflict (Klein et al., 2023). Additionally, interventions targeting refugees have started to engage fathers by designing father-inclusive programs and encouraging their active participation alongside mothers (Bunn et al., 2022; K. E. Miller et al., 2023; Miller-Graff & Cummings, 2022); while it was commonly reported that it is difficult to recruit and engage refugee fathers, studies have generally suggested that refugee fathers' involvement can enhance family resilience and mitigate the negative impacts of stressors on children.

1.2. Viewing 'refugee parenting' in a vacuum

Close to the same period in which research on parenting began to include fathers in the general parenting research, criticism grew regarding the ecological validity of much of the child development research being conducted (Hughes & Haynes, 1978; Vaughn et al., 1980). Many studies observed children in controlled laboratory environments that lacked real-world applicability—psychologist Urie Bronfenbrenner famously remarked that the developmental science of the 20th century is the science of the “strange behaviour of children in strange situations with strange adults for the briefest possible periods of time”

(Bronfenbrenner, 1996, p. 5). Following this, Bronfenbrenner developed the ecological systems theory, which posits that child development occurs through complex interactions between multiple environmental layers—such as family, school, community, cultural norms, and political systems (1974, 1977, 1979, 1996, 2006).

The ecological theory uses systems thinking—an approach that invites us to consider factors not in isolation but as part of different connecting parts that each influence each other. This approach has proved especially valuable for studying ‘refugee parenting,’ where displacement creates disruptions across multiple interconnected ecologies, from immediate family environments to broader societal structures (Scharpf, Kaltenbach, et al., 2021). Building on Bronfenbrenner’s work, key theorists such as Michael Ungar and Forma Walsh, best known for examining systematic and family resilience, expanded its practical applications. Ungar (2009) focused on parenting-environment interactions: for instance, introducing the idea of providing children with the “right amount of risk” to build resilience to the real-world environment. More recently, Ungar and colleagues expanded the systems thinking approach to explore ‘refugee parenting’ (Kuru & Ungar, 2021); and currently, Kristin Hadfield, once Ungar’s post-doctoral fellow (Hadfield & Ungar, 2018; Ungar & Hadfield, 2019), along with her colleagues, is building upon systems thinking ideas of resilience to better understand how refugee children cope with and adapt to adversity (Hadfield, Al-Soleiti, et al., 2024; Hadfield, Dajani, et al., 2024; Hadfield et al., 2022). As for Walsh, she employed systems thinking to challenge Western-centric perceptions of what constitutes a family (Walsh, 2006), reconceptualising families as dynamic support networks where multiple adults inevitably take care of the same child (Irace, 2011; Walsh, 2012, 2021).

Expanding the definition of a family to include the “collection of individuals who love and care for [the child]” brings the view that “virtually every child will be

coparented, continuously or episodically, from birth through adolescence” (Irace, 2011, p.1). Such perspectives, along with growing research on the factors that shape parenting dynamics—such as maternal gatekeeping and its impact on father involvement—has contributed to decades of coparenting studies, which is broadly defined as “the ways in which parents interact with each other in their roles as parents and coordinate their parenting” (Campbell, 2022, p. 1). Yet while the importance of “coparenting relationship [as] foundational to young children’s development” (Nandy et al., 2021, p.1) is well-established in scholarship, a notable scarcity of studies exists on this topic within refugee parenting research. Importantly, the gap stems not due to a lack of data about multiple refugee parents for the same child—several ‘refugee parenting’ studies included both father and mother data but did not analyse their parental interactions together (e.g., Bergset and Ulvik, 2021; Smetana and Ahmad, 2018). Methodological challenges in analysing coparenting may be a factor, but it is possible that the gap is reflective of deeper biases toward analysing parenting from an individual perspective, focusing on either ‘mothering’ or ‘fathering’ while overlooking how caregiving involves relationships between multiple adults.

Systems thinking is not only useful for studying the family-level factors influencing parenting behaviour. It is also key for understanding individual-level factors, for example, between parenting practices and parental mental health (Sager et al., 2024). In the context of conflict and displacement, the recognition that parental mental health is important to parenting followed several notable works, including the early empirical works of Kenneth Miller (1996) and Ragnhild Dybdahl (2001), which showed caregiver mental health predicts children’s wellbeing in war-affected populations. Betancourt and Khan (2008) underscored critical connections between parent’s mental health and displacement environment using Bronfenbrenner’s ecological theory, specifically the

“ecological transitions” concept. Catherine Panter-Brick and colleagues provided that poor mental health among conflict-affected parents must be understood within contexts of ongoing, pervasive poverty, social inequality, and persistent violence (Eggerman & Panter-Brick, 2010; Panter-Brick et al., 2009). Thus, this body of evidence was important for understanding that parental psychological symptoms in displacement were not isolated clinical phenomena, but expressions embedded in broader sociopolitical realities.

2. Overlooking common displacement contexts

2.1. Late recognition of the importance of context

A landmark contribution to the study of refugees’ behaviour and mental health was provided by Miller and Rasmussen (2010), who provide a useful framework for studying daily stressors in refugee mental health. Tol et al. (2011) underscored the importance of daily stressors to refugee mental health by indicating that trauma-focused approaches have low priority by practitioners. These insights catalysed several efforts on ‘refugee parenting’ interventions and parental mental health (Tol et al., 2013). Such works, among others, emphasised context and systems thinking, then coalesced into one framework called the ecological model of refugee distress (Miller & Rasmussen, 2017), which represents a cornerstone in this field. In the subsequent year, Fazel and Betancourt (2018, p.1-3), echoed the value of supporting displaced children by focusing on their parent’s mental health, which may be effectively done by addressing the “daily stressors” and the “broader contextual factors such as language barriers [and] poor living conditions.” As more research was conducted, the model proved repeatedly useful in explaining and predicting the mental health and behaviour of refugees, notably in relatively large-scale studies in Lebanon led by Michael Pluess and colleagues (Karam et al., 2019; McEwen et

al., 2023; Popham et al., 2022, 2023) and in several parenting interventions among displaced people (Bunn et al., 2023).

The ecological model of refugee distress addresses how diverse ongoing stressors (legal uncertainties, employment restrictions, housing challenges, discrimination, etc.) affect refugees' overall psychological adjustment and functioning in various life spheres. Meanwhile, the family stress model specifically examines how economic hardship and related stressors impact family dynamics (Conger et al., 1994; Conger & Donnellan, 2007). Originally developed to examine economic pressure on families during the "Iowa farm crisis," the model identifies pathways through which external stressors impact parenting practices and, consequently, child wellbeing (Conger et al., 2010). Although not as widely applied as the ecological model of refugee distress, the family stress model has been used in the 'refugee parenting' literature, such as the study conducted by Sim et al. (2018), which used this model to demonstrate how displacement-related stressors among Syrian refugee parents in Lebanon shape 'refugee parenting' behaviour by affecting the parents' mental health, community relationships, and economic opportunities.

The increasing application of the ecological model of refugee distress and the family stress model in the 'refugee parenting' literature happened at a time when most studies in the field were conducted in resettlement context (Boruszak-Kiziukiewicz & Kmita, 2020; Bunn et al., 2022; Eltanamly et al., 2021; Gillespie et al., 2022; Kouta et al., 2022; Merry et al., 2017; Sangalang & Vang, 2017a; Scharpf, Kaltenbach, et al., 2021; Sousa et al., 2025; Timshel et al., 2017). Merry et al. (2017) exemplified this limitation in a systematic review of qualitative studies on the experiences of parenthood among refugees, asylum-seekers, and undocumented migrants. They found that only three out of 108 studies were conducted in low- or middle-income countries (which do not represent a resettlement context), which is all the more striking when taking into account that less

than one per cent of refugees get resettled annually to a high-income country (UNHCR, 2020).

2.2. Bias towards framing displacement as a past event

It is worth briefly to look back at the roots of the “trauma-focused framework developed in the West” adopted by “researchers trained in Western psychiatry” that Miller and Rasmussen's (2010, p. 10) landmark paper was set to challenge. The framework origins date back to the 17th and 18th centuries, with terms like “nostalgia” or “Da Costa’s syndrome” to “shell shock” in World War I (WWI) and “combat fatigue” in WWII (Jones & Wessely, 2005). Post-Traumatic Stress Disorder (PTSD) was built around these ideas and was later introduced in the 1980s, as a way to explain Vietnam War veterans’ psychological struggles (Young, 1995). Traumatic events in the context of conflict, in other words, were long-assumed to be discrete and bounded in past-events. This is at odds with the continuous trauma exposure common among the large number of unsettled refugees who often live in environments characterised by ongoing violence and insecurity (Eagle Prof. & Kaminer, 2013; Kira et al., 2013).

Assumptions about trauma being confined to the past mirror widely held assumptions about the common nature of forced displacement, which for long was perceived as temporary and stemming from past events (Loescher, 2001). Originally, the Refugee Convention (1951) definition explicitly delineated this assumption, as it was limited to European refugees fleeing WWII events that had ended. Although the (1967) Protocol subsequently removed the temporal and geographical limitations of the Refugee convention (1951), the assumption of temporariness of displacement persisted in international refugee regime structures and policies (Loescher, 2001). This is most evident in the institutional design of the United Nations High Commissioner for Refugees

as a time-limited agency (until 2003, UNHCR mandate had to be renewed every three years). This presumption of temporary displacement, shaped by the European refugee experience post-WWII, has proved increasingly disconnected from realities where protracted refugee situations became the norm rather than the exception (Crisp, 2003). At a time when 90% of all refugee situations were protracted, UNHCR finally acknowledged protracted refugee situations as a significant global phenomenon (UNHCR, 2004).

Protracted refugee situations are those that include “25,000 refugees from the same country [who] have been living in exile for more than five consecutive years” (UNHCR, 2020). These situations are widespread and have existed in more than 70 countries (UNHCR, 2004). Notable examples include three of the world’s largest and most enduring refugee populations: Palestinian refugees, displaced in 1948 across multiple countries including Jordan, Lebanon, Syria, Egypt, Gaza, and the West Bank; Rohingya refugees, who have sought refuge primarily in Bangladesh’s Cox’s Bazar region in the 1970s; and Syrian refugees, who since 2011 have created one of the largest refugee crises in recorded history, with millions displaced across neighbouring countries like Turkey, Lebanon, and Jordan (UNHCR, 2024c). Currently, approximately 78% of refugees live in protracted refugee situations (UNHCR, 2020), and millions of refugee children will spend their entire lives in such situations (UNICEF, 2023).

Though frequently occurring in Low- to Middle-Income Countries (LMICs), protracted refugee situations are nevertheless unique in that they typically emerge from specific state policies and practices regarding refugee populations, including restrictive encampment policies, limitations on freedom of movement, and barriers to formal employment (Aleinikoff & Poellot, 2014). These restrictions may be acceptable limitations during initial emergency phases, but they become increasingly harmful when many years pass, as refugees’ focus shifts from immediate survival to rebuilding

sustainable lives. Since the LMIC designation is strictly dependent on the Gross Domestic Product (GDP), which is a country-level broad economic indicator, it fails to capture the particular legal and political restrictions that may be specific to the refugee population who are in a protracted refugee situation (UNHCR, 2022). Protracted refugee situations can also occur in relatively wealthy host states, including some parts of Europe, thus the phenomenon transcends simple economic categorisation.

Obviously, protracted refugee situations are not identical in every aspect. However, they universally impose restrictions on refugees' right to move and work. As Edward Said, an influential literary theorist, notes when describing the Palestinian refugee population, who are currently living in seven protracted refugee situations (see Supplementary table 1): "these refugees have no passports, they are unable to work, they are unable to travel, you know there is seven million Palestinians in the world today more or less facing the same problems" (Media Education Foundation, 1998).

Even in countries widely praised for progressive refugee policies, the pattern of restriction on refugees persists. Consider Uganda, often regarded as granting refugees the most rights globally (Betts et al., 2019). However progressive the provisions are in Uganda, refugees there remain legally distinct from citizens in many matters relating to work and freedom of movements. For example, while many in a protracted refugee situation are allocated plots in settlements (e.g., Nakivale, Rhino Camp), they are only getting small pieces of land, and many do not get any land at all (Betts et al., 2019). Accordingly to Grześkowiak (2024, p.103), "because the plots of land and the vast majority of all the remaining assistance (including food distribution) are provided in one of the 13 official settlements, freedom of movement is, to a large extent, only illusory." These constraints extend beyond settlements, with urban refugees also facing significant work and mobility restrictions (Monteith et al., 2017).

Decades of theoretical work discussed earlier suggests that the legal and political disadvantages characteristic of protracted refugee situations may infiltrate family dynamics and parenting in refugees. However, relatively little empirical research is conducted in protracted refugee situations while the limited existing work is not being contextualised within these situations. Umbrella and systematic reviews suggest that a majority of refugee research, including studies focusing on parenting, originates from high-income countries, which most commonly do not have a protracted refugee situation (Carpiniello, 2023; Merry et al., 2017). As for the relatively smaller literature on refugees in LMICs, which often is conducted during a protracted refugee situation, its findings are seldom framed in connection to that context, even when the definition accurately applies (S. R. Meyer, Lasater, et al., 2017; Reed et al., 2012). This trend continues in recent studies that target parenting in conflict-affected settings, many of which are conducted in a protracted refugee situation without ever being framed within such context (Sousa et al., 2025).

3. Methodological and representational issues

3.1. Social desirability in ‘refugee parenting’ research

‘Refugee parenting’ research has shown significant strides in developing sophisticated and psychometrically robust survey instruments (e.g., Alleyne-Green et al., 2019; Miller et al., 2025; Puffer et al., 2017), but has yet to fully embrace the methodological complementarity that became standard practice in the general parenting studies after the 1930s (Orlansky, 1949)—specifically, the crucial addition of observational methods to strengthen conclusions made on parenting data. Live and video observational methods superseded earlier exclusive reliance on survey due to issues with social desirability and

self-report biases (Gardner, 2000; Morsbach & Prinz, 2006; Paulhus & Vazire, 2007).

Parents were often found to modify their responses to present themselves as ‘better’ parents, systematically overreporting their positive parenting behaviours, such as patience and consistent discipline, and underreporting their negative behaviours, like yelling or harsh punishment (Tourangeau & Yan, 2007). This issue was especially common in surveys that asked questions that were impressionistic or retrospective (Carey, 1982). However, even meticulously designed surveys sometimes fell short as parents frequently misremembered the frequency and specific details of daily interactions with their children (Rosenman et al., 2011). Surveys have another critical problem—they are ‘contextually blind’ to the crucial situational factors that influenced a parenting behaviour (Brenner & DeLamater, 2016).

A significant reason for the almost exclusive reliance on self-reports is practicality: research samples involving refugees are more challenging to recruit and observe due to logistical constraints (Sulaiman-Hill & Thompson, 2011), and these costs increase even further in conflict settings (Seagle et al., 2020). Once observational data are collected, there is the issue of accessing a standardised coding scheme. Most coding manuals do not cater to the refugee population, and those that do typically lack rigorous if any validation among these groups (Dalgaard et al., 2020; Hein et al., 2020; Lebiger-Vogel et al., 2022; Lembcke et al., 2020; Scharpf et al., 2023; van Ee et al., 2013; Wiepkema, 2013). Developing coding manuals is labour and time intensive, and that is why, in general, coding manuals—particularly the ones that examine attachment and have been previously used on refugees—are typically proprietary, entail substantial licensing fees, and require certification training programs or annual renewal costs for manual usage in addition to mandatory reliability testing fees (Biringen et al., 2014; Crittenden, 1981). These cost barriers create exclusivity in access to observational research methods, and

since these methods are considered the ‘gold standard’ for studying parenting (Hawes & Dadds, 2006), those who can afford using them have greater influence in shaping the field itself.

3.1. Cost barriers for Global South researchers

While progress has continued, having diversity in academia generally, and more so in psychology (Kite et al., 2001), remains an ongoing challenge (Maphalala & Mpofu, 2017). These diversity concerns extend beyond gender to encompass various ethnic and social backgrounds. A particularly pressing issue is the underrepresentation of researchers from the Global South, referring to lower-income regions (Twum-Danso Imoh et al., 2022). This form of imbalance is particularly relevant to the ‘refugee parenting’ research, as most displaced populations reside in low-resource settings (UNHCR, 2024), where researchers from the Global South face substantial barriers to conducting research. Many structural barriers are responsible for this representational imbalance (Twum-Danso Imoh et al., 2022), but one critical yet rarely acknowledged challenge is the inequity in research synthesis capabilities.

The exponential growth in most research areas, which has also been seen in the ‘refugee parenting’ literature (Fiorini et al., 2018; Gillespie et al., 2022), is making synthesising knowledge highly time consuming and labour-intensive (Michelson & Reuter, 2019). As a result, researchers in low-resource institutions may face more barriers to fully synthesise the literature at large or situate their work within the broader academic conversation because they lack the institutional resources to conduct comprehensive reviews. This is best illustrated in Figure 1 below, which shows how a country’s Gross Domestic Product (GDP)—a reflection of its institutional wealth—closely mirrors the number of systematic reviews produced from that country on a given topic. Without

exception, to date, all the systematic reviews related to ‘refugee parenting’ are led by authors from institutions based in North America or Europe (Boruszak-Kiziukiewicz & Kmita, 2020; Bunn et al., 2022; Eltanamly et al., 2021; Gillespie et al., 2022; Kouta et al., 2022; Merry et al., 2017; Sangalang & Vang, 2017a; Scharpf, Kaltenbach, et al., 2021; Sousa et al., 2025; Timshel et al., 2017).

One way to reduce the high cost of systematic reviews is by greater involvement of machine learning approaches to automate labour-intensive tasks (Michelson & Reuter, 2019). Previous evaluations on the use of Artificial Intelligence (AI) in synthesising knowledge showed that they require substantial human oversight and validation (Blaizot et al., 2022). One possible reason for these results is that these early AI tools split text into fixed segments, which may hinder their ability to understand context. Newer large language models (LLMs) based on the transformer technology, such as Generative Pre-Trained Transformer (GPT), may overcome this problem since they capture more contextual information (Vaswani et al., 2017). Whether these new approaches could be used for literature synthesis efforts would not just be about finding a technical solution, but rather a way of ‘levelling the playing field’ by amplifying the voices of researchers from contexts where displacement often happens.

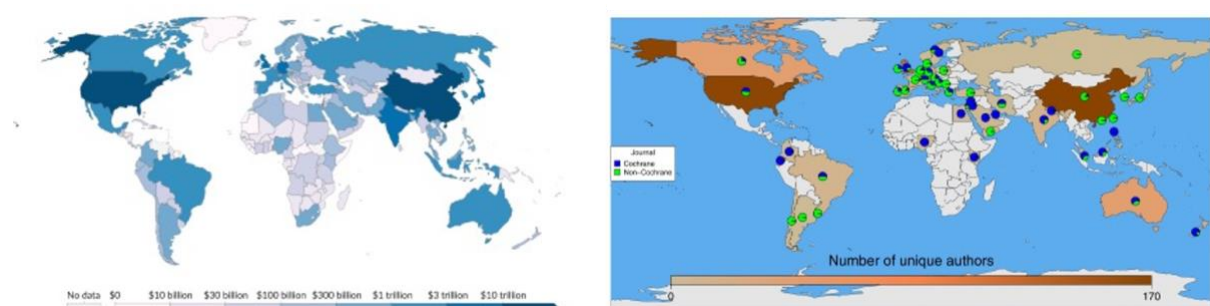


Figure 1. Country's Gross Domestic Product (left) and Authorship Diversity Among Systematic Reviews (right).

Note. The topic of systematic reviews being referenced here is eyes and vision, and data were taken from Qureshi et al., (2020). Data regarding the Gross Domestic Product (GDP) was sourced from ourworldindata.org.

4. Addressing the three gaps in this thesis

I take a threefold approach to addressing these three gaps in ‘refugee parenting’ research:

(i) redirecting attention to research conducted on the 78% of displaced people in protracted refugee situations, (ii) expanding conceptual frameworks to include ‘refugee coparenting,’ and (iii) optimising methodological approaches to data collection and synthesis. In my first study (*Chapter 2*), I systematically identify and synthesise the existing literature on parenting in protracted refugee situations, the most common displacement context. Once this literature is gathered through a systematic review, I will then analyse it using the family stress model (Conger et al., 1994; Conger & Donnellan, 2007), which has proven valuable in understanding how ‘macro’ environmental stressors affect family dynamics and child development across various challenging contexts (Masarik & Conger, 2017).

My next study (*Chapter 3*) will address the narrow conceptual framing issue by conducting a qualitative study examining coparenting to enable exploration of how refugee mothers and fathers navigate parenting responsibilities relationally in protracted refugee situations. Thematic analysis here provides an ideal analytical framework due to its theoretical flexibility (Braun & Clarke, 2006, 2021) and ability to identify shared patterns of meaning while balancing individual experiences with group-level insights (Kiger & Varpio, 2020). This study aligns directly with family systems theory’s emphasis on relationships (Bowen, 1966; Gare, 2000; Minuchin & Fishman, 1981; Walsh, 2006).

The final two studies (*Chapter 4-5*) address methodological limitations. In the first study (*Chapter 4*), I directly address the methodological limitation relating to video observational coding systems by developing and validating the Observational Behavioral System Evaluating Sensitivity and Synchrony (OBSESS). Both sensitivity and synchrony are fundamental concepts in attachment theory and developmental psychology (Deans, 2020; Leclère et al., 2014). They are also relevant to a child's emotional security, self-regulation, social competence, and cognitive growth (Davis et al., 2017; Rodrigues et al., 2021). The final study (*Chapter 5*) addresses the methodological limitation relating to knowledge synthesis in this field, responding to calls to use more machine learning approaches in systematic reviews to make them more efficient (Michelson & Reuter, 2019). In the final *Chapter 6* of my thesis, I discuss how these studies collectively advance our understanding of 'refugee parenting.'

Chapter 2: Parenting and mental health in protracted refugee situations: A systematic review

Khraisha, Q., Abujaber, N., Carpenter, S., Crossen, R. J., Kappenberg, J., Kelly, R.,
Murphy, C., Norton, O., Put, S. M., Schnoebelen, K., Warratch, A., Roney, S., &
Hadfield, K. (2024). Parenting and mental health in protracted refugee situations:
A systematic review. *Comprehensive Psychiatry*, 152536.
<https://doi.org/10.1016/j.comppsy.2024.152536>

This chapter is an exact copy of the published journal paper.

Abstract

Refugees' parenting behaviour is profoundly influenced by their mental health, which is, in turn, influenced by their situation of displacement. Our research presents the first systematic review on parenting and mental health in protracted refugee situations, where 78% of all refugees reside. We pre-registered our protocol and screened documents in 22 languages from 10 electronic databases, reports by 16 international humanitarian organisations and region-specific content from the top 100 websites for each of the 72 countries that 'host' protracted refugees. Our criteria were empirical papers reporting parenting and parental mental health data on refugees who are in a protracted refugee situation. Studies including only internally displaced or stateless persons were excluded. A total of 18,125 documents were screened and 30 studies were included. We identified a universal pathway linking macro-level stressors in protracted refugee situations, such as movement restrictions and documentation issues, to symptoms of depression and anxiety, which, in turn, led to negative parenting practices. Additionally, culture-specific pathways were observed in the way parental mental health and parenting were expressed. Situational (e.g., overcrowding) and relational factors (e.g., spousal dynamics) modulated both of these pathways. Biases in the research included the over-representation of specific protracted refugee situations, overreliance on self-reported data, and a heavy focus on mothers while neglecting fathers and other caregivers. Longitudinal research is needed to clarify the directionality and causality between specific macro-level stressors in a given protracted refugee situation and parental mental health and practices. Refugees were rarely consulted or involved in the design of research concerning their parenting and parental mental health. In recognising the existing links between protracted refugee situations, parental mental health, and parenting, our systematic review calls for a shift in thinking: from focusing solely on the micro aspects that affect 'refugee parenting' to

understanding and tackling the broader macro-level stressors that drive them. We urge larger and long-term research efforts that consider diverse protracted refugee situations, greater investment in science communication and diplomacy with governments, and stronger implementation of durable solutions by states to alleviate the roots of refugee parents' distress and negative parenting practices.

1. Introduction

War and armed conflict have displaced over 120 million people from their homes (UNHCR, 2024c). Given that a substantial proportion of this population are refugee children, with millions more born every year into displacement, ‘refugee parenting’ has become a pressing issue (H. Clark et al., 2020; UNHCR, 2024c). The parenting behaviour of refugees is deeply affected by their mental health (Eggerman & Panter-Brick, 2010; Hein et al., 2020; K. E. Miller, Koppenol-Gonzalez, et al., 2020; Miller et al., 2023), which is further deteriorated by the situation of their displacement (Alisic et al., 2014; Carpinello, 2023; Dattani et al., 2023; Koenen et al., 2017; Mesa-Vieira et al., 2022). Approximately 78% of all refugees are in protracted refugee situations—defined by the United Nations High Commissioner for Refugees (UNHCR) as “25,000 refugees from the same country [who] have been living in exile for more than five consecutive years” (UNHCR, 2020). These situations exist throughout the world and vary in compositions (UNHCR, 2004). Some examples include the 1.5 million Syrian refugees dispersed throughout Lebanon, comprising one-third of the population (UNHCR, 2024c, 2024d), and the 280,000 Somali refugees in Kenya, including nearly 10,000 born in the Dadaab camp to parents who were themselves born in the camp (UNHCR, 2012). While protracted refugee situations are widespread and the subject of many studies (Adugna et al., 2022; Etzold & Fechter, 2022; Jacobs et al., 2022; Mielke & Etzold, 2022; Papatzani et al., 2022; Tobin et al., 2022), no systematic review has yet specifically focused on them, which, by extension, means that parenting and mental health in such contexts have not been systematically reviewed.

Regardless of their size or where they happen, all protracted refugee situations share a common feature: they are difficult to resolve on a macro-scale (Aleinikoff & Poellot, 2014). This difficulty often leads states to implement oppressive measures, which

“typically involve restrictions on refugee movement and employment possibilities and confinement to camps” (UNHCR, 2004). Such broad, systemic-scale policies existing at political, social, and economic dimensions act as macro-stressors that inevitably impact refugee children through the family, the child’s own microsystem (Bronfenbrenner, 1977, 1979, 1996). Studying the factors that influence parental behaviour in such contexts is thus particularly important because parenting plays a key role in moderating and mediating the cascading effects between the macro-level stressors and the child (Bronfenbrenner, 1977, 1979, 1996). Evidence of this macro-micro connection can be found across multiple protracted refugee situations: for instance, protracted Somali refugees in the UK faced housing de-prioritisation by the state, resulting in crowded conditions, which then bred harsher parenting (Hickman, 2008); protracted Syrian refugee parents in Turkey were required by the government to pay special fees that stopped them from enrolling their children in school (Doktorlari, 2019); and protracted Palestinian refugees in occupied territories faced work restrictions, which led to poor living conditions that increased psychological control over their children (Ahmad & Smetana, 2021; Barber, 2001). All these examples show that macro-level factors are closely connected to parenting in protracted refugee situations, but they do not provide a clear understanding of how this relationship may unfold for parents.

This is where parental mental health becomes crucial to consider as a mechanism underlying the link between ‘refugee parenting’ behaviours and macro-stressors in protracted refugee situations (Masarik & Conger, 2017). Converging evidence suggests that the uncertainty created by the inability or unwillingness of host states to secure a durable solution may lead to complex and ambiguous emotional responses in refugee parents (Bentley et al., 2022; Bogic et al., 2015; Côté-Olijnyk et al., 2024; Delilovic et al., 2023; Haas, 2017; Hinchey, Chammaa, et al., 2023; Hvidtfeldt et al., 2022; Kamal &

Killian, 2016; Tatton, 2023). This, in turn, can result in parenting practices that are more about managing a continuous state of liminality (Bryant et al., 2018; Miller et al., 2022; Panter-Brick et al., 2014). However, due to the lack of a systematic review focusing on protracted refugee situations, the ‘refugee parenting’ literature tends to underemphasise the impact macro-level factors have on refugee parents’ mental health, instead frequently extrapolating conclusions from systematic reviews on trauma-exposed parent populations that themselves do not highlight these broader influences (Babcock Fenerci et al., 2016; Barel et al., 2010; Christie et al., 2019; L. R. Cohen et al., 2008; Lovejoy et al., 2000; Wilson & Durbin, 2010). As a result, these systematic reviews have become the foundational knowledge of several interventions serving refugee families by targeting parental health, potentially leading to shortcomings in their conception and implementation (Bunn et al., 2022; Miller & Jordans, 2016; Pedersen et al., 2019; Rubenstein et al., 2020; UNHCR, 2020).

Earlier attempts to integrate the literature on ‘refugee parenting’ have not only ignored the context of displacement, but have also been narrow in their choice of database sources (e.g., none have examined grey literature) and vague in their inclusion criteria (Eltanamly et al., 2021; Merry et al., 2017; Sangalang & Vang, 2017a; van Ee et al., 2016). This methodological flaw meant that previous systematic reviews on the topic have failed to find many studies on refugee parents, and those studies they found lumped refugees together with other groups, such as people whose refugee status ceased, internally displaced persons (IDPs), citizens, and migrants (Eltanamly et al., 2021; Merry et al., 2017; Sangalang & Vang, 2017a; van Ee et al., 2016). Such conflation not only overlooks the unique macro-level stressors faced by many refugees, but also hinders accountability. Host states have a legal responsibility to address protracted refugee situations, a mandate that does not apply to other migrants and non-displaced persons

(Aleinikoff & Poellot, 2014). In line with this responsibility, many problems in protracted refugee situations are tied to state (in)actions (Aleinikoff & Poellot, 2014), meaning that the effective redress of these problems must come from the state itself. This also suggests that ‘refugee parenting’ interventions cannot be only targeting the parents on the individual or the family level, but must also include an element of ‘science diplomacy,’ which involves “providing scientific advice and evidence to inform and support decision-making in policies” (European External Action Service, 2022). Therefore, a systematic review explicitly targeting parenting and mental health in protracted refugee situations is not only urgently needed to improve the ecological validity of ‘refugee parenting’ research and intervention, but also as a tool that would help researchers engage in science diplomacy to redirect policymakers’ attention towards the macro-stressors the majority of refugee families experience.

Our systematic review is the first to comprehensively capture the literature on ‘refugee parenting’ and mental health in protracted refugee situations, employing the most extensive search strategy to date for any systematic review examining any aspects of parenting in adversity. It is also the first to exclusively focus on the refugee parent population. Our aim is twofold: (i) to understand what is known about parenting and mental health in protracted refugee situations to inform future research and (ii) to use the synthesised evidence to inform policy and interventions serving refugee families.

1. Methods

This pre-registered systematic review followed the Joanna Briggs Institute’s guide to mixed methods systematic reviews (Stern et al., 2020), for an overview see Fig. 2. We reported our findings in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Page et al., 2021). All the relevant materials for

this review are available online (<https://osf.io/tuv6q/>) on the Open Science Framework (OSF), with the protocol accessible via the International Prospective Register of Systematic Reviews (PROSPERO) database (CRD42022289404). Note this work is part of a larger systematic review targeting all research on parenting in protracted refugee situations. This paper presents the findings on parental mental health as one of the outcomes of this larger review, while other outcomes will be reported elsewhere.

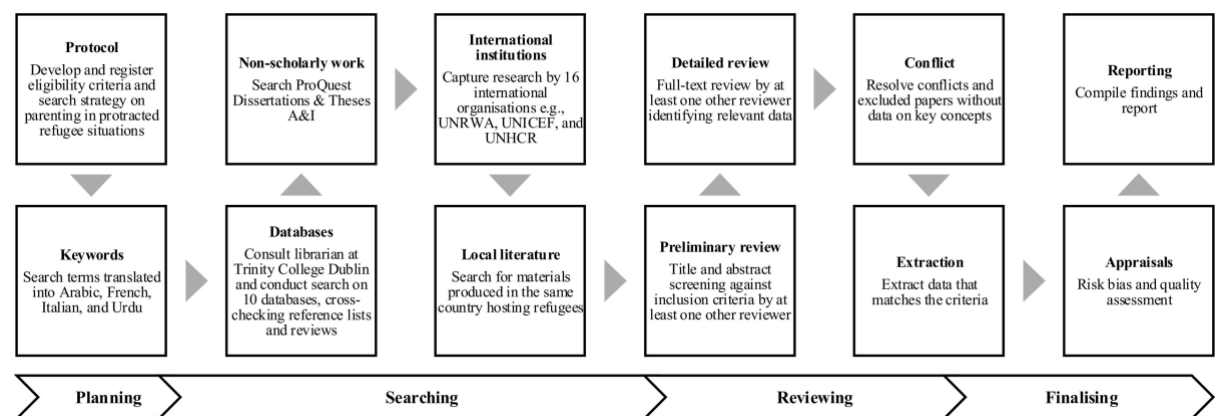


Figure 2. Stages of Systematic Literature Review Process on Parenting and Parental Mental Health in Protracted Refugee Situations.

Note. UNRWA (United Nations Relief and Works Agency for Palestine Refugees in the Near East); UNHCR (United Nations High Commissioner for Refugees); UNICEF (United Nations Children’s Fund); PROSPERO (The International Prospective Register of Systematic Reviews); QuADS (Quality Assessment with Diverse Studies). The diagram represents a systematic literature review workflow in four phases, focusing on ‘refugee parenting’ in protracted situations up to July 2022. It details a strategy that searches ten databases and includes grey literature and international publications across 72 countries, followed by thorough screening and quality checks for data extraction.

2.1. Eligibility Criteria

Our systematic review targeted empirical papers (design) reporting parenting and parental mental health data (outcome) on refugees (population) who are in a protracted refugee situation (exposure). Inclusion of a particular paper occurred when all of these criteria were met together (Fig. 3).

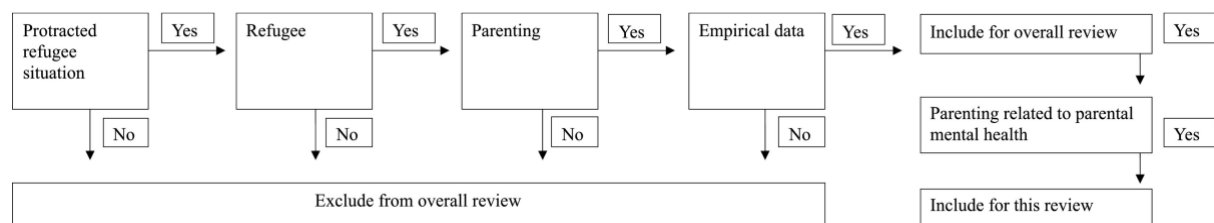


Figure 3. Decision Tree for Inclusion in a Systematic Review on Parenting and Parental Mental Health in Protracted Refugee Situations.

Note. This figure illustrates the decision-making process we took to select studies for our review; the sequence of decisions is an example only, as any of the four primary criteria may have been screened first. Studies had to directly explore observable parenting behaviours, practices, or styles in relation to parental mental health within refugee populations in a protracted refugee situation (living in exile in a host country with 24,000+ refugees of the same nationality displaced for 5+ years). Finally, we only included studies that presented original, empirical data – excluding theoretical papers, protocols, and literature reviews.

2.1.1. Population

We specifically screened for studies providing data on refugee parents. The data needed to speak to participants who were both refugees and parents, but we had separate criteria for each.

a) Refugee

Our target population for this review were refugees defined in accordance with the ‘Refugee Convention’ as persons who “owing to well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group or political opinion, is outside the country of nationality and is unable or, owing to such fear, is unwilling to avail of the protection of that country” (UNHCR, 1951). Studies referring to people in their country of origin as refugees were excluded since a refugee must be “outside of their country.” People currently undergoing the asylum process were considered refugees because these people are de facto self-declared as refugees, and they were only missing an official body’s recognition. These screening criteria are in alignment with the Handbook on Procedures and Criteria for Determining Refugee Status (UNHCR, 2019). Internally displaced and stateless persons were excluded.

b) Parent

Refugees must be parents, defined as any individual ascribed a legal or customary responsibility for caregiving, including biological fathers and mothers, parents living apart, single parents, same-sex parents, parents with adopted or fostered children, stepparents, and other relatives living together in ways outside the ‘nuclear’ formulation (e.g., grandparents, cousins, siblings, or unrelated individuals who are customarily responsible for a child within their family). This definition aligns with how parents are described in the Convention on the Rights of the Child (Convention on the Rights of the Child, 1989). Studies with samples that partly consisted of refugee parents were included as long as the data was disaggregated.

2.1.2. Outcomes

We were interested in parental mental health and parenting as outcomes. Although we had separate criteria for each, the data needed to speak to participants reporting on both parental mental health and parenting.

a) Parenting

Parenting was defined as any directly observable parental behaviours, including overall parenting practices or styles done to a child aged 18 or less (Darling & Steinberg, 1993). Examples of directly observable parenting practices include disciplining a child, showing affection, or engaging in activities that support their children or control them (Darling & Steinberg, 1993; Schaefer, 1959). Parenting styles refer to any combination of these parenting practices (Baumrind, 1966, 1971; Kuppens & Ceulemans, 2019), including authoritative (high demandingness and high responsiveness), authoritarian (high demandingness and low responsiveness), indulgent (low demandingness and high responsiveness), and neglectful (low demandingness and low responsiveness). Although the primary focus of this review is on parenting, the studies did not require child or parent participation (e.g., other stakeholders could report on observed parenting behaviours); it is just that the data itself had to be about parenting. In cases where the word ‘child’ was mentioned without specifying their age, we assumed they were 18 or younger.

b) Parental mental health

We characterised parental mental health broadly in accordance with the World Health Organisation’s (WHO) and Centres for Disease Control and Prevention’s (CDC) definitions as a “state of mental wellbeing that enables people to cope with the stresses of life” (WHO, 2014) which encompasses “emotional, psychological, and social”

functioning (CDC, 2023). These definitions encapsulate experiences of general emotional distress, including fatigue or guilt, and the spectrum of mental health conditions characterised by significant changes in thinking, emotion, and behaviour, such as depression, anxiety disorders, post-traumatic stress disorder (PTSD), bipolar disorder, schizophrenia, and substance abuse (5th ed.; DSM–5; American Psychiatric Association, 2013). Of course, mental health is not merely the absence of mental disorders (WHO, 2014), so we also consider any positive aspects of emotion, such as parental wellbeing. Consistent with the WHO’s and CDC’s definitions above, a one-time emotional reaction did not constitute mental health. Lastly, we did not include attitudinal data that described what parents thought they would feel if their child behaved in a certain way; rather, we focused on parents’ mental health in relation to their past behaviour as parents rather than their stated intentions.

2.1.3. Exposure

We only included studies conducted with parents exposed to protracted refugee situations. Since UNHCR does not publish a comprehensive list of protracted refugee situations (UNHCR, 2004), we followed a systematic process to define and locate them.

a) Defining protracted refugee situations

Protracted refugee situations are defined by UNHCR as “25,000 refugees from the same country [who] have been living in exile for more than five consecutive years” (UNHCR, 2020). While our criteria were largely identical to this definition, we adjusted the sample threshold to no less than 24,000 refugees to account for settings where precise refugee population tracking is challenging, because UNHCR rounds figures to the nearest thousand among these populations (UNHCR, 2007). Note that the definition covers entire

situations rather than individual cases. This means that these situations are not necessarily unique in terms of the length of displacement each refugee person experiences; rather, they are unique in that refugees become part of a large, identifiable group that is unable to access a durable solution.

b) Mapping protracted refugee situations

Refugees cease to be refugees only by either becoming citizens in their asylum country or safely returning to their country of origin (UNHCR, 2020). Our assessment of protracted refugee situations was, therefore, based on statistics for refugee populations that were not able to access voluntary repatriation (choosing to return to their home country) or resettlement (those who obtained citizenship of the asylum country). These were indicated as the total number of refugees registered with either the UNHCR or the local government each year, as obtained from the UNHCR website using the Refugee Data Finder tool (UNHCR, 2021d). Supplementary Table 1 provides a visualised list of all the protracted refugee situations that have existed or remain ongoing to date.

c) Identifying studies conducted in protracted refugee situations

To confirm if a particular study mapped onto the protracted refugee situation list (Supplementary Table 1), we searched for information regarding the refugee country of origin, study geographical location, and the time in which data was collected. For the time of data collection, we first considered relying on the publication year. However, some studies had a long gap between data collection and publication, so we instead only relied on data collection year(s) for accuracy. Data collection dates needed to be either (i) explicitly mentioned or clearly deducible in the studies or, where they were not, (ii) retrievable through contacting the study's authors. We deduced the date only when there

was no ambiguity, such as the cases for studies where the COVID-19 pandemic or other time-bound situations were cited. Fig 4. provides the case of Chadian refugees in Sudan as an example of how we used the UNHCR data to identify protracted refugee situations: they exceeded 24,000 in 1984–1990, again in 1993–1994, and in 2006–2014 (UNHCR, 2021d), which meant that we only included studies that collected data on Chadian refugees in Sudan in 1988–1990 and 2010–2014.

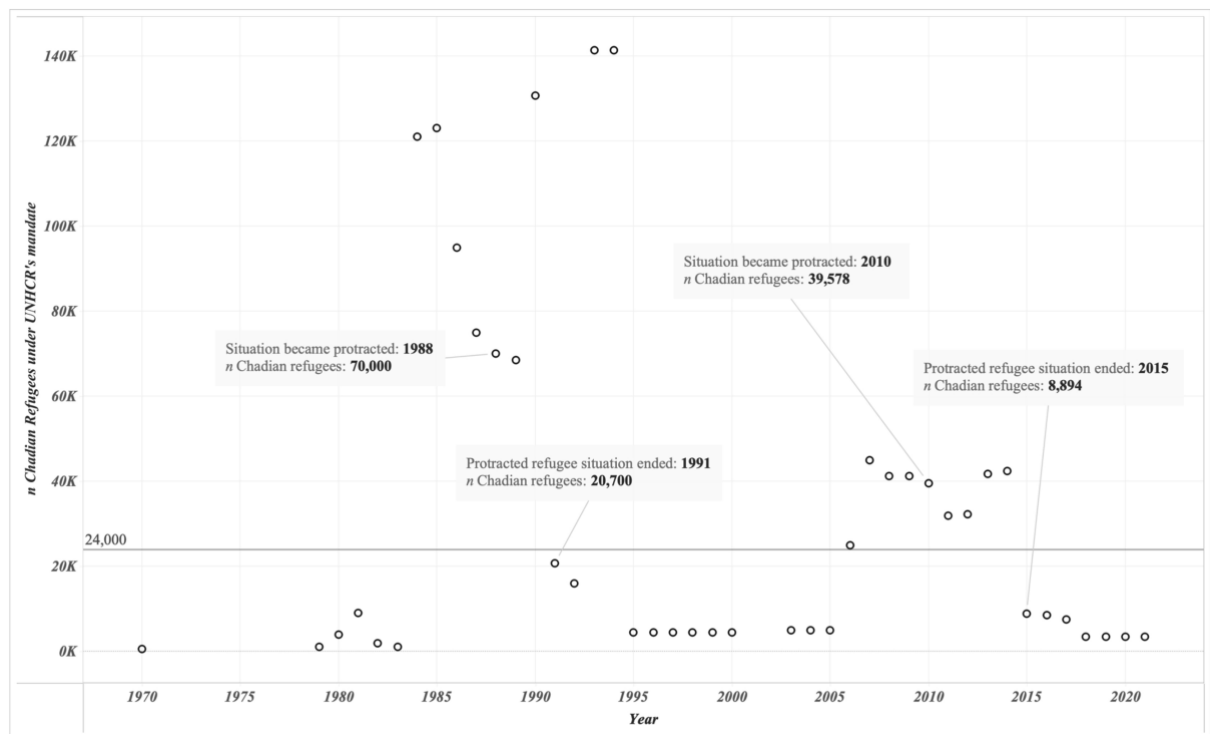


Figure 4. Protracted Refugee Situation of Chadians in Sudan.

Note. This figure provides an example of how studies were included based on UNHCR's definition of protracted displacement in the case of Chadian refugees in Sudan—who exceeded 24,000 in 1984, and again in 1993–1994, and from 2006 to 2014. In that case, our systematic review included studies that had data collected from Chadian refugees in Sudan between 1988 and 1990 and from 2010 to 2014. Only studies collecting data in the 5th year and the subsequent years of a given protracted refugee situation were included. The data was extracted from the United Nations High Commissioner for Refugees

(UNHCR) on the 18th of November 2021 using the Data Finder tool. Data Finder provides population figures for refugees and individuals in refugee-like situations officially recognised by UNHCR.

2.1.4. Study design

We included all documents with empirical data on parenting, regardless of the type, arrangement, or category of that data. However, only baseline results from intervention studies were included, as our goal was to capture ‘organic’ parenting behaviours, practices, and styles. Theoretical papers, protocols, and reviews were excluded. Books and book chapters were included if they reported primary data, along with the date of data collection in the abstract or summary text for the book. We did not state any time limitations on the age of publication in our protocol. However, in practice, included studies could not be published prior to 1951 because, otherwise, the population within these studies would not be considered "refugees" as defined by the 1951 Refugee Convention (UNHCR, 1951).

2.2. Search strategy

2.2.1. Peer-reviewed literature

We searched ten electronic databases for peer-reviewed literature: the Medical Literature Analysis and Retrieval System Online (MEDLINE; via EBSCOhost), the American Psychological Association’s APA PsycINFO (via EBSCOhost) and APA PsycArticles (via EBSCOhost) databases, as well as the Education Resources Information Centre (ERIC; via EBSCOhost), the Cumulative Index to Nursing and Allied Health Literature (CINAHL; via EBSCOhost), the Maternity and Infant Care Database (MIDIRS; via Ovid), Sociological Abstracts (via ProQuest), Embase, the Cochrane Central Register of

Controlled Trials (CENTRAL), and the Web of Science Core Collection. Our search strategy for the electronic databases combined search terms and keywords for the concept of ‘refugee’ and ‘parenting,’ with ‘protracted refugee situations’ being entered as a list of keywords in that same query (Fig. 5).

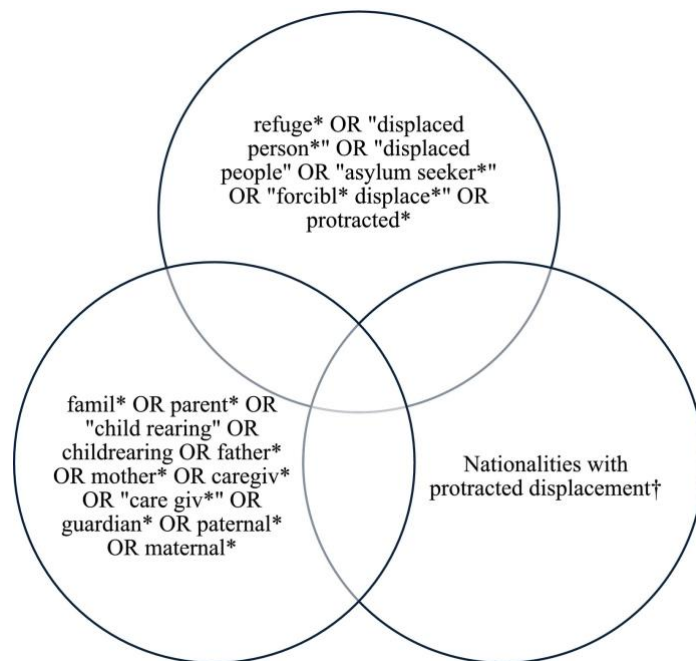


Figure 5. Boolean Search Query Keywords Framework.

Note. †Based on data drawn from UNCHR outlined in Supplementary Table 1: (Afghan* OR Afgan*) OR (Angola* OR Angolese) OR (Armeni* OR Armani OR Armen) OR (Azerbaijan* OR Azerbajdzhan OR Azerbaidjan) OR Bangla* OR (Bhutan* OR Butan OR Boutan) OR (“Bosnia and Herzegovina” OR Bosnia-Herzegovina OR “Bosnia Herzegovina” OR “Bosna Hercegovina” OR Bosnia*) OR Burund* OR (Cambodia* OR Cambogia OR Kampuchea OR Kampuchean) OR (Cameroon OR Cameroun OR Cameron* OR Cameroones) OR (“Central African Republic” OR “Central African” OR “Central Africans”) OR Chad* OR (China OR Chinese) OR Colomb* OR Liberia* OR (Croatia* OR Croatan OR Croat* OR Hrvatska) OR (Congo* OR Zaire) OR Salvador*

OR Eritrea* OR (Ethiop* OR Yaltopya OR Abyssinia) OR (Georgia* OR Sakartvelo) OR Guatemala* OR (Guinea-Bissau OR Bissau-Guinean OR “Guinea Bissau” OR guinea*) OR (Haiti* OR Hayti) OR (Iran* OR Persia*) OR (Iraq* OR Irak*) OR Kazak* OR (Mali* OR “French Sudan”) OR Mauritania* OR (Myanmar* OR Burma* OR Burmese) OR Namibia* OR Nigeria* OR Palestin* OR (Philipp* OR filip* OR pilipinas OR pinoy OR pinay) OR (Russia* OR Rossiya OR russo- OR “Soviet Union”) OR (Rwand* OR Ruanda) OR (Serbia* OR Serb* OR Srbija) OR Kosov* OR “Sierra Leon*” OR Lenoeans OR Salone OR Somali* OR (Sudan* OR Soudan) OR (“Sri Lanka*” OR Ceylon OR Sinhalese OR Singhalese OR Cingalese) OR (Syria* OR Levant) OR (Tajik* OR Tadzhik* OR Tojikiston) OR (Tibet* OR Xizang OR Sitsang) OR (Timor* OR Sitsang OR Maubere) OR (Turk* OR Turco-) OR (Ukrain* OR Ukrayina OR Ukranian) OR (Viet* OR Cochinchina OR Annam OR Tonkin) OR (“Western Sahara*” OR “Sahrawi Arab Democratic Republic” OR Saharawi OR “spanish sahara”) OR (Zimbab* OR Zimbo OR Rhodesia) OR Nicaragua* OR (“Lao People’s Democratic Republic” OR Laos OR Laotian*). The keywords in other languages are available on our PROSPERO page.

The search terms and their synonyms were identified from key articles related to the research questions and were used in combination with the controlled vocabulary (e.g., MeSH term or Subject Headings). The search strategy was slightly adapted for different databases to account for the difference in the interface of databases. We developed and ran the search strategy in English, Arabic, Urdu, French, and Italian. These languages were chosen because every member of our research team involved in the protocol design is fluent in at least one. The search strategy in English was developed in consultation with a subject librarian (GF) at Trinity College Dublin. Search strategies in other languages

were translated from English by the co-authors who were fluent in that language. No filters for language or geographic regions were applied. All the search strategies were prospectively uploaded to the OSF page before conducting the search in databases. In addition to electronic databases, we searched the reference lists of all the review papers on refugees that were captured for additional eligible studies. The reference lists of included papers were downloaded using Scopus and were exported to Covidence for screening. We completed the search in July 2022.

2.2.2. Grey literature

The grey literature search was threefold and aimed to target: (i) non-published scholarly work, (ii) documents published by international organisations working in the humanitarian sector, and (iii) local materials produced by governments in their respective countries. We conducted the searches in this order, and we entered the results into Covidence manually, as the platform did not support grey literature at the time of our review.

a) Non-published dissertations

The electronic database for this component was ProQuest Dissertations & Theses A&I, and it was searched in English, Arabic, Urdu, French, and Italian.

b) Reports by international organisations

We began by identifying the most significant international organisations that research refugee populations based on a previous desk review (M. C. Greene et al., 2017). We then searched the websites of 16 organisations. This search was conducted in English only. This is because, based on preliminary searches (made available on the OSF page:

<https://osf.io/tuv6q/>), we found that such organisations largely publish their reports in English and then translate them into other languages. The 16 websites we searched were as follows: reliefweb.int; refworld.org; mhpss.net; alnap.org; ecoi.net; acaps.org; unhcr.org; unrwa.org; who.org; cvt.org; amnesty.org; hrw.org; icrc.org; iom.org; state.gov; and freedomhouse.org.

a) Data on local websites

Lastly, we targeted grey literature produced by local or national institutions by specifying the web domain of each country we identified to have ever hosted a protracted refugee situation (see Supplementary Table 1). For example, for locally produced material in Jordan—such as documents published by the Jordanian Department of Statistics—we ran our search strategy on Google with the specified country domain (.jo). The country domain name was changed for all 72 countries that host(ed) a protracted refugee situation (Supplementary Table 1). To avoid duplicates, we excluded any websites listed with a different domain that were already included in the list of websites for international sources of grey literature (e.g., we excluded unrwa.jo because we had included unrwa.org). We searched the local material using keywords in English and at least one other official language (if multiple exist) in any of the other four languages we employed for the peer-reviewed search (Arabic, Urdu, French, and Italian). For example, when searching local material in Lebanon, we searched in English and in Arabic; when searching local material in Pakistan, we searched in English and in Urdu, and so on. At least one second reviewer was assigned to conduct a grey literature search on at least one refugee nationality in English and in another language to reduce the risk of bias in the search process. The search record and studies included for all components are available on our OSF page.

2.2.3. Selection process

QK screened all English and Arabic language materials across all literature types and review stages, with AW doing the same for Urdu, NA for French, and BV for Italian. For quality control, a second reviewer randomly assessed 32% of titles/abstracts and 10% of full-text articles in each of our main languages (English, Arabic, Italian, Urdu, and French). Our search captured the following additional languages: Armenian, Chinese, Croatian, Danish, Dutch, German, Hebrew, Iranian, Japanese, Polish, Portuguese, Russian, Serbian, Slovenian, Spanish, Turkish, and Vietnamese. We excluded materials in Croatian, Iranian, Japanese, Serbian, and Vietnamese without translation, as we could not access the full texts. However, members of our research team fluent in the following languages reviewed Chinese, Danish, Dutch, German, Portuguese, Spanish, and Turkish articles (SP reviewed Danish, Dutch, German, and Spanish, whereas CM reviewed Chinese, SC reviewed Spanish and Portuguese, and MC reviewed Turkish). We found full texts for Armenian, Hebrew, Slovenian, Russian, and Polish but lacked native speakers on the team. As a solution, we used AI-assisted translation of documents in these languages. Any conflicts or uncertainties were addressed through discussion and, when necessary, consultation with KH or team members with the necessary language expertise.

2.2.4. Use of large language model: GPT-4

In a separate study, we investigated the capabilities of GPT-4 in screening documents, covering both peer-reviewed publications and grey literature across various languages. Our findings (Khraisha, Put, et al., 2024; last chapter of thesis) indicated that GPT-4 surpassed human performance when supplied with accurate prompts. GPT-4 was used to screen the full texts of 9 studies, which accounted for ~0.001% of total studies screened,

in languages that were not spoken by any team member (Slovenian, Russian, Polish, Hebrew, and Armenian). To further reduce any errors, we translated the documents and viewed them using DeepL (DeepL, 2024) except for Hebrew, which was translated using Google Translate (DeepL did not support Hebrew).

2.2.5. Data extraction

A group of 23 research assistants extracted data from selected articles, with all work re-checked by QK. For robust quality control, a third reviewer independently screened approximately 10% of the studies and QK and the quality appraisal team (detailed below) further reviewed all data. This data included identifiers (title, authors, DOI, publication date, data collection dates), participant details (number of participants, countries of origin/residence, age, gender, education, and whether there are additional non-refugee participants), and study design (cross-sectional, longitudinal, etc.). Information on sampling techniques, types of parenting studied, measurement instruments, and qualitative and quantitative findings specifically related to protracted refugee situations were also captured. Comparisons with other populations were noted if present.

2.2.6. Quality appraisal

We used the Quality Assessment with Diverse Studies (QuADS) tool (Harrison et al., 2021) to evaluate the methodological rigour and reporting quality of included studies. The QuADS tool emphasises a contextualised approach to quality assessment, evaluating studies not just based on a single, general standard, but in relation to their specific aims and contexts. This aligns well with the nature of the diverse literature we captured on parenting within protracted refugee situations, which included studies with various objectives in different settings. Using this tool, studies were appraised by at least two

reviewers trained by QK, who also reviewed all findings. Training included a series of seminars on various research designs and sampling techniques, including hands-on practice on several studies and group discussions conducted online and in-person (example recordings of training can be found on OSF). Quality appraisal included—among other things—evaluation of the strength of the theoretical framing, research setting descriptions, sample appropriateness, and how stakeholder perspectives were incorporated in research design and instrument development. In line with our protocol, we did not exclude studies based on their quality scores but incorporated insights drawn from the quality assessment throughout the results and discussion sections (for more details, see Supplementary Table 2, and our PROSPERO page, CRD42022289404).

2. Results

We identified 18,125 non-duplicated papers in various languages. After screening, a total of 30 records were included in the review: 24 peer-reviewed studies and 6 works of grey literature (Fig. 6). Where data was obtainable, the most common exclusion reasons were a lack of data specifically on parenting (802) and data being collected from non-protracted refugee situations (554). We obtained excellent agreement between reviewers (Cohen's $Kappa = 0.89$, adjusted for imbalance = 0.90) across both peer-reviewed and grey literature (see Table 2 for a detailed breakdown of exclusion reasons and agreement scores).

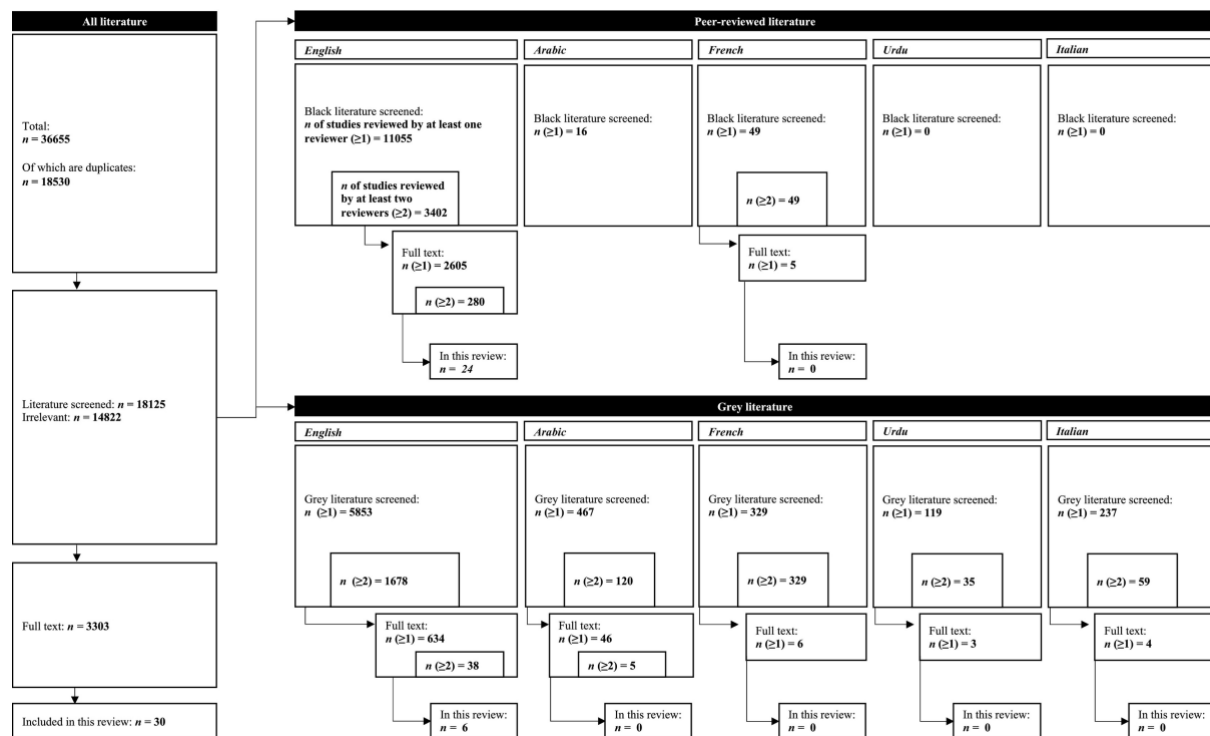


Figure 6. Systematic Literature Review on Parenting and Parental Mental Health in Protracted Refugee Situations. PRISMA Flow Diagram.

Note. This figure illustrates the screening process to identify relevant literature on parenting and mental health in protracted refugee situations. We developed and ran the search strategy in English, Arabic, Urdu, French, and Italian. However, our search in these five languages also captured the following additional languages: Armenian, Chinese, Croatian, Danish, Dutch, German, Hebrew, Iranian, Japanese, Polish, Portuguese, Russian, Serbian, Slovenian, Spanish, Turkish, and Vietnamese. Only one non-English study was relevant to our review, written in German.

In terms of our findings, our systematic literature review highlighted that past and ongoing stressors in protracted refugee situations play distinct roles in affecting refugee parents' mental health and their ability to parent effectively. We start by examining research on refugee parents' war experiences and other past stressors, and then move on

to discuss the literature's findings on ongoing stressors. Table 1 provides an overview of the studies included and the methodologies used in each.

Table 1

Studies on Parental Mental Health and Parenting in Protracted Refugee Situations

Paper	Origin	Residence	<i>n</i> = <i>parent</i> (<i>n</i> = <i>female</i>)	Instrument(s) of measure	Publication status	Quality ratings
Adfam (2009)	Somalia	United Kingdom	23† (15†)	Focus groups with parents	Grey literature	72%
Afifi et al. (2016)	Palestine	Lebanon	40 (40)	Semi-structured interviews	Peer- reviewed	85%
Akesson & Sousa (2020)	Syria	Lebanon	80 (44)	Semi-structured interviews	Peer- reviewed	69%
Akhter & Kusakabe (2014)	Myanmar	Bangladesh	70† (40†)	Open-ended interviews	Peer- reviewed	54%
Al-Natour et al. (2022)	Syria	Jordan	16 (16)	Semi-structured interviews	Peer- reviewed	69%
Alnuaimi (2018)	Syria	Jordan	21† (**)	Semi-structured interviews	Thesis	74%
Atallah (2017)	Palestine	West Bank	80† (**)	Semi-structured interviews	Peer- reviewed	95%
Babuç (2022)	Syria	Turkey	13† (9†)	Semi-structured interviews	Peer- reviewed	74%
Bellamy et al. (2017)	Syria	Turkey & Jordan	106† (50†)	Semi-structured interviews	Grey literature	62%

Bunn (2022)	Syria	Jordan	30* (18†)	Semi-structured interviews	Thesis	97%
Doktorlari (2019)	Syria	Turkey	2 (1)	Focus groups with parents	Grey literature	46%
Fellmeth et al. (2018)	Myanmar	Thailand	9 (9)	Semi-structured interviews	Peer-reviewed	74%
Hecker et al. (2022); Scharpf et al. (2021)	Burundi	Tanzania	452 (226)	Structured clinical interview	Peer-reviewed	79%; 90%
Ingabire & Richters (2020)	Congo	Rwanda	9‡ (6‡)	Open-ended interviews	Peer-reviewed	62%
Jabbar & Zaza (2019)	Iraq	Jordan	12 (12)	Fieldwork observation; survey; semi-structured interviews	Peer-reviewed	77%
Lembcke et al. (2020)	Syria & Iraq	Germany	44 (44)	Video observation: survey; biomarkers	Peer-reviewed	82%
Lewando Hundt et al. (2004)	Afghanistan	Germany	—	Focus groups with children	Peer-reviewed	46%
Lewek & Naber (2017)	Afghanistan & Syria	Germany	24 (**)	Semi-structured interviews	Grey literature	46%
Meyer et al. (2013)	Burma	Thailand	—	Focus groups with key informants	Peer-reviewed	77%
Meyer et al. (2017)	DRC & South Sudan	Rwanda & Uganda	—	Survey with children	Peer-reviewed	95%

Miller et al. (2020)	Syria	Lebanon	144 ‡ (72)	Focus groups with parents	Peer- reviewed	90%
NRC (2015)	Palestine	Gaza	123 (113)	Focus groups with parents	Peer- reviewed	49%
Puffer et al. (2017)	Burma	Thailand	513 (426)	Survey	Peer- reviewed	100%
Rhee et al. (2012)	Vietnam	U.S.	111 (46†)	Case reports	Peer- reviewed	87%
Sim et al. (2018, 2019)	Syria	Lebanon	291 (291)	Survey	Peer- reviewed	82%; 85%;75
Syam et al. (2019)	Syria	Lebanon	28 (13§)	Semi-structured interviews	Peer- reviewed	69%
Yoosefi Lebni et al. (2022)	Afghanistan	Iran	30	Semi-structured interviews	Peer- reviewed	77%

DRC; Democratic Republic of the Congo. NRC; Norwegian Refugee Council.

* Data was deduced as the exact number was not explicitly mentioned.

** Data cannot be determined.

† Parents were included in the sample, but it is not clear if all the sample were parents.

§ Report inconsistent within the same study.

‡ Part of the multi-stage sampling process.

Note. This table provides descriptive data, such as sample size and type of instrument(s) used to capture parenting and mental health, of the studies included in the systematic review on parental mental health and parenting in protracted refugee situations. We used the Quality Assessment with Diverse Studies (QuADS) tool (Harrison et al., 2021) to evaluate the methodological rigour and reporting quality of included studies. As the

QuADS tool emphasises a contextualised approach to quality assessment, the total scores reflect the extent to which methodology and reporting quality were appropriate to the specific aims of the study.

Table 2

Data Exclusion Summary and Inter-rater Reliability Assessment

Category	Total	English peer-reviewed (grey literature)	Arabic peer-reviewed (grey literature)	Instrument(s) of French peer-reviewed (grey literature)	Urdu peer-reviewed (grey literature)	Italian peer-reviewed (grey literature)
Data Unobtainable	1127	857 (257)	0 (12)	0 (0)	0 (0)	0 (1)
No Parenting Data	802	659 (114)	0 (23)	1 (2)	0 (1)	0 (2)
No Empirical Data	328	255 (65)	0 (6)	1 (0)	0 (1)	0 (0)
No Refugees	226	175 (44)	0 (4)	2 (0)	0 (1)	0 (0)
Not Protracted	554	429 (118)	0 (1)	1 (4)	0 (0)	0 (1)
After the Cut-off Date	1	1 (0)	0 (0)	0 (0)	0 (0)	0 (0)
No parental mental health	235	205 (30)	0 (0)	0 (0)	0 (0)	0 (0)

Cohen Kappa	.89	.76 (.84)	—	—	—	—
Adjusted Kappa	.90	.89 (.95)	—	—	—	—

Note. This table highlights reasons for exclusions and the consistency of reviewer agreement. We used Cohen's kappa (κ) to measure the agreement between reviewers. We also employed PABAK (prevalence-adjusted bias-adjusted kappa) to measure the agreement more accurately (Byrt et al., 1993). Due to small sample sizes in other languages, to calculate agreement, we lumped all the peer-reviewed non-English literature in one group ($\kappa = .48$, adjusted $\kappa = .94$) and the all the grey non-English literature in another group ($\kappa = .63$, adjusted $\kappa = .86$).

2.3. Past stressors

The research in this section paints a complex picture of how past trauma and stress experienced by refugee parents influence parenting and parental mental health in protracted refugee situations. Although parents generally reported that past stressors play a direct role in shaping their mental health in protracted refugee situations, these stressors did not consistently influence their parenting behaviours (Hecker et al., 2022; Lembcke et al., 2020; Scharpf, Mkinga, et al., 2021; A. Sim, Bowes, et al., 2018; A. Sim et al., 2019; A. Sim, Fazel, et al., 2018). That said, the evidence surrounding this relationship was limited, and, at times, discrepancies arose between parent and child reports, suggesting a risk of bias (Hecker et al., 2022). Parenting measures also lacked validation and studies relied mostly on cross-sectional, self-reported data (Hecker et al., 2022; Lembcke et al.,

2020; Scharpf, Mkinga, et al., 2021; A. Sim, Bowes, et al., 2018; A. Sim et al., 2019; A. Sim, Fazel, et al., 2018).

2.3.1. War and unspecified past trauma

War trauma and its influence on parental mental health and parenting in a protracted refugee situation was explicitly researched only in one cohort of Syrian refugees in Lebanon (A. Sim, Bowes, et al., 2018; A. Sim et al., 2019; A. Sim, Fazel, et al., 2018). In that sample, war trauma was correlated with increased depression and anxiety symptoms among parents (A. Sim, Bowes, et al., 2018). These symptoms were predictive of harsher punishment methods but not of parental rejection (A. Sim et al., 2019). While war exposure strongly predicted maternal posttraumatic symptoms, these symptoms did not directly lead to harsh parenting or parental rejection (A. Sim, Bowes, et al., 2018). This was further supported by the qualitative data narrated by the same participants (A. Sim, Fazel, et al., 2018), which indicated that they mostly perceived daily stressors, rather than past war exposure, as defining their ‘refugee parenting’ practices. While this work does not speak to the role of the child, other research on protracted Syrian refugees in Germany found that posttraumatic symptoms in children, as reported by the mother, diminish positive maternal affection, as noted through live observation (Lembcke et al., 2020). Similar to earlier, the mothers’ posttraumatic stress symptoms overall did not correlate with observed maternal affection (Lembcke et al., 2020). However, in cases where the child’s posttraumatic symptoms exceeded those of the mother, reported maternal posttraumatic symptoms were negatively associated with observed maternal affection (Lembcke et al., 2020).

Although not explicitly linking data to war exposure, two additional studies, both using the PTSD Checklist for DSM-5 (PCL-5), were conducted on the same group of

Burundian refugees in Tanzania and reported mixed findings (Hecker et al., 2022; Scharpf, Mkinga, et al., 2021). In the first of these studies, parental PTSD symptoms did not predict changes in parental violence (Hecker et al., 2022), consistent with Syrian refugees in Lebanon (A. Sim, Bowes, et al., 2018). However, parents' perceptions of their children's posttraumatic symptoms predicted a reduction in maternal violence (Hecker et al., 2022), contrasting with findings from Syrian refugees in Germany that showed a decrease in affection (Lembcke et al., 2020). This discrepancy might have arisen from methodological differences: one study used direct observation to measure maternal affection (Lembcke et al., 2020), while the other relied on self-reports to measure maternal violence (Hecker et al., 2022). It may also be true that increased parental neglect in response to children's posttraumatic stress symptoms was common across the two studies (which would explain why both studies reported reductions in parenting styles). Furthermore, among the Burundian refugees in Tanzania, the children's reports contradicted their parents, with fathers' PTSD symptoms predicting children's perceived paternal and maternal violence (Hecker et al., 2022), indicating possible social desirability bias in parental reports or unreported inter-parental conflict.

In the second study involving Burundian refugees in Tanzania, parental PTSD and psychological distress were grouped under one concept of 'parental psychopathology', and it was found that maternal psychopathology initially affected mother-child attachment, which subsequently influenced maternal parenting behaviour (Scharpf, Mkinga, et al., 2021). These findings are difficult to compare with the Syrian refugees in Lebanon where posttraumatic symptoms did not predict maternal behaviours (A. Sim, Bowes, et al., 2018; A. Sim et al., 2019; A. Sim, Fazel, et al., 2018), given the grouping of symptoms, but they appear somewhat contradictory. Finally, the psychopathology of Burundian refugee fathers in Tanzania directly impacted their child's wellbeing but did

not do so through paternal parenting behaviour (Scharpf, Mkinga, et al., 2021), suggesting that there may be different pathways depending on the refugee parent's gender or the particular role they assume.

2.3.2. En-route trauma

In the only study examining the impact of flight on refugee parents' mental health and parenting (Lembcke et al., 2020), it was shown that children who had to flee from danger received significantly less maternal affection than those born post-arrival, although the children sample of the former group was significantly older and very small ($n = 6$). However, one notable strength of this study was its use of non-self-reported data, which included measuring cortisol levels to proxy mental health and employing a live observational tool (i.e., no use of cameras) to assess maternal affection towards the child. Still, the reliability of the observational tool, as measured through the inter-coder consistency, was only marginally better than poor (it had a reliability score of 0.51, where 0.49 is considered poor).

2.4. Ongoing stressors

Across different protracted refugee situations, there was a variety of evidence suggesting that macro-stressors affecting the mental health of refugee parents and their parenting practices manifested in universal and cultural-specific ways, each shaped by relational and situational factors (Akesson & Sousa, 2020, 2020; Al-Natour et al., 2022; Alnuaimi, 2018; Atallah, 2017; Babuç, 2021; Bellamy et al., 2017; Bunn, 2019; Doktorlari, 2019; Fellmeth et al., 2018; Hecker et al., 2022; Ingabire & Richters, 2020; Lewando Hundt et al., 2004; Lewek & Nawer, 2017; Miller, Ghalayini, et al., 2020; NRC, 2015; Puffer et al., 2017; Rhee et al., 2012; Syam et al., 2019; Yoosefi Lebni et al., 2022). Specifically,

restrictive measures and policies universally resulted in depressive symptoms and negative parenting practices, and in some cultures, refugees turned to substance abuse as a coping mechanism to deal with their stress and trauma, while in others, they tended to conceal their emotional distress. Compared to the above section, the evidence in this section was clearer and more plentiful. However, the absence of longitudinal research impeded interpretations about directionality and the lack of large-scale, cross-national studies limited quantitative comparability of ‘refugee parenting’ practices and parental mental health outcomes.

2.4.1. Universal pathways

Macro-level stressors imposed on refugees (restrictions on residency, movement, and documentation) led to parental depression and negative parenting practices among Syrian refugees in Lebanon (Akesson & Sousa, 2020; Syam et al., 2019), Burundian refugees in Tanzania (Hecker et al., 2022) Vietnamese refugees in the United States (Rhee et al., 2012) and Burmese refugees in Thailand (Fellmeth et al., 2018). For Palestinians in Gaza, severe movement restrictions due to the ongoing conflict and subsequent symptoms of hopelessness related to instances of parental neglect (NRC, 2015). An almost identical finding was reported among Congolese refugees in Rwanda, although the Congolese parents’ depression symptoms further prompted suicidal ideation in addition to negative parenting (Ingabire & Richters, 2020). That diverse populations in different protracted refugee situations display a pattern between macro-level stressors, parental depression and negative parenting, suggests that a universal pathway may be at play.

However, this pathway is likely to affect the parenting system as a whole rather than solely the individual parent. This was first noted among Burmese refugee mothers who reported experiencing symptoms of depression when witnessing fathers’ aggression

towards their children (Fellmeth et al., 2018). Equally, refugee Burundian fathers' depressive symptoms and psychosocial impairment correlated with maternal violence towards children (Hecker et al., 2022). Among Syrian refugee families, depression in one parent related to verbal and physical inter-parental conflict—and witnessing domestic violence between parents also increased aggressive behaviour among children, leading to a cycle of violence and emotional trauma within the family (Al-Natour et al., 2022; Syam et al., 2019). These findings are further echoed in Gaza, where mothers' feelings of hopelessness prompted them to seek support from their partners, which then influenced their own parenting capabilities; if their partners were also stressed, the emotional strain was amplified back to the mother, limiting her responsiveness (NRC, 2015). Among Palestinian refugees in the West Bank, children's behaviour contributed back to the feelings of hopelessness among parents (Atallah, 2017). More recently, during the COVID-19 pandemic, restrictions on movement made Syrian and Afghani refugee mothers feel overwhelmed by placing pressures on spousal ties (Babuç, 2022; Yoosefi Lebni et al., 2022). These pressures were exacerbated by isolation from vital community support systems and a lack of legal documentation, resulting in reduced responsiveness and availability towards their children (Babuç, 2022; Yoosefi Lebni et al., 2022).

In many cases, parental depression coincided with parental anxiety; many Syrian refugee parents who reported depression in response to movement restriction measures also reported anxiety over their children's safety, influenced by their living conditions (Akesson & Sousa, 2020; Alnuaimi, 2018; Bellamy et al., 2017). Such parental anxiety frequently led to more restrictive parenting practices (Akesson & Sousa, 2020; Alnuaimi, 2018; Bellamy et al., 2017; Bunn, 2019; Doktorlari, 2019; Lewek & Nawer, 2017), including forcibly marrying the children (Bellamy et al., 2017). Children in Gaza have also reported that limited space within the household increases their parents' stress and

dismissiveness, contributing to harsher parenting styles (Lewando Hundt et al., 2004; A. Sim, Fazel, et al., 2018). Restrictions on employment in Lebanon were linked to the father's stress, which was related to the father's neglect (Miller, Ghalayini, et al., 2020). Nonetheless, there was one study with contradicting findings, where Burmese refugee parents in Thailand reported rarely exhibiting behaviours such as taking anger out on their children or feeling too stressed to spend time with them (Puffer et al., 2017). In that study, no participants were recruited from refugee camps and most of the sample was employed (Puffer et al., 2017). Taken together, these studies suggest that the presence or absence of certain situational factors (e.g., camp overcrowding, lack of employment) could be responsible for the co-occurrence of parental depression and anxiety.

2.4.2. Cultural-specific pathways

A small but important body of evidence suggested that parenting and parental mental health differed among culturally-distinct protracted refugee populations. Parental reports of substance abuse linked to macro-stressors were largely shaped by cultural norms and beliefs among Burundian refugees in Tanzania and Burmese refugees in Thailand (Hecker et al., 2022; Meyer et al., 2013). Specifically, parents in these settings noted that it was culturally common, and sometimes encouraged, to drink alcohol, while also reporting that fathers' alcohol abuse led to paternal violence (Hecker et al., 2022; Meyer et al., 2013). Money spent on alcohol, in turn, exacerbated financial problems, which then led to increased stress that triggered more drinking, creating a vicious cycle that resulted in more parental neglect and abuse (Meyer et al., 2013). Situational factors may have also played a role in shaping this cultural-specific pathway among Congolese refugees in Rwanda and among South Sudanese refugees in Uganda; children observed their parents in overcrowded camps consuming alcohol and then often mimicked this behaviour,

leading to several adverse outcomes, including diminished school attendance, social withdrawal, and aggressive behaviour (Meyer et al., 2017). In the same vein, among protracted Somali refugee fathers who lived in crowded conditions in the United Kingdom, the use of khat (a stimulant that is considered a “traditional social activity among Somali men back home”) was associated with parental neglect (Adfam, 2009).

The absence of parental substance abuse in studies involving Syrian and Palestinian refugees, notwithstanding methodological differences, suggests that there may be cultural pathways linking protracted refugee situations to parenting and mental health. However, such culture-specific pathways may not necessarily just be about the expression of parental mental health; they may also relate to parenting communication styles. For instance, in one study, Syrian refugee mothers in Jordan concealed their distress from their children, which may have been due to a cultural taboo associated with talking about trauma (Jabbar & Zaza, 2019). Palestinian refugee parents displaced to Lebanon, who share the same language as Syrians, were also found to conceal their emotions to cope with chronic uncertainty and try to maintain normalcy, which hindered their ability to communicate with their child (Afifi et al., 2016).

3. Discussion

Our systematic review is the first to comprehensively highlight the existing evidence on refugee parents’ mental health and parenting behaviours in protracted refugee situations, where 78% of all refugees reside (UNHCR, 2023a). The focus on these situations addresses long-standing concerns regarding decontextualised research in the domains of parenting (Cuellar et al., 2015), mental health (Saxena et al., 2006), and refugee studies (Tol et al., 2023). We identified a universal pathway linking protracted refugee situations

with parental depression and anxiety which, in turn, predicted negative parenting styles. This may be because the macro-stressors that refugee parents experience subject them to repeated aversive stimuli beyond their control, potentially inducing a ‘learned helplessness’ behavioural response, which has been long-established to be central in the aetiology of depression (Miller & Seligman, 1975). At the same time, protracted refugee situations are likely to continuously evoke worries about safety and the future, leading to a state of constant vigilance and agitation that reinforces the behavioural response of ‘learned restlessness,’ which is recognised as a key factor in the development of anxiety (Fogle, 1978). Additionally, cultural-specific pathways emerged, manifesting in parental mental health through substance abuse and in parenting behaviours through the concealment of emotional distress. Our findings pertaining to the universality and cultural-specificity of refugee parental mental health symptoms and parenting behaviours align with data in previous systematic reviews (Creech & Misca, 2017; Goodman et al., 2011; Lovejoy et al., 2000; McLeod et al., 2007; van Ee et al., 2016), whereas our results pertaining to the links between protracted refugee situations, parental mental health, and parenting are in support of the family stress model and multiple ecological models of child development and refugee distress (Bronfenbrenner, 1977, 1979, 1996; Masarik & Conger, 2017; Miller & Rasmussen, 2017).

Macro-stressors in protracted refugee situations that affected ‘refugee parenting’ in universal and culture-specific ways were largely characterised as ongoing, rather than stemming from the past (Akesson & Sousa, 2020, 2020; Al-Natour et al., 2022; Alnuaimi, 2018; Atallah, 2017; Babuç, 2021; Bellamy et al., 2017; Bunn, 2019; Doktorlari, 2019; Fellmeth et al., 2018; Hecker et al., 2022; Ingabire & Richters, 2020; Lewando Hundt et al., 2004; Lewek & Nawer, 2017; Miller, Ghalayini, et al., 2020; NRC, 2015; Puffer et al., 2017; Rhee et al., 2012; Syam et al., 2019; Yoosefi Lebni et al., 2022). While the few

studies that investigated how posttraumatic stress symptoms affect parenting in protracted refugee situations (Hecker et al., 2022; Lembcke et al., 2020; Scharpf, Mkinga, et al., 2021; A. Sim, Bowes, et al., 2018; A. Sim et al., 2019; A. Sim, Fazel, et al., 2018) showed that past stressors play a direct role in shaping parental mental health, these stressors did not consistently influence parenting. This aligns with the theoretical framework proposed by Miller and Rasmussen (2017), suggesting that ‘daily stress’ impacts mental health and social relations more than war trauma. This is also consistent with research on non-displaced individuals with prolonged exposure to war in Afghanistan, where chronic poverty and social inequality had a greater impact on family dynamics and social suffering than the direct effects of war (Eggerman & Panter-Brick, 2010).

Several studies pointed out that universal and culture-specific pathways linking ongoing macro-stressors with refugee parental mental health and parenting practices were strongly shaped by relational dynamics (Fellmeth et al., 2018; Hecker et al., 2022; A. Sim et al., 2019). This finding harmonises with several systematic reviews and meta-analyses that show that mothers’ and fathers’ responsiveness (Deneault et al., 2022) and mental health (Postler et al., 2022; Sanders et al., 2014) are closely connected. Yet, while relational dynamics were sometimes associated positively, by buffering parental stress and promoting positive parenting (Babuç, 2022; A. Sim et al., 2019), at other times they had a negative association, as evidenced by the common reporting of parental substance abuse problems in cultures that normalise these substances (Hecker et al., 2022; S. Meyer et al., 2013). Macro-level stressors also affected relational dynamics by catalysing and maintaining long-term shifts in traditional parental roles, which then disrupted parental mental health and parenting (Akhter & Kusakabe, 2014; Çelik & İçduygu, 2019; Ingabire & Richters, 2020). An illustrative example of this in our review is the Rohingya refugee situation in Bangladesh, where camp restrictions on fathers force mothers to become

breadwinners, sometimes resulting in shame and negative parenting styles among fathers (Akhter & Kusakabe, 2014).

Relational aspects are not divorced from the immediate situational factors in which they operate. Such interconnectedness was consistently observed in our systematic review, and was further evident among third-person reports from children and stakeholders, where inter-parental conflicts were commonly observed (Lewando Hundt et al., 2004; S. Meyer et al., 2013; S. R. Meyer, Yu, et al., 2017). In the confined spaces of refugee camps, inter-parental conflict may become not only more common, but also harder to hide, increasing the ‘spill over’ of parental stress to children (Martin et al., 2017). Such an understanding aligns with systems thinking and family systems theories that emphasise the synergy of personal and environmental factors in determining parents’ mental health and parenting behaviour (Ungar, 2010; Walsh, 2021). To bring together the findings from our systematic review, Fig. 7 provides a visual conceptualisation of the links the literature suggests to exist between protracted refugee situations, parental mental and parenting.

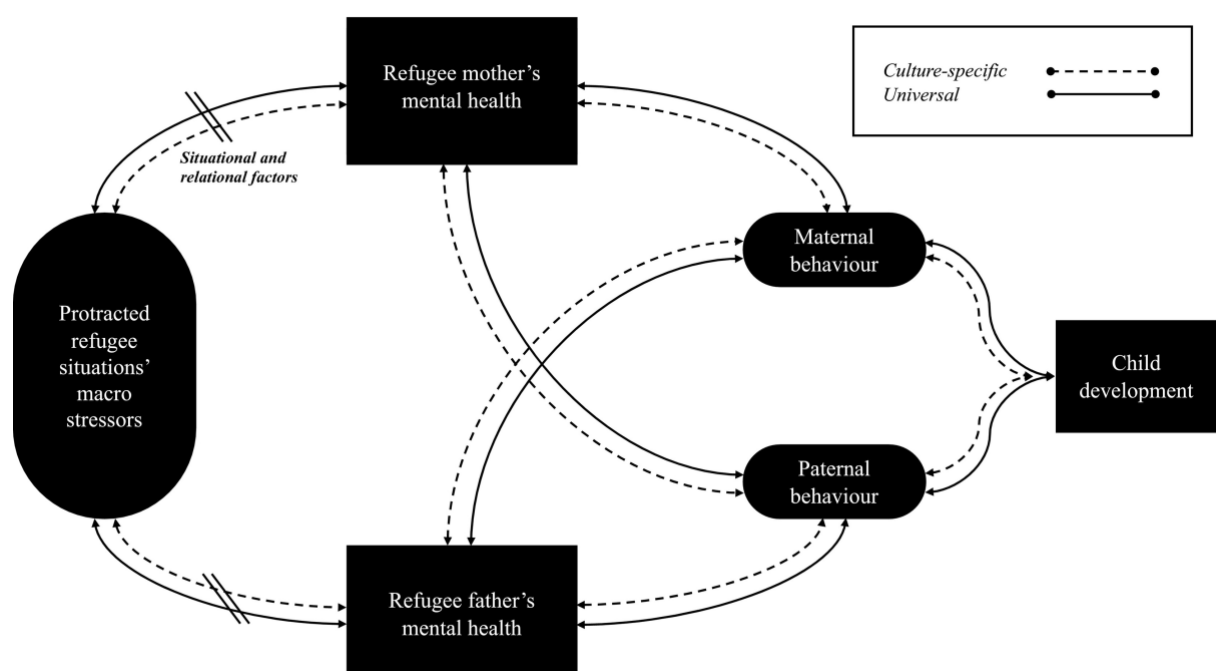


Figure 7. Pathways Linking Protracted Refugee Situations with Parenting and Parental Mental Health: A Conceptual Model.

Note. This figure represents a conceptual model illustrating the impact of protracted refugee situations' macro-stressors on the mental health of refugee parents and the subsequent effects on parenting and child outcomes. Situational and relational factors may play a role in modulating the impact of macro-level stressors on parental mental health and parenting.

2.5. Limitations

Given the diversity of protracted refugee situations and the well-established cultural variability of mental health and parenting practices (Gopalkrishnan, 2018; Lansford, 2022b), it is surprising that refugee parents' understanding of either of these concepts was rarely integrated into research methodologies. Rather, much of the existing research we found employed a top-down understanding of mental health and parenting, often using closed-ended surveys translated without validation, which not only can be unreliable (McLeod et al., 2007), but which also fail to capture culture-specific practices and beliefs (Hinchey, Khalil, et al., 2023). These may include the practice of prioritising firm swaddling of babies (to prevent the 'Moro reflex,' a sign of fright) and avoiding corporal punishment as parental protection from the 'Ceeb,' fright illness in children, seen among protracted 'Hmong' parents from Laos in Thailand (Capps, 1999). Other examples of overlooked cultural practices include 'Naseehath' (religious storytelling) and supervised reading to instil moral values among protracted Afghani refugee children in Pakistan (Sadiq, 2022). While these examples are about refugees originating from East and Central Asia, literature reviews also show that validation issues are particularly prevalent when studying parenting (Ghandour et al., 2024) and mental health (Hinchey, Khalil, et al.,

2023) in the Middle East, where the majority of studies in our systematic review took place.

In our systematic review, all studies except one relied on self-report data (Lembcke et al., 2020). Overreliance on self-report questionnaires, even when validated, introduces a significant possibility of bias, as parents are found to alter their responses to conform to social norms (Hawes & Dadds, 2006; Law & Roy, 2008; Van de Mortel, 2008). Studies elsewhere have also found that stress impairs parents' cognitive functions and emotional stability, affecting the accuracy of their self-reported data (Hayden et al., 2010). Lower socio-economic status and educational levels can influence how parents perceive and respond to questions, potentially leading them to alter their responses to align with perceived social expectations or in hope of receiving additional support (H. M. Feldman et al., 2005; Meisenberg & Williams, 2008; Reese & Read, 2000). Given that protracted refugee situations expose parents to high levels of stress [31,32], and that refugee parents often possess low educational attainment (UNHCR, 2023d), the risk of bias in self-reports in our reviewed studies can be a serious limitation. Indeed, it could explain the discrepancy found between parent and child reports in highly powered studies, such as that of Puffer et al. (Puffer et al., 2017) and Hecker et al. (Hecker et al., 2022).

Self-report bias among the studies in our review is compounded by the fact that most research recruited only mothers, which means that their biases also affected the understanding of refugee fathers and the overall family's functioning and cohesion in protracted refugee situations. This is a frequent issue in parenting studies (Cabrera et al., 2018; R. Feldman, 2023; Panter-Brick et al., 2014), but is particularly common among war-affected populations, where, despite concentrated efforts to incorporate trauma-exposed fathers to understand their mental health, most of that incorporation has entailed

the mothers being asked to report about the fathers (van Ee et al., 2016). In fact, evidence of such bias emerging can be found in our review (Adfam, 2009). Future research should be wary of this, as refugee mothers' self-reports on fathers can skew our understanding of fathers, particularly as refugee couples can substantively disagree about the paternal role (Hadfield, Al-Soleiti, et al., 2024).

It is important to evaluate the methodological rigour of our systematic review in light of the limitations of the literature. Although one strength of our review was comprehensiveness, providing a broad perspective on the topic and allowing for a more credible conclusion on existing knowledge gaps, this came at the cost of speed, and we have not included studies published after July 2022. Second, while our multilingual, multi-source search strategy was designed with the aim of reducing the dominance of English-language perspectives and promoting inclusivity, the language choice still did not fully represent all refugee populations in all protracted refugee situations, and the peer-review literature search strategy specifically targeted electronic databases for publications in English. Third, despite searching a high number of electronic databases, access restrictions due to paywalls and limited institutional subscriptions prevented the inclusion of some potentially relevant non-English studies.

2.6. Future research directions

Future research should explore how parenting and mental health interact across diverse protracted refugee situations (see Supplementary Table 1 for an overview of protracted refugee situations). This should include examining how relational dynamics vary with the geographical context of the refugee situation relative to the country of origin. For instance, a study comparing protracted Vietnamese and Mexican refugee parents in the United States suggests that Mexican parents had better access to social support networks

because they were in proximity to their country of origin (Menjivar, 1997). Another multi-country study on protracted Palestinian refugee parents indicates that parental responsiveness is lower in Gaza than in Jordan or the West Bank (Itani et al., 2017); the unique situation in Gaza suggests that geographic remoteness from violence is another crucial factor. This may be because exposure to violence can diminish the emotional availability of the social network, which has been discussed under the concept of continuous traumatic stress (Eagle Prof. & Kaminer, 2013).

Additionally, future research may examine how relational factors in a given protracted refugee situation may form a cultural interaction effect, where prevailing parental behavioural norms are either weakened or strengthened depending on the host country's norms. Evidence of such an effect may be traced in other literature: for example, studies repeatedly find that Syrian refugee parents are more likely to force child marriages in Jordan and Lebanon compared to Sweden, despite all these countries having laws against child marriage (El Arab & Sagbakken, 2019; Rabo et al., 2021). A potentially powerful approach for understanding the relational dynamics related to the refugee communities that make up a given protracted refugee situation is by leveraging longitudinal designs with network analysis approaches (An et al., 2022; Ogburn et al., 2024).

As many of the studies in our systematic review examined mental health in a broad sense, there is a need for more focused research on specific disorders. Standardised diagnostic assessments may be particularly helpful in informing interventions targeting extremely high-risk refugee parents, as well as in facilitating comparisons across different populations and protracted refugee situations (Meijer, Finkenauer, et al., 2023; Meijer, Franz, et al., 2023). Given that post-traumatic stress disorder (PTSD) is a topic extensively investigated in refugee literature (Bedaso & Duko, 2022; Blackmore, Boyle,

et al., 2020; Mesa-Vieira et al., 2022), yet relatively understudied in relation to ‘refugee parenting,’ future research may prioritise studying this disorder among refugee parents. However, our review suggests that ‘post’ reactions to stress have little or weak impact on parenting in protracted refugee situations. Instead, future research could examine how PTSD moderates symptoms of depression and anxiety, which, in turn, shape ‘refugee parenting’ and parental mental health in protracted refugee situations. This research direction may necessitate a nuanced approach to trauma sub-types; previous studies indicate that specific trauma sub-types, such as death threat trauma, have a stronger predictive validity for mental disorders like PTSD, but are less relevant for conditions like anxiety (Arnetz et al., 2014; Hinchey, Chammaa, et al., 2023).

Finally, our systematic review found only two studies that investigated positive parenting practices (Lembcke et al., 2020; A. Sim et al., 2019), and none that examined positive mental health. Therefore, shifting the traditional ‘negative lens’ to include more research on mental wellbeing and positive parenting, including the possible buffering effects of positive parenting practices on refugee child development (Kira et al., 2012; Thabet et al., 2009), is a crucial future research direction. Such a shift is not only academically relevant but also aligns with the desires of refugee communities, who often wish to highlight their resilience and capabilities rather than just their struggles (Panter-Brick, 2022). Again, this underscores the importance of adopting participatory methodologies, an approach that our systematic review shows to be ignored (see supplementary table 2, supplementary material), yet highly feasible in humanitarian settings (Panter-Brick et al., 2020).

2.7. Policy and intervention implications

The ‘first wave’ of interventions targeting refugee parents and other displaced persons has exclusively focused on delivering parenting knowledge, and, as a result, has been largely

successful in increasing reported parental knowledge about optimum parenting practices (Gillespie et al., 2022). However, these interventions have simultaneously shown little to no effects on refugee children's behaviour and emotions (Gillespie et al., 2022), suggesting a gap between enhanced refugee parental knowledge and its application. One possible reason for this disparity is that refugee parents may feel too overwhelmed by their displacement situations to effectively act on acquired knowledge and appropriately implement nurturing practices throughout their day-to-day lives. This hypothesis is supported by a recent systematic review of mental health interventions targeting refugee families, which found that promoting parental mental health within parenting interventions leads to improved parenting practices and child outcomes (Bunn et al., 2022). Similar observations can be found among single intervention studies conducted in protracted refugee situations (Lakkis et al., 2020; Miller, Koppenol-Gonzalez, et al., 2020; Miller-Graff & Cummings, 2022; Singh et al., 2021). Post-intervention data from these studies suggests that what may be unique to enhancing the effectiveness of 'refugee parenting' interventions in protracted refugee situations is addressing parental depressive symptoms related to macro-stressors (Miller-Graff & Cummings, 2022), integrating cultural variants of parenting and distress (Singh et al., 2021), and strengthening relational dynamics within families (Lakkis et al., 2020), which align with findings discussed in our systematic review.

Still, this 'second wave' of parenting interventions misses the targeting of situational factors, which our data synthesis suggested to play a key role in modulating the impacts of macro-stressors on refugee parental mental health. Interventions conducted in protracted refugee situations do recognise the importance of situational factors, such as overcrowding, in determining parental mental health (Miller, Koppenol-Gonzalez, et al., 2020), but they often fail to directly address them (Bunn et al., 2022). One way they could

address situational factors is by integrating home organisation workshops into the existing intervention program to teach refugee parents how to maximise their limited living space through simple, low-cost solutions like creating storage from recycled materials. People delivering a ‘refugee parenting’ intervention could also be trained on how to conduct a brief housing assessment during the intervention sessions so they could offer more targeted strategies to improve housing conditions. These may include practical advice on mitigating issues like dampness, poor ventilation, or heating by using readily available materials and methods. Whatever approach is taken, addressing situational factors could potentially help ensure that gains in refugee parental mental health and parenting are sustained post intervention delivery (Bryant et al., 2022). It could also have a profound direct impact on children’s physical health (Chaudhuri, 2004), as well as refugee families’ livelihoods, economic inclusion, and physical safety (Shergold et al., 2019; UNHCR, 2018), which can assist in mitigating the continuous negative impact of various macro-stressors that characterise protracted refugee situations.

Finally, current ‘refugee parenting’ interventions often address poor parental mental health by treating it as a symptom of the ‘daily stress’ that refugee parents experience. However, doing so threatens to oversimplify the issue, given that ‘daily stress’ is directly tied to macro-level policies in protracted refugee situations. Narrowly framing it as ‘daily stress,’ therefore, risks obscuring the role policymakers and governmental bodies have in creating, perpetuating, and maintaining restrictive measures that contribute to that stress (UNHCR, 2004). Re-framing ‘daily stressors’ as ‘macro-stressors’ may help underscore the need to not only target parental mental health and parenting, but also the wider policies that negatively impact both—potentially ushering in a new ‘third wave’ of parenting interventions targeting refugee families that, for example, integrate advocacy and legal actions to solve documentation issues. Such an

approach of incorporating policy-level interventions to improving the lives of refugees not only aligns with calls-to-action suggested by prominent scholars in the field (Fazel & Betancourt, 2018), but is also supported by preliminary evidence from our systematic review, which highlights the impact of progressive refugee policies in enhancing refugee parental mental health and parenting. Specifically, our review identified a study indicating higher instances of household violence related to parental substance abuse among Congolese refugees in Rwanda, as opposed to their South Sudanese counterparts in Uganda (S. R. Meyer, Yu, et al., 2017). When considering the sharp contrast in the policies applied to protracted refugee situations between the Ugandan government, more progressive (UNHCR, 2023b), granting refugees unrestricted work rights and freedom of movements; versus the Rwandan government, more restrictive (UNHCR, 2021a), confining refugees to congested camps with limited employment opportunities and movement rights, an argument for the importance of policy-level intervention gains significant traction.

4. Conclusion

We are the first to systematically review parenting and mental health in protracted refugee situations in order to inform research, intervention, and policy. Our findings highlight the need to rethink ‘refugee parenting’ support, moving beyond a narrow focus on negative parenting behaviours and parental mental health to a broader understanding of the underlying contexts from which these behaviours and emotions arise. Negative parental mental health and parenting is not just a standalone problem; it reflects a larger issue of family systems strained by macro-level stressors. While strengthening relational and situational factors can buffer the harmful impacts of these stressors, it is insufficient without structural support and guaranteed rights. States play a fundamental role in

ensuring that refugees' family rights and welfare are safeguarded, as their policies shape parenting and parental mental health in protracted refugee situations. We call for more research to expand our understanding of the situational and relational factors that contribute to family systems resilience, while also urging researchers to invest in science diplomacy efforts with governments. Finally, we call on states to comply with their obligations to implement durable solutions that would alleviate macro-level stressors in ongoing protracted refugee situations.

Chapter 3: Coparenting, mental health, and the pursuit of dignity: A systems-level analysis of refugee father-mother narratives

Khraisha, Q., Sawalha, L., Hadfield, K., Al-Soleiti, M., Dajani, R., & Panter-Brick, C. (2024). Coparenting, mental health, and the pursuit of dignity: A systems-level analysis of refugee father-mother narratives. *Social Science & Medicine*, 340, 116452. <https://doi.org/10.1016/j.socscimed.2023.116452>

In the previous Chapter, studies reviewed revealed that stressors characteristic to populations in protracted refugee situations, such as movement restrictions and employment limitations, have both universal and culturally specific impacts on parents' mental health and parenting behaviours. These impacts on 'refugee parenting' were mediated by the 'family system,' suggesting that the response of other family members plays a significant role in shaping the emotion and behaviour of the refugee parent—this insight therefore informs the subsequent research direction by highlighting a need to investigate family relational interactions as a whole rather than individual parenting behaviours in isolation.

This chapter is an exact copy of the published journal paper.

Abstract

Research on coparenting is virtually absent from the refugee literature, despite its importance for family systems, children's bio-behavioural and emotional development, and intergenerational responses to social change. In 2022, we conducted 30 semi-structured interviews with Syrian refugees in Jordan and used thematic analysis to examine how fathers and mothers ($n = 15$ dyads) enacted parenting together. We identified four approaches characterising how couples navigated coparenting interactions, family cohesion, and intergenerational change. These were negotiation, mirroring, anchoring, and transformation. Specifically, Syrian couples negotiated how to balance responsibilities, sought emotions and behaviours that reflected calm and respect, prioritised family togetherness over education or resettlement opportunities, and, strikingly, adopted gentler parenting approaches to transform intergenerational experiences. Underpinning these four themes were efforts to uphold family dignity. Syrians described themselves as ordinary parents, eschewing the label of refugee parents and building a normal life for their families after war and displacement. Our thematic analysis offers methodological and conceptual advances in exemplifying how to capture a dyadic understanding of coparenting and why refugees strive to parent in ways that sustain mental health and dignity. This systems-level analysis of coparenting in dignity is specifically relevant to strengthening the processes of family-level communication and to designing integrated programs that support caregiving, wellness, and family unity. Our findings lay the groundwork for developing a relational, agentic model of family caregiving systems in the context of precarity and forced displacement.

1. Introduction

Over one hundred million people are now forcibly displaced, accounting for more than one per cent of all humanity (UNHCR, 2023c). Often, refugee families face protracted displacement without a clear path toward voluntary returns, local integration, or resettlement (Assali et al., 2022). Such adversities are typically associated with compromised wellbeing, child development, and family functioning (Arakelyan & Ager, 2021; Miller & Rasmussen, 2017), necessitating systems-level work to foster transformational change (Panter-Brick, 2023; Ungar et al., 2023). Systems-level research goes beyond mapping problems or evaluating intervention efficacy—it asks us to understand how individual, family and community-level factors interlink, calling for developing a transdisciplinary, practice-based research agenda to address systems-related implementation questions (Tol et al., 2023). Key to this discourse is family systems theory, which posits that “serious crises and persistent life challenges have an impact on the whole family, and in turn, key family processes mediate adaptation (or maladaptation) for individual members, their relationships, and the family unit” (Walsh, 2021, p. 256). As argued in this paper, a systemic lens is needed to better understand the relational, agentic nature of family caregiving systems and generational shifts in how refugees are parenting their children.

It is often thought that parenting difficulties stem from caregiver misunderstandings of the effects of specific parenting behaviours, such as harsh discipline, on child wellbeing (e.g., Bjørknes et al., 2012). Consequently, many interventions have specifically targeted the skills and knowledge of refugee caregivers (Gillespie et al., 2022). However, recent literature has seen a paradigm shift—from a deficit model, attributing poor caregiving practices to poor skills and inadequate knowledge—to a stress-compromised parenting model that emphasises the role of poor

mental health and chronic stress in inhibiting the acquisition or implementation of effective parenting skills (F. L. Brown et al., 2022; Miller, Koppenol-Gonzalez, et al., 2020). Discussion of stress-compromised parenting initially emerged from studies that explicitly documented the links between caregiver mental health and parenting behaviours in war-affected and refugee settings (Bryant et al., 2018a; Panter-Brick et al., 2014; A. Sim, Fazel, et al., 2018). Subsequent research highlighted that understanding the intersection of caregiver mental health and parenting skills is critical to improving early childhood experiences, which set the stage for human bio-behavioural systems underlying the development of love, attachment, social cohesion, and peace (Leckman et al., 2021). Recognising such links, a new wave of interventions is now seeking to improve *both* caregiver mental health and parenting skills (Jensen et al., 2021; Miller, Koppenol-Gonzalez, et al., 2020; A. L. Sim et al., 2021).

However, research and interventions on ‘refugee parenting’ remain predominantly focused on mothers (e.g., Bryant et al., 2018; Sim, Bowes, et al., 2018), especially when evaluating impacts on children’s mental health (Sangalang & Vang, 2017b). Only a handful of studies have sought to demonstrate how fathers are consequential to children’s health and development in refugee contexts (Bergnehr, 2022; Bulford et al., 2022; Hadfield, Al-Soleiti, et al., 2024; Vaage et al., 2011; van Ee et al., 2013; Wissö & Bäck-Wiklund, 2021b). Systematic reviews of family-based interventions show that, in refugee settings, no interventions have explicitly focused on coparenting (Bunn et al., 2022; Panter-Brick et al., 2014; Pedersen et al., 2019b), while refugee studies that have included father and mother data (e.g., Bergset & Ulvik, 2021; Smetana & Ahmad, 2018) have not analysed these dyadically. This reflects pervasive biases towards an analysis of (co)parenting as ‘mothering’ which overlooks the importance of father-child relationships in human development and parenting interventions (Cabrera et al., 2018; R. Feldman,

2023; Panter-Brick et al., 2014). It also leaves unnoticed how fathers and mothers negotiate parenting together—which hinders a relational understanding of family caregiving systems. Research gaps on the relational dimensions of caregiving are surprising - and critical - considering the decades of work that have established the importance of coparenting, defined as “the ways in which parents interact with each other in their roles as parents and coordinate their parenting” (Campbell, 2022, p. 1). In sum, research on how coparenting is approached and enacted dyadically is vital to understand, yet virtually absent from the refugee literature.

Also relatively uncommon, yet much needed in the social sciences, are bottom-up examinations of how migrants and refugees strive for a dignified life under conditions of precarity (Barber et al., 2014; Kamruzzaman et al., 2022; Panter-Brick, 2021; Willen, 2019). In war-torn Afghanistan and border areas of Pakistan, longitudinal research with parent-child dyads highlighted that household levels of unity and harmony impacted both caregiver and child mental health; Afghan narratives showed that family unity was the bedrock of personal and social dignity (Eggerman & Panter-Brick, 2010; Panter-Brick, 2021, 2023). For Syrians, a “crisis of dignity” is at the heart of lived experiences: the revolution in Syria was “for dignity and freedom,” and the flight to safety compelled Syrian families to endure “humiliation,” while after resettlement, Syrian men and women articulate that “dignity is the most important thing” to “feel like we have worth” (Pearlman, 2021, pp. 281–282). In Lebanon, refugees have framed conversations about displacement and resettlement as a search for a “safe and dignified future” (Shawaf & El Asmar, 2017, p. 6). In Jordan, Syrian refugees have portrayed themselves as having the courage and agency to realise their dreams and ambitions, the human right to migrate to another country, and the desire to be successful and respected (Panter-Brick, 2021). Across diverse war-affected contexts, such narratives highlight a deeply rooted search for

family and social dignity. Yet, when it comes to understanding the pursuit of dignity, the evidence base for refugees, particularly in low- and middle-income countries, is very limited.

Insights from qualitative research are especially valuable to build knowledge on what matters to people in conflict and forced displacement and how individual, family and community-level factors might interplay in everyday parenting. For example, qualitative research has offered an in-depth understanding of the nexus between mental health and parenting capacities (Christie et al., 2020) and provided evidence on the links between caregiver depression, violence towards children, and parenting intervention efficacy (Cluver et al., 2020) in non-Western contexts. Qualitative insights are also valuable for systems-level work to understand how humanitarian intervention programs strengthen agency, dignity, and resilience through a ‘cascade of effects’ from one generation to the next (Panter-Brick, 2023). Importantly, qualitative data can be useful to facilitate the next paradigm shift—from a deterministic model of refugee parenting to an *agentic model of coparenting in dignity*. An agentic model is one that “structures scientific inquiry around the trajectories of individual lives and decisions” (Acolin & Fishman, 2023, p. 5). It has gained more recognition in the wake of the COVID pandemic, as the field of population health has turned away from a causal, deterministic model of health to an agentic model that views individuals as intentional and meaning-making actors (Acolin & Fishman, 2023).

This study addresses knowledge gaps at the intersection of parenting, mental health, and dignity in the refugee literature. Through in-depth interviews with mother-father dyads, we investigated how coparenting is enacted in protracted refugee contexts and why mental health and family values matter to parenting dynamics. Such evidence is critical to

strengthening family systems perspectives, supporting wellbeing, and promoting healthy child development.

2. Methods

This study received formal approval from the Prime Minister's Office of Jordan and ethical approval from Yale University (1502015359) and Trinity College Dublin (01E/2020/10). All study materials, including the qualitative interview guides and information on the piloting process, are available online on Open Science Foundation (https://osf.io/gcv5z/?view_only=8cb81ef7dd754217a36212082018d324).

2.1. Participants

This qualitative study ($n = 15$ father-mother dyads) drew from a cohort study of Syrian refugee families, taking part in a father-mother-child ($n = 160$ triads) survey undertaken to quantify how closely spouses agreed with respect to the father's levels of engagement at home, caring for his children, and supporting his wife in parenting activities (Hadfield, Al-Soleiti, et al., 2024). These 160 families had also participated in a cluster randomised trial undertaken with mothers and their 4-8-year-old children ($n = 322$ dyads) to assess the impacts of a reading intervention. This intervention did not influence family relationship outcomes. All families in the cohort were drawn from local community-based organisations that served socio-economically disadvantaged households in the city of Amman and in Zaatari refugee camp. For this study, we originally reached out to sixteen couples; we excluded from analysis one mother interview because we could not reach her husband, whose phone had been disconnected. Our sample size was determined by time and financial constraints in the field: we designed the qualitative study after the

completion of the father-mother-child survey because fieldwork suggested interesting data on coparenting. Participants provided informed consent to be contacted for qualitative interviews during their survey participation. No families refused to participate in the study.

The selection of respondents followed a quota sample (~10%) of the father-mother-child triad survey. One-third of dyads were couples for whom fathers had indicated *higher* levels of parenting involvement relative to what the mother reported the father's levels of involvement to be (i.e., couples were discordant in father-mother reporting). Another third were couples for whom fathers reported *lower* levels of parenting involvement relative to what mothers reported (these couples were also discordant). A final third were couples who agreed with respect to levels of father involvement in the home (i.e., couples were concordant). Such data were collected by implementing the Father Involvement Questionnaire (FIQ), a 20-item instrument developed for use with refugee and low-resource communities in Lebanon (Hein et al., 2020), which we adapted for implementation with both spouses. With these data collected just prior to the start of our qualitative study, we were able to include families representing diverse levels of spousal agreement on everyday parenting practices. Our purpose was to reach a balanced sample rather than to analyse qualitative data in concordance with father-mother reporting.

2.2. Data collection

Our interview guide, developed in English and Arabic, featured open-ended questions on mental health and parenting practices. We translated, back-translated, and piloted the interview guide, reviewing all data with the field team before final implementation (April to July 2022). We asked respondents to describe two parents in their community they

knew personally (asking women about mothers and men about fathers). We asked them to describe one parent with “bad” mental health and one with “good” mental health, then describe one parent who provided “engaged” parenting and another who did not. While storytelling was first developed as a culturally grounded approach to mental health assessment with war-affected Afghans (K. E. Miller et al., 2006), we used it to elicit in-depth conversations about coparenting. For both mental health and parenting narratives, we asked respondents to contrast the mindset, behaviours, and circumstances characterising their examples. Next, we asked open-ended questions about the respondents’ own parenting experiences and how these were shaped by exposure to war, displacement, and other challenges. Finally, we prompted each participant to discuss how they addressed potential disagreements related to raising children.

Fathers and mothers were interviewed by three Jordanian researchers, two males and one female, with whom they were already familiar, given prior participation in the cohort study. Both fathers and mothers were independently contacted by phone to explain study procedures, gain informed consent, and arrange a time for the interview. To respect local preferences and address issues of privacy and convenience, fathers were interviewed by phone, while mothers were interviewed during daytime visits to their homes: Syrian fathers spent much of their time outside the home, working, looking for work, socialising, or running errands for the family, whilst Syrian mothers were mostly homebound with their children. All interviews were conducted in Arabic, the first language of respondents and field staff. They ranged between 35 and 127 minutes in duration ($M=55.8$, $SD=21.6$); they were shorter, on average, for mothers ($M=50.3$, $SD=8.7$) than fathers ($M=61.3$, $SD=28.7$). Interviews were audio-recorded, transcribed (by co-author LS), pseudonymised, and then translated to English by two co-authors (LS and QS) fluent in

Arabic and English. Participants were compensated for their time with a small gift (mobile phone cards for fathers, ten Jordanian Dinars for mothers).

2.3. Thematic analysis

We conducted a thematic analysis, following Braun & Clarke's (2006) guidelines. First, two co-authors (LS and QK) immersed themselves in all the transcripts, independently coding them line-by-line, using open coding; this involved close reading to create codes that summarised sections of text. The narratives from spouses were coded together to capture the data's dyadic nature. After LS and QK independently coded 44% of the interviews—meeting weekly to review work and resolve coding discrepancies—LS coded the remaining interviews, whilst QK reviewed new coding. Unresolved disagreements would have required the input of other co-authors, but there were none.

Second, we employed an inductive approach to uncover the patterns within the text codes (Braun & Clarke, 2006). Three co-authors (QK, LS, and MAS) independently generated sub-themes and then reached a consensus on their content by comparing notes. We grouped all codes under sub-themes and discussed them (in Arabic and English) in team-wide Zoom meetings. In this way, we checked for the emergence of new codes and themes identified within textual data and for depth and breadth of theoretical insights, two processes known as inductive thematic saturation and theoretical saturation (Saunders et al., 2018). Third, we visualised these data, using Obsidian (v. 2022) software, to picture the overlap between father-mother data (as exemplified in Figure 8) and better see the relationships between codes, sub-themes, and themes (Figure 9). That is, we supplemented textual analysis with visual data to better consider dyadic interactions and thematic saturation. Visualising dyadic overlap (Figure 8) provided us with robust ways to identify what was common or shared among different couples and to analyse these

shared aspects in more detail within the text data. An example of dyadic textual data is shown in Table 3. All names have been changed, and identifying information removed.

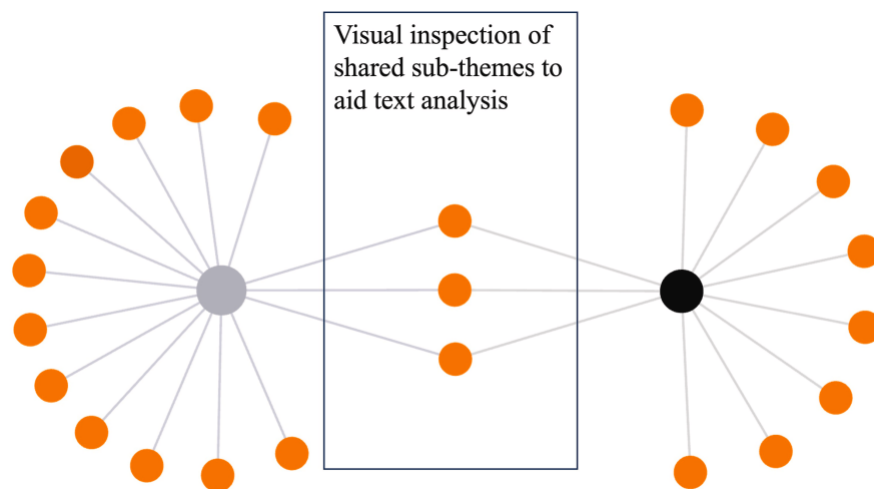


Figure 8. Data visualisation of topic overlap from dyadic mother-father narratives.

Note. Textual data from Hosam (father, grey circle) and Hend (mother, black circle), organised into sub-themes (orange nodes).

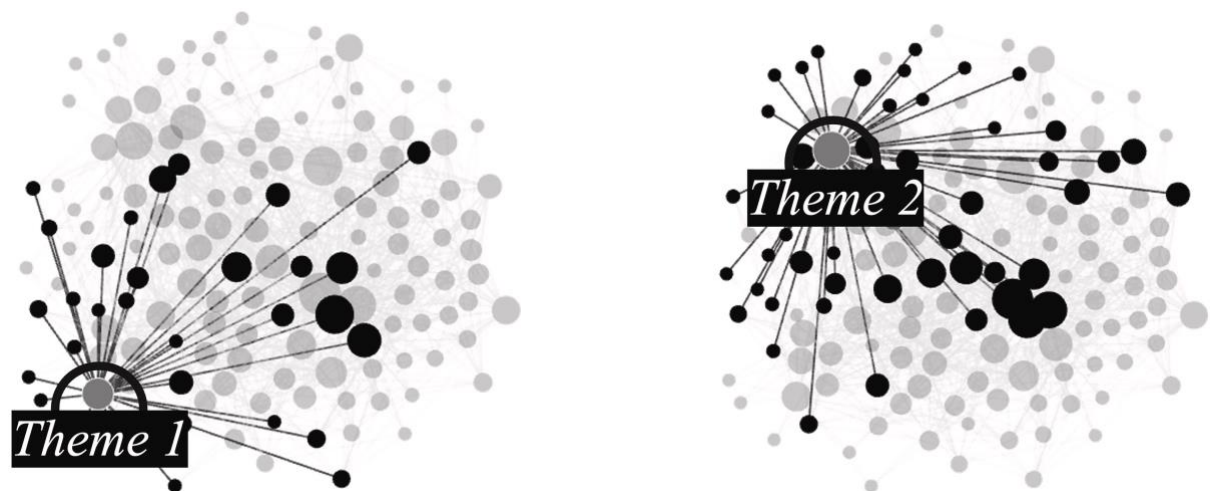


Figure 9. Data visualisation showing how different themes captured different patterns in the textual data.

Note. Theme 1 and theme 2 captured markedly different sub-themes (shown as black nodes).

Table 3

Example of dyadic narratives of war-related experiences and responses after forced displacement

Hosam (Father)	Hend (Mother)
I told you that after the incident [son's death], my wife got really tired, I mean, she was a different person before. [...] God blessed me with the child I wanted – he was very attached to me, and I was very attached to him. At that time, I was scared of dying, I was scared that I would die and not see my son grow up. [...] The incident happened on the first day of Eid. I: May God bless his soul. How old was he? F: A year and a half. I was very attached to him, he died in my arms. I thought that his brother also died, I even thought my wife had died. We were hit by a missile when we were getting back home. I: May God bless his soul, and praise to God that God saved the rest of you, and gave you patience. It was an immediate and sudden death because of the missile?	My husband tried not to upset me, but he used to tell me 'I know your pain, your pain and my pain are one and the same. He tried to take on the pressure, so that nothing happens between us. I've really been in a difficult mental condition - whenever I remember my son, I can still hear his voice, I start crying. I mean, when I feel that I miss him, I get very stressed out. It affected my home ... Gradually I started telling myself that's enough, I want to get out of this state. I mean, the last year we spent in Syria, the children didn't go to school, and when we came here, we enrolled them in school. Because of the incident, I was scared there, I had just lost a son, my other son stayed bedridden for around a month. We suffered. My husband and I were like a shelter (support) for each other, <i>alhamdulillah</i>, I was standing by

F: Yes, of course. The missile was shot, and his side, and he was standing by mine ... The seconds after that, it hit us. *Alhamdulillah*, pressures just piled up in our heads, but when they were both [two sons] dying in my arms, one gets away from the pressure, it gradually I was trying to rescue both, and I didn't decreases, but it leaves a trace. When the know what had happened to my wife and children went to school for example ... my 17-year-old son saw everything (the incident). He daughter because I couldn't see them. I tried was 8. He saw his brother in front of him to save the two boys ... My other son was He saw a horrible scene ... He saw part of his badly injured too, he had a severe injury to brother's brain coming out of his mouth. He his genitals, his left testicle was completely saw that and he kept waking up at night gone. My older son, a fragment of the missile screaming for a long time, he used to scream 'I hit him directly in the head and got to his want Omar' [brother], and he used to keep brain directly, his brain was shred into crying. He was affected, but because he's full pieces. of hope, he got himself out of that state by

I: There's no power and no strength except with Allah. himself.

F: And I got shot in my leg all the way down. And there were fragments that got stuck, *subha nallah* if they were to hit me, they could've killed me or something. Umm, regarding the emotional effect, I was broken. I started saying God why didn't you take us all? My other son kept bleeding for a month, his wounds would reopen and bleed, it drove me crazy. I started to ask God to put him out of his misery, I said things My mental status was affected, I stopped praying, even though I can't live without praying, but I

stopped praying. Was it anger? I don't know,
but I stopped praying, may God forgive me. I
used to cry all day. I used to go to his grave
every day and spend 4 or 5 h there, I would
stop crying when I got there, and sit by his
grave and talk to him, I became a crazy
person One moment I would be crying
the other I would start laughing. I tried to let
it out with poetry. I started writing poetry. I
started writing lament poetry. I wrote one I
named: 'The first day of Eid.'

3. Results

3.1. Demographics

Participants were displaced to Jordan from Syria between 2011 and 2014. Fathers averaged 37.9 years ($SD=4.8$) of age and had 5.1 ($SD=2.5$) children, while mothers averaged 31.5 years ($SD=6.5$) of age and had 4.5 ($SD=2.2$, range: 2-9) children (Table 3, supplementary). One father had two wives; we interviewed just one. Families lived in flats in Amman ($n=10$) or mobile caravans ($n=5$) in Zaatari, with between 1.3 and 5.5 people per room ($M=3.0$, $SD=1.3$) and household sizes ranging from four to eleven people ($M=6.9$, $SD=2.1$). Five families living in Amman had previously lived in a refugee camp. Regarding education, two mothers had never attended school, whilst eight fathers and six mothers stopped school somewhere in the primary level (grades 1-6). All other

respondents had attended at least some secondary-level schooling (grades 7-12), including one father who had attended university. One father and four mothers were non-literate. Twelve fathers were employed, mostly part-time in construction; only one mother worked outside the home full-time. In all but one family (where the mother, Ibtisam, also worked), there was only one person in the household with paid employment.

3.2. Themes

We identified four main themes: negotiation, mirroring, anchoring, and transformation (Figure 10). Spousal interactions reflected concerted efforts to reconcile past traumas, adapt to current challenges, and sustain aspirations for the future. Specifically, we saw an intertwining of mental health, coparenting style, and cultural values in the ways spouses discussed what was best for their children. In structuring everyday caregiving responsibilities, coparenting efforts endeavoured to offer a sense of normality amid the chaos of displacement.

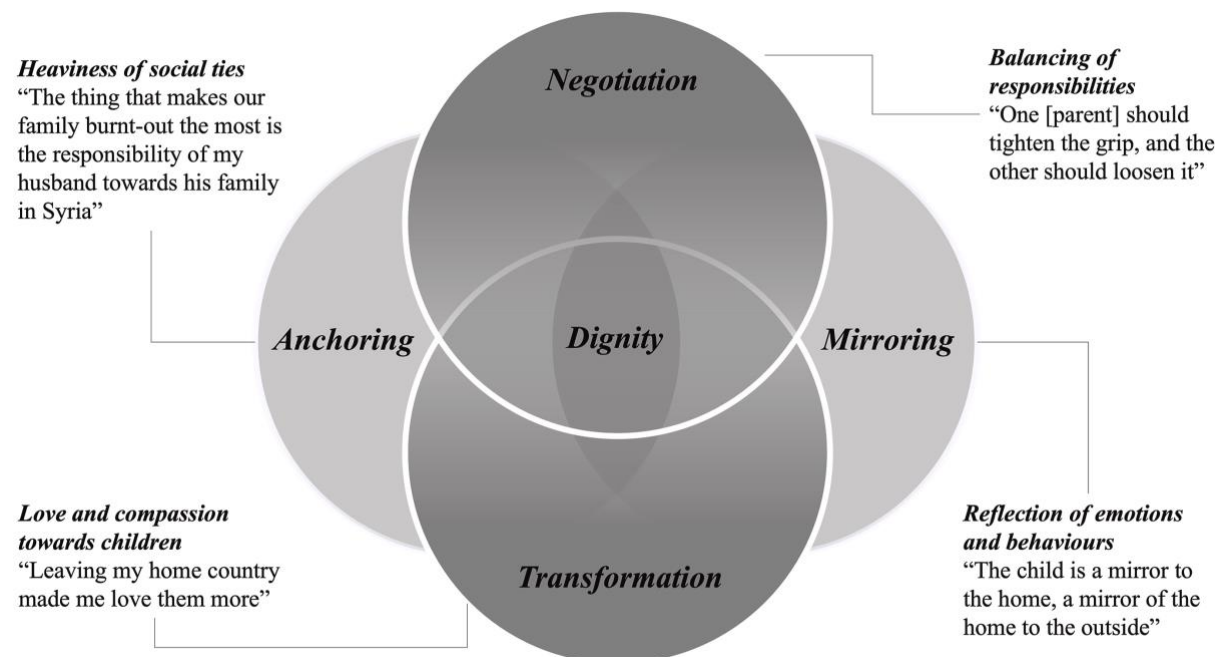
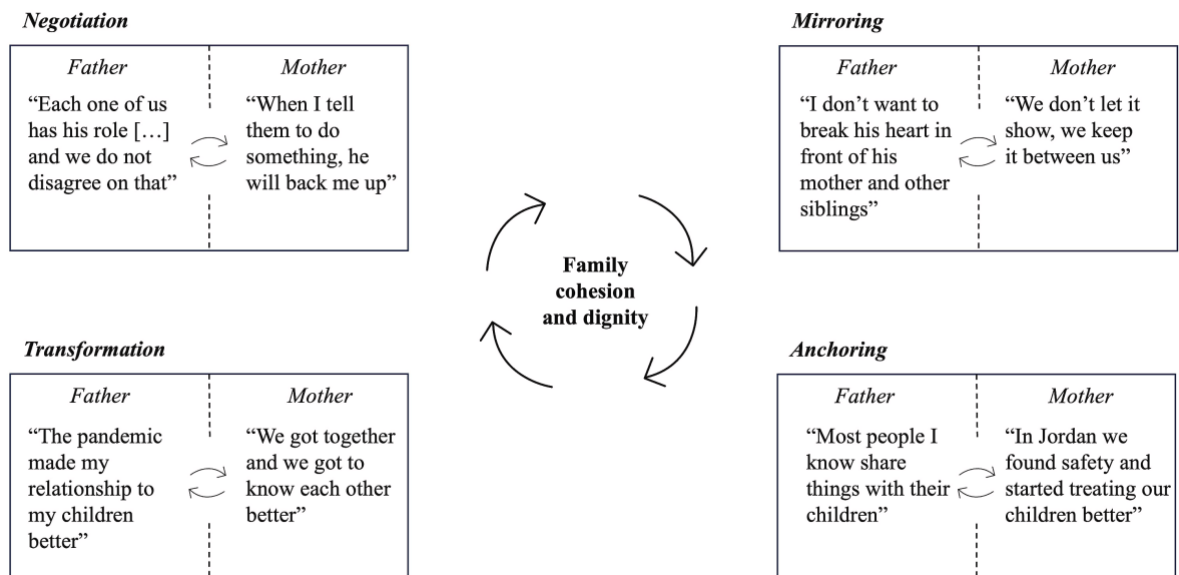


Figure 10. Four dimensions of dyadic father-mother narratives and family caregiving systems, with representative quotes.

Father-mother narratives revealed instances where the family achieved a sense of cohesion and dignity through interactions that prioritised stability and family togetherness (Figure 10a), as well as instances where the household was marked by evident disunity (Figure 10b). Faris and Fatma, for example, were clear with respect to coparenting responsibilities. As Faris stated, "each one of us has his role... and we do not disagree on that," meanwhile, Fatma knew that her husband would "back [her] up" when she instructed the children "to do something," suggesting that the family placed cultural value on respect and mutual support. They were careful to mirror harmony in family interactions, prompting them to hide potentially disturbing conflict: when Faris said "I don't want to break his heart in front of his mother and other siblings," it complemented Fatma's strategy: "We don't let it show, we keep it between us." In turn, the family had responded to the COVID-19 pandemic in ways that strengthened their relationships, as they "got together [and] got to know each other better" (Figure 11a). Together, these processes reinforced the family's sense of cohesion and sense of dignity. Conversely, couples like Fawaz and Fayrooz were trapped in negative family interactions that destabilised the family (Figure 11b). They were "always in disagreement," and their children got into fights at school, causing the father "trouble with other people" and the mother "blame." Each negative incident intensified the next, leading Fawaz to say, "at the end, I hit him," when his son would not listen, and Fayrooz to leave home, which meant the "children stopped school and started living like barbarians." The heaviness of family ties was shared with neighbours, where fathers had "always problems" with their children

and mothers cried “every time” they spoke of certain children. In these families, disagreement led to disunity, and disunity morphed into social indignity.

a)



b)

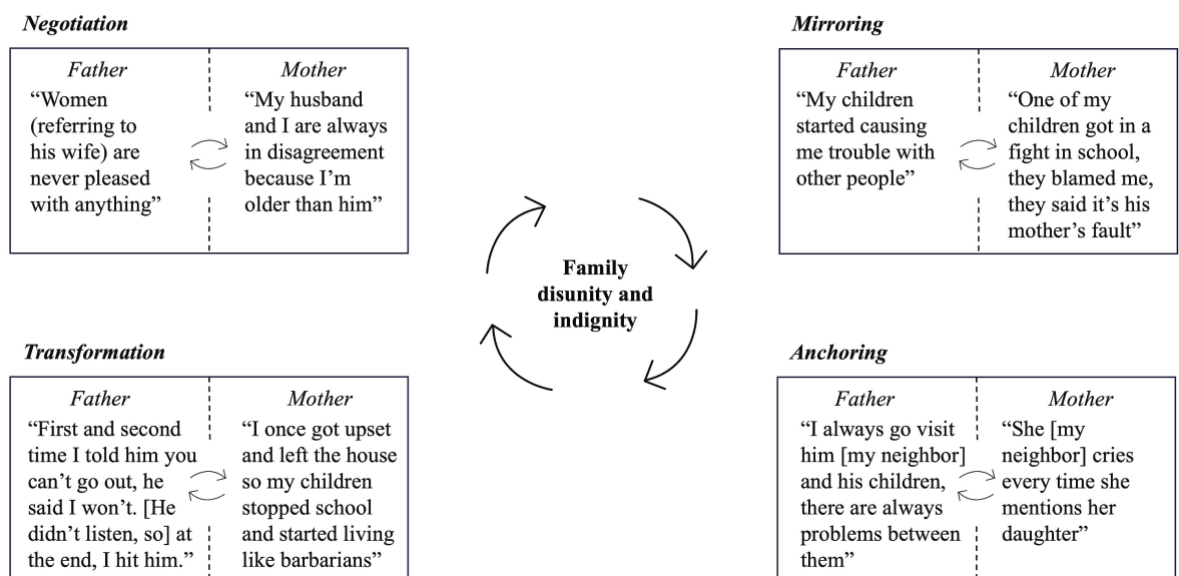


Figure 11. Examples of coparenting narratives: (a) stabilising family interactions between Faris (father) and Fatma (mother), (b) destabilising family interactions between Fawaz (father) and Fayrooz (mother).

Note. Quotes were grammatically corrected; original transcripts are available upon request.

3.2.1. *Negotiation*

Parenting as a couple meant, in essence, negotiating with each other to decide on how to raise children. Syrian refugee couples did this by “discussing everything” or “sharing opinions” with one another, explaining that discussions (مناقشات) would help achieve agreement. The aim was to achieve one’s desired caregiving outcomes through a balance of parenting responsibilities. Negotiations about childrearing required strong communication, the ability to compromise, and a willingness to work together despite differences in positionality.

For example, when Fatma asked her husband Faris to physically punish their young child, who had been disobedient, Faris encouraged her to use a gentler approach to “win the boy over.” Fatma followed his suggestion to use a less harsh approach to discipline. Later on, Faris told his wife he had decided to withdraw their children from school with a view to protecting them from bullying. Fatma disagreed with this decision and negotiated with Faris until they agreed to a compromise: their sons would be withdrawn from school as long as their daughters were allowed to resume their education. In matters of parenting decisions, especially concerning their children’s schooling, Fatma emphasised the importance of mutual support. She valued the everyday moments when she and her husband reinforced each other’s approaches: “when I, for example, tell the children to study, [Faris] tells them, ‘yes, study, listen to your mother.’ When I tell them to do something, he will back me up, which makes me happy.” However, while striving for shared decision-making, Fatma was clear about which matters were non-negotiable. For example, she explains that if her husband caused their daughters to leave school again,

It would lead to a divorce between us, because these are girls... they will be the ones responsible for teaching their children. You know, fathers go to work and who's responsible for teaching the children? The mother. I mean, whatever it takes, there's no way, even if I had to get a divorce, I would keep them in school and find work to pay for their teacher but never get them out of school.

Negotiations sought to balance different parenting approaches to win the child over. In her role as mother, for instance, Marwa remarked, "how will you [reason with the child] if the father isn't by your side?" Spouses wanted to balance nurturing with discipline, often stating that fathers should be the ones to enforce discipline when children misbehave. In the words of Dawod, a father of six, the role of the father is to "scare" the children "if they do something wrong." But couples would also try to balance each other by taking on distinct roles. Joud explained:

When we're alone and the children are in school, my husband and I sit together and talk about these things. I tell him that one of us should tighten (يشد) and the other should loosen (يرخي) the grip, because if we both loosen the grip, it will ruin the child.

The 'grip' refers to the parental control, guidance, or style (اسلوبنا) necessary to steer a child's behaviour relating to personal conduct and social commitments. The aim was to navigate parenting challenges with just the right amount of grip, which involved careful negotiation between spouses in response to specific demands. For example, parents were often anxious when their children asked to see friends outside the safety of the home. One couple, Rashdan and Rawan, explained that they "have an understanding:" if their daughter "asks her father for permission to go out, and he says yes," then Rawan (mother)

says she will set “a specific time for her to return home.” Parents were clear that both “goals and means” of parenting had to be well-aligned to better safeguard family functioning and children’s wellbeing. For her part, Ibtisam (mother) explained:

There was an example in which both parents wanted to raise their children well, but the mother wanted to do it with understanding and the father wanted to use force, so they agreed on the ends but didn’t agree on the means. I’m sure they should agree on the means, the same way they agree on the ends.

When prompted to explain how such ‘disagreement’ would affect children, Ibtisam did not hesitate: “they will be lost in the middle,” and it would “affect their mental status; it would make them less confident and make them feel unsafe.”

Couples were keenly aware that successful negotiations between them improved family harmony. At times, couples became distant from each other or interfered with each other’s parenting. As Layan stated, when her husband is “angry and not in a good mood, he wouldn’t be in the mood to talk to anyone.” Poor mental health underlined these dynamics. This was most evident at times of overwhelming stress, where some couples could not resolve important parenting issues, resulting in one partner adopting a negative style of parenting and the other disapproving of that approach without being able to negotiate a better dynamic. This happened to Zaytona and her husband during the COVID-19 pandemic: “he started interfering with everything, and paying attention to [things] he didn’t use to pay attention at all.” Other women, like Marah, felt a husband’s interference prevented her from being the parent she wanted to be:

Marah: My husband is like that, he likes to be the head of the family and responsible for everything. If I tell him to let me go out and get things, he says, “no, you stay in your home.” He likes to be responsible for everything.

Interviewer: Does this bother you?

Marah: I feel it does; I don't know, like some people like it and think it's a nice thing that the father is responsible for everything, but for me, I feel it bothers me. I mean, sometimes when he's at work, and I want to go get something that I need to get, he doesn't let me, for example, if I want to go get my daughter vaccinated, he says, "you can't go alone, it's forbidden, you have to have [a male chaperone] with you." This thing bothers me a lot.

Marah faced such parenting challenges as a result of her husband's strict adherence to conservative cultural ideas pertaining to social respectability. Disagreement was described as a chronic issue in the spousal relationship – Marah's attempts to navigate parenting dilemmas were met with persistent inflexibility on the part of her husband.

3.2.2. *Mirroring*

Parents described their parenting feelings and behaviours as a process of mirroring (انعكاسات), whereby their emotions and actions were "reflected" from one family member to another. Many parents were keenly aware of how emotions and experiences would be shared between themselves and their children. For example, Fawaz explained how witnessing spousal disagreements could have a profound and direct impact on his children:

When the children see that their mother and I are laughing together, they become happy, and when they see that there is tension between their mother and me... they become unhappy ... it reflects (بينعكس) on them directly, sometimes I sleep in a separate room, and they notice that, and say "why is my dad sleeping alone!" And when they see that we're laughing, they become in a better mood, they become happy.

Parents took steps to avoid conflicts in the presence of their children, mindful of their impacts on the next generation. For example, Dawod (father) ensured that the “kids aren’t there” during marital conflicts. Sahar and her husband Saeed discussed all disagreements privately, “not in front of the children.” Majed explained that because his daughter reacts viscerally to domestic conflict and “immediately starts crying,” he must actively work to rectify the situation by making amends with his wife and “laughing together in front of her, so that [their daughter] doesn’t think we have problems.” Through their actions, parents like Majed demonstrate an understanding of how emotions generated through the spousal relationship can impact children and spill over into the parent-child relationship. Children, too, sought to shield negative emotions from their parents. During COVID restrictions, for example, families were confined to the home and worked hard to manage their feelings in order to avoid negatively impacting other family members. At that time, Zaid realised that his children wanted to shield him from negativity:

If the children have a [mental health] thing, they always try to hide (يخبو) it from me so they do not bother me. I know what they are hiding, but they [keep it from me] so they do not increase the burden on me. Meaning that they like to spare me, as in not to increase the worry on me.

Zaid built on this awareness to foster love and harmony within his household. His approach led him to describe what he perceived as a “surprisingly” positive experience during COVID-19; his family “were very supportive” and “actually got closer.”

Parents also drew on a mirroring analogy when discussing how parenting reflected social standing in the community. Dawod, for example, saw his child as “a mirror (مرايه) to the home, a mirror of the home to the outside.” He explained: “your child reflects (صورتك) your image fully. The way you treat him reflects your outer image (الخارجية).” Thus, parents were concerned that if they treated their children poorly or were

not sufficient disciplinarians, their children would not be a credit to them in public:

“[children] would think it’s normal if their parents are yelling and fighting, and it’s ok for them to yell and fight and hit people too” (Ibtisam, mother). In this way, parents highlighted two kinds of mirroring: one pertained to the emotional and behavioural interconnectedness between family members, and the other linked family-level behaviours to social respect and dignity at the community level.

Notions of mirroring were deeply entrenched in local perspectives on mental health and social dignity. For example, Fatma described how mirroring presented itself in two of her neighbours’ families. In one household, she commented, the mother simply “can’t control” her children, and the children’s unruly behaviours reflect her state of “depression.” Emotions and behaviours were closely connected in that children who “never stay still” reflected poor mental health and parenting guidance. In contrast, the other household had a mother with “good mental health,” where mere eye contact prompted obedience and calm:

God willing, even if she just looks at them, the children understand her. They’re very calm, they’re not like other children who keep jumping around. When she visits me, they just sit beside her, they don’t move. If she tells them to move, they do, if she doesn’t, they just stay seated. You don’t feel like they’re children from this generation!

Fatma turned to describing emotional and behavioural patterns seen in the fathers she knew in the community. She described one as “always stressed out,” yelling to gain control over the children. This was counter-productive, as the children “stopped listening” to the point of defiance: “if they’re told not to do something, they do it on purpose.” Another father was calm and mentally composed, using “good words” and treating

children in a manner that elicited respectful behaviour. His children reciprocated with a calm demeanour and respectful obedience:

[The one father,] he's always stressed out because of his children, if they're going out, or if they're ruining things, or if they get out of the house. They got lost more than once, and [the family] started looking for them [in the neighbourhood]. But the other one is more comfortable, he sits and talks. He doesn't pay much attention to that [issue of getting lost] because he knows that his girls are calm and wouldn't get out of the house. He is very calm, and his mental status is... very calm.

In essence, good mental health and engaged parenting went hand-in-hand, fostering cooperation, interpersonal and social respect within the family and local community. By contrast, "bad" mental health went hand-in-hand with distant or disruptive parenting; it resulted in family dynamics that were considered disorderly, unsettling, and undignified.

3.2.3. *Anchoring*

After resettlement in Jordan, many families remained anchored to Syria — they were fearful for the safety and wellbeing of their relatives back home and felt an obligation (تلازم) to socially and materially support them. The pressures of fear, insecurity, and social obligations directly impacted what they described as "burnout," wanting to withdraw from parenting activities. As Ibtisam explained:

The thing that makes our family burnt out the most is my husband's responsibility towards his family in Syria. It causes us burnout because we think about it a lot, because his parents live alone in Syria, so he keeps thinking about them. When he

feels short on his duties towards them, he gets really upset, and I get upset when he gets upset.

For his part, Saeed explained how relationships in Syria affected his parenting experience:

I have my mother and my sisters in Syria; you have to send them [money] because you are obligated (ملزوم) to them. For example, when you feel you're broke and short on rent, you become withdrawn, and your mind gets occupied with bigger things, so you become withdrawn from your children.

Social networks greatly mattered to fathers' sense of worth as a caregiver and socioeconomic provider. Comparing his situation to another parent who was financially better off, Mohammed described himself as "a good and a bad father at the same time; I mean, I try to give the [children] everything I can give, sometimes they ask for things that I can't get." The other Syrian parent, in Mohammed's opinion, "always gets out and has fun; he rents a car and goes to a farm for a couple of days. He has relatives in Europe who send him and his wife money. He's always living life." Mohammed felt like a good parent because he tried to give his children what they needed – but he also felt like a bad parent in comparison to someone with more economic resources.

Sometimes, parents served as a positive anchor for each other, offering support and advice. Dawod shared a story about a time he stepped in to help a Syrian neighbour and his son, who was showing emotional distress. He recounted the man's gratitude for noticing his difficulties and for offering concrete parental guidance:

He was very surprised, his eyes teared up, and he asked me what to do. I told him, "walk with me," and we took [the son] to a bike place, that was [his son's] wish,

to buy a bike. In two days, the boy was different, from sky to earth. He told me, “God bless you.”

Social ties could both secure and weigh down a family when having to decide where to live, where to work, and how to stay safe. For example, Faris and Fatma had been offered to leave Jordan for resettlement in the U.S. They refused this opportunity because the people they knew had warned them that life in the U.S. would negatively impact their culture, values, and ability to raise children effectively. The mother, Fatma, said:

We had the opportunity to move to the United States, we refused it, and [Faris] regrets it; he keeps saying we shouldn't have said no ... people were telling us that [Syrian] people [in the U.S.] weren't comfortable. That one thing makes up for the other, that even if they provide everything for us, the children would be ruined and lost (بنتز عوا), and the wife would be lost; that's why we refused it.

They had discussed their options together, deciding that they could not bear another experience of “ruin” and “loss.” They wanted to shield their children from further loss and instability, and to live a dignified life, rooted in family togetherness. For Faris, the local Syrian community represented a place where “most people [I know] share things with their children.” For her part, Fatma found safety in Jordan and “started treating [the] children better.” Nonetheless, for Faris and Fatma, the process of anchoring had cut two ways: it had offered stability but also denied broader opportunities. A sense of regret began to surface, given the challenges of living in Jordan, knowing that resettlement in the U.S. might have offered better educational prospects for their children.

3.2.4. Transformation

Father-mother narratives were imbued with a sense of transformation, as parenting strategies changed over time in response to war trauma and ongoing challenges. In discussing how parenting behaviours had been transformed across generations, for example, parents used the word “softening” (تحنُّن). Here is how Sami described his relationship with his children following his departure from Syria:

Leaving [Syria] made me love them more because they became my country, my religion, and my whole life (عيشتي). My children compensated [the loss of] my country. When I see them around me, it makes up for everything... When I go to work, I say to myself, “I can’t wait to get back home and sit with them.” I feel full of longing for them. Sometimes, when I visit my sister [elsewhere in Jordan], I feel like I can’t wait to be done with that to get home – it feels like I’m returning to my home country.

The experience of war is what caused this transformation for Sami:

My heart softened because of the shock. I mean, God forbid, if someone you hold dear dies, you’ll find that your heart softens towards other people you hold dear. I mean, I lost a brother in Syria. You feel that this is your brother, and your other brother makes up for it; you give all the longing (الشوق) for the one who died to the one who’s still alive.

Such experiences of loss led many parents – especially fathers – to transform their parenting styles from what they had experienced as children. Fathers described adopting warmer, kinder strategies to protect the next generation from negative experiences and promote their children’s sense of safety, wellness, and development. Other fathers had to

respond to the ongoing conditions of precarity and material poverty. Jabir (father), who lived in Zaatari, explained:

The problem is that the environment we're in currently is different from the one we used to live in in many ways. We're in a camp, and it consists of caravans and tents. In Syria, things were different; we had a house and other things and necessities... sometimes you get emotional, you feel compassion, and you worry about your children... I think, for us, it made us more kind, honestly.

Parents also underlined the significance of early experiences in shaping mental health and caregiving behaviour. For instance, Fayrooz (mother) discussed how a father's mental health could result from early experiences: "the way his father treated him when he was a child. If his [own] father was uptight, he'll have a psychological complex and treat his children the same way he was treated." Similarly, Hosam, who was physically abused as a child, said he heard people comment about him, early on, that he "will have a psychological complex when he grows up, he will become crazy" due to the harshness of his upbringing. Hosam, however, was determined to break the cycle and not "let that harshness control (القسوة تسيطر علي)" him. He consciously chose to adopt a softer demeanour with his children, even if it initially felt "fake" to him. He found it difficult to change, but over time, his choice away from harshness became a habit, allowing him to create a more positive environment for his family. In this way, positive change could be nurtured, and habits could be reshaped, where needed, across generations.

Strikingly, many families spoke about creating a sense of normalcy when seeking to characterise their caregiving behaviours. Fathers like Majed and Zaid declared: "We are normal parents." As Saeed explained, parenting and coparenting were "for the children's own good. I mean, at the end of the day, we wish the best for our children."

Notably, they did not describe themselves as an unusual group of parents but as ordinary parents adapting to extraordinary circumstances. This sense of normality was profoundly challenged when, as detailed in Table 3, Hosam and his wife Hend had to care for the family in the aftermath of a missile attack that killed one of their sons and severely injured another. Reflecting on the links between the past and present, Hosam and Hend detailed how feelings of fear and exhaustion and tragedies of death and injury resulted in them “sheltering” and “standing by” each other as a couple and family. Hosam emphasised the significance of establishing a sense of normality through a safe and secure home environment for their children. He implemented a weekly family ritual of staying home every Thursday to promote stability. He prioritised this routine over socialising, stating, “I have to decline [social invitations] on Thursdays because that day is reserved for my wife and children.” Hend indicated that Hosam creates a “special atmosphere” for the children, “discusses things with them”, and gives them “freedom” to pursue their passions. Hosam believes this has resulted in open communication with his children: “When they want to tell me a story, I sit and listen, even if they repeat it a thousand times.” What makes these experiences transformative is Hosam’s reflection on changes across generations:

I never dared to sit and talk with my father. This is one of the differences (الاختلافات). I got married and had two children, but even then – if I wanted to talk to my father or seek his advice, he didn’t look me in the eye. If we’re watching television and I ask him about his opinion, he looks at me sideways, and if he doesn’t like the conversation, he just moves his hand this way, telling me to go away; he doesn’t use words. For example, I don’t remember ever kissing my father other than the time when I joined the army.

In acknowledging differences across contexts and generations, Syrian parents were transforming past experiences of caregiving, which had lacked fatherly affection. In doing so, they became more open to their children's hopes and dreams: some fathers encouraged their children to learn new skills and pursue their interests and passions in sports, art, or music. For example, Hosam explained:

It got me excited because when I saw [my son] write, I saw that he is talented. So, I'm helping him develop this talent despite it not being consistent with my values or environment. But I just don't want my son to say, "I had a dream, and my dad deprived me of it. In my life, I never dared to go to my father and read one of my poems to him."

Parents also sought out new opportunities. Some had moved into cities because they thought that refugee camps offered limited schooling possibilities. Others sought out resettlement options further afield: Jabir, for example, wanted to leave Jordan and go to America or Canada because he believed that people in those countries "focus on education." Parents sought out opportunities for themselves, not just for their children. For example, Marwa took a parenting program offered in her local community, and thought this benefited her whole family: "I started treating the children well, my husband and I started understanding each other better. We used to fight all the time. The course helped us a lot. My relationship with my husband got much better." In expressing her concerns about the future, Marwa emphasised the value of "sticking together" as a family and show social dignity. In much the same vein, another mother, Zaytona, emphasised the importance of long-term patience to ensure that the "family doesn't get broken."

Building togetherness, counteracting experiences of loss and ruin, was at the forefront of concern: when narrating the story of a father unable to provide for his family, Dawod

emphasised: “this man fights with his wife all the time ... his family is broken up (مفككة).”

Family disunity – and social indignity – was there for all to see.

4. Discussion

This study focused on the lived experiences of coparenting among Syrian refugees and the interconnections between coparenting, caregiver mental health, and family values. Our data speak to *how* fathers and mothers navigate caregiving experiences together in times of protracted displacement and *why* they strive to engage with their children and spouse. In narrating how they communicated and related to one another, Syrian mothers and fathers exemplified the inner workings of family caregiving systems, highlighting some of the key hallmarks of caregiving and coparenting in refugee communities who strive to express a sense of family togetherness and dignity. This grounded understanding of refugee lives addresses important gaps in current knowledge, moving us towards an agentic model of coparenting and intergenerational change.

We identified four themes in our analysis: dyadic negotiation, mirroring, anchoring, and transformation (Figure 9). As shown in Figure 9, negotiation was necessary to balance coparenting responsibilities. This finding aligns with literature on family systems theory, such as Walsh’s work on family resilience (2021), highlighting the role of relationships and effective communication in overcoming the challenges of life adversity. It also resonates with research indicating that disruptions to social structures compel mothers and fathers to renegotiate their roles and responsibilities both within and outside the family (Asaf, 2017; Merry et al., 2017; Yalim & Critelli, 2023). In our study, while families emphasised that fathers were economic providers and wives shouldered home-bound responsibilities, they did so within the practice of dyadic negotiation, one

that evaluated character strengths, cultural values, educational background, opportunities for education or resettlement, mental health, and power differentials.

Parents were also concerned with the mirroring of relationships and anchoring of family ties. Often, these concerns emerged in the context of effectively managing the stress of everyday life and the burden of social responsibilities. Both mirroring and anchoring are constructs identified in the wider literature. The concept of mirroring, originally introduced in a psychotherapeutic context (Carrilio & Walter, 1984), pertains to the bidirectional interplay between children's emotions and parents' behaviours (Chartrand & Bargh, 1999; Eisenberg et al., 1998), a process embedded in transformative social learning theory (Wals, 2010). Acts of mirroring highlight the relational frame established between family members: they may reflect supportive family ties to shield each other from emotional issues or, conversely, reflect an inability to communicate and address emotional burdens. What emerged from this study is that Syrian refugees spoke of mirroring in two different ways: first, a mirroring of emotions and behaviours across generations, and second, a mirroring of parenting behaviour and social standing in the community (Figure 9). Notably, attaining the right level of "grip" on their children improved family dynamics and enhanced social respectability.

Syrian parents expressed the notion of anchoring in terms of the heavy weight of social ties. Post-resettlement, refugee households found themselves tethered to both their home country and local community: they felt "burnt out" from enduring responsibilities towards relatives left in Syria, and locally, deliberated education and resettlement opportunities as trade-offs against securing family cohesion and safety. Research shows that anchoring (Grzymala-Kazlowska, 2018) and strong social ties (Campos et al., 2014) are generally associated with positive experiences. By contrast, work by Stewart (2015) found that in marginalised communities, social ties are associated with the burden of

social responsibility. Our study shows that this burden was hard to escape for participants who had family in Syria. It also reveals the difficult trade-offs refugee families must make, for example, prioritising family togetherness and relative stability in Jordan versus relocation to another country with new opportunities for education and employment. These trade-offs echo the twin facets of entrapment evidenced in Afghanistan (Eggerman & Panter-Brick, 2010), where families were unable to both fulfil social obligations and realise their personal aspirations.

One of the most striking results of our study relates to the theme of transformation. As a result of their experiences of war and displacement, many parents shifted towards a more compassionate approach to parenting - a “softening” - breaking from a pattern of harsh parenting in earlier generations. These findings are consistent with work suggesting that parental empathy increases during challenging times (Bryant et al., 2021; Eltanamly et al., 2021, 2022). Yet, people who experience harsh or abusive parenting may find it difficult to engage in positive parenting (C. A. Greene et al., 2020). The parents in our study acknowledged this issue: many worked hard to change harsh parenting approaches experienced in their childhood. With effective negotiation and communication with their spouses, men who had hardly ever embraced or established eye contact with their own fathers now engaged with children with love, emotional support, and open communication. Here we see one of the clearest examples of agentic coparenting, anchored in the experiences of war, forced displacement, and the value of “doing the best” for one’s children.

Underpinning these four themes were efforts to uphold family dignity during precarity. After forced displacement, refugees see efforts to keep the family together and to uphold social dignity as a cultural and moral duty. This finding is in line with the narratives reported in several other studies on Syrian refugees (Pearlman, 2021; Shawaf

& El Asmar, 2017). In our study, fathers like Majed, Zaid, and Hosam insisted they were “normal parents” – eschewing the pejorative label of ‘refugee parents’—and wanted to build a normal life for themselves and their families in displacement. Families sought to construct a safe, predictable, and nurturing environment for their children to mitigate the mental health signatures of death, injury, and loss, carving a space for normal lives and livelihoods in the wake of displacement. These endeavours echo what we know to be vital for children’s learning and wellbeing (Fiese et al., 2002; Gentzler et al., 2005; Michalek et al., 2022), neurobiological growth, empathetic love, and community-level peacebuilding (Baranov et al., 2022; Leckman et al., 2021; Panter-Brick, 2023). Put differently, in the context of refugee lives, participants saying “we are normal parents” is both a defiant and dignified expression of agentic coparenting.

This brings us to evaluate how our study results can be used to facilitate better parenting interventions in systems-led research, practice, and policy. Our research specifically highlights the importance of enhancing refugee father-mother communication in ways that uphold mental wellness and family cohesion by taking into account the larger sociocultural dimensions of coparenting that include the balance of mutual responsibilities, the weight of social ties, the reflection of emotion and behaviours, and the pursuit of dignity amidst structural challenges. In terms of practice and policy, these findings call for strengthening coparenting systems in ways that help caregivers improve their relational agency, namely, the power to act together in light of their emotional and social relatedness to others. As perceptively discussed by Burkitt (2016), the concept of relational agency is critical to understanding social interactions, given that decision-making is nested within interpersonal and social relations. This relational perspective of agency, as existing in networks of interdependence, is indispensable to understanding

family caregiving systems in displacement, especially when addressing the evolving needs of coparenting practices amidst intergenerational and social change.

5. Conclusion

Our analysis offers methodological and theoretical advances to better analyse family caregiving systems and coparenting in the context of protracted displacement. It exemplifies how to capture a dyadic narrative and explains why, for Syrian refugees, living a dignified life is at the core of efforts to parenting together and efforts to parenting well. Such insights from qualitative research can help enrich the evidence base and transform existing paradigms on refugee lives, moving us towards a more agentic understanding of parenting decisions and intergenerational transformation. They are specifically relevant to research and interventions that seek to improve family-level communication and design integrated programs that support caregiving, wellness, and dignity.

Chapter 4: OBSESS to measure interactions transcending word meaning in refugee mother-child dyads

Khraisha, Q., Fan, M., Lee, M., Himsi, M., Raguragui, M., Chuhan, C., Put, S., Volta, B., Wadhwa, A., Dos Santos, E., Çakmak, M., Qasim, A., Sulowska, M., El Kharouf, A., von Stumm, S., Marechal, I., Dajani, R., Vining, R., Michalek, J., & Hadfield, K. (under review). OBSESS to measure interactions transcending word meaning. *Child Development*.

This Chapter directly builds on the twin insights highlighted previously—the near-absence of observational methods in the systematic literature review (Chapter 2) and the critical importance of non-verbal mirroring behaviours suggested in the qualitative study (Chapter 3). I address these limitations by developing and validating the Observational Behavioural System Evaluating Sensitivity and Synchrony (OBSESS).

This chapter is an exact copy of the one being reviewed in the journal.

Abstract

Observational coding systems are invaluable for understanding parent-child interactions, but their reliance on language comprehension limits their use with linguistically diverse populations. We developed the Observational Behaviour System Evaluating Sensitivity and Synchrony (OBSESS) to measure refugee parent-child interaction quality through paralinguistic cues (e.g., tone) and kinetic behavioural signals (e.g., gestures). Here OBSESS was tested on 1,092 videos featuring 287 Syrian refugee mother-child pairs in Jordan (children aged 4-8 years, gender-balanced). Most coders were unable to understand the language the dyads spoke, yet all demonstrated excellent intra- and inter-rater reliability across pre-registered analyses. Coding patterns also showed convergence, divergence, sensitivity to contextual factors, and temporal stability. OBSESS is made freely available to facilitate its implementation in low- to middle-income countries.

1. Introduction

By the end of 2025, it is estimated that over 139.3 million people will be displaced due to war and conflict, marking another rise in global displacement for the thirteenth consecutive year (UNHCR, 2025). This escalating forced displacement crisis disproportionately affects displaced children, who face heightened psychological distress and educational challenges (Blackmore, Gray, et al., 2020). As their vulnerabilities become amplified, displaced children may become highly reliant on their caregivers for their overall wellbeing and development (WHO, 2020). While there is a considerable number of studies that have advanced our understanding of responsive parenting in displaced populations (Khraisha, Abujaber, et al., 2024), they primarily rely on self-reported data, which presents significant limitations. Language barriers, varying literacy levels, and a lack of locally validated measures can skew self-reported information (Hayden et al., 2010; Hinchey, Khalil, et al., 2023), reducing the accuracy of data on parenting interactions.

Behavioural observation

Observational methods, including videotaped interactions, offer a more precise view of parenting behaviours than self-reported methods, as they capture them directly (Hawes & Dadds, 2006). However, due to resources barriers, the use of these methods is constrained in low- and middle-income countries (Bozicevic et al., 2024), where most displaced people reside (UNHCR, 2024). In these locations, displaced caregivers also face more significant mental health challenges that, in addition to reducing their emotional availability, distort their self-reported responses by hindering memory retrieval and influencing their perceptions or interpretations of past events (Meisenberg & Williams,

2008; Prime et al., 2023; WHO, 2020). Though not entirely free from biases, observation of parental behaviours in displacement settings could circumvent recollection bias and reduce the likelihood of socially desirable responses.

Unlike the relatively simple task of completing a self-report checklist or questionnaire, observational tasks require parents to manage their behaviour in real-time, imposing a higher cognitive load, which makes it more challenging for individuals to consistently conceal socially undesirable actions (Eysenck et al., 2007). Consequently, observational methods can provide a more authentic representation of parent-child dynamics, capturing behaviours that might be omitted or misrepresented in self-reports, which is especially important in displacement settings, where people usually have insufficient resources and legal protections that increase the pressure on them to respond in ways that align with perceived social or cultural expectations (Aleinikoff & Poellot, 2014; Bozicevic et al., 2024; Meisenberg & Williams, 2008).

1.1 Sensitivity and Synchrony

In addition to bypassing certain cognitive barriers, video observation can capture nuanced paralinguistic-kinesic aspects of parental behaviours that typically fall outside the scope of self-reports. This is particularly true when evaluating Synchrony—defined as the “matching of behaviours, emotional states, and biological rhythms” between parents and children—which is measured almost exclusively through video observation (Leclère et al., 2014, p.3). Other interactions, like sensitivity, or “the attunement of parents to their child’s emotional and cognitive states,” are only partially captured by survey data (Ainsworth et al., 1974; Borairi et al., 2024, p.2). This is because parental behaviours that support a child’s autonomy or provide guidance often include paralinguistic and kinesic aspects, which are responsive to the child’s actions or the context of the interaction, for example, stepping back to allow independent problem-solving (Borairi et al., 2024; Mathers et al., 2024).

Sensitivity and Synchrony are both fundamental concepts in attachment theory and developmental psychology (Deans, 2020; Leclère et al., 2014), and are relevant to a child's emotional security, self-regulation, social competence, and cognitive growth (Davis et al., 2017; Rodrigues et al., 2021). Though all parents demonstrate some degree of Sensitivity and Synchrony with their children, cultural differences shape how these behavioural domains manifest (e.g., universality does not equal uniformity; Black et al., 2024). As suggested by recent observational research, in one culture, mothers may maintain high arousal levels to keep their infants positively engaged (Palestinian Arab), while in another, mothers may modulate their arousal to match their infants' emotional states (Israeli Jewish; Abu Salih et al., 2023). These findings also illustrate how observational data can capture the 'hidden choreography' of cultural parenting practices—subtle, automatic, or complex behavioural patterns that typically elude self-report measures, which often presume uniform expressions of parenting across cultures (Lansford, 2022a; K. E. Miller et al., 2025).

1.2 Limitations in literature

Although observational data can provide a more ecologically relevant and comprehensive understanding of key parent-child dynamics, implementing such methods raises several challenges, including ethical and privacy concerns that are acutely felt in displaced populations, who often have traumatic histories, may distrust authority figures, and may be wary of being recorded (Hall & Werner, 2022). In some instances, such distrust has led researchers to rely on live observation after participants refused video recording (Lembcke et al., 2020). Cultural differences in gender interactions can further complicate not just video data collection but also analysis; for example, in our study, some mothers were uncomfortable with men on the research team viewing or analysing their videos.

Additionally, analysing video data requires extensive resources, as it involves detailed coding by trained professionals using expensive software. This becomes even more challenging and resource-intensive when analysing observational data from non-English speaking participants, who represent the majority of the displaced population (UNHCR, 2024b). According to the United Nations High Commissioner for Refugees' latest figures, displaced individuals currently originate from 205 UN-recognised sovereign states, each with its own languages or dialects (UNHCR, 2024). Often, multiple linguistic groups of the same nationality coexist within the same country. Jordan alone currently hosts refugees from 34 nationalities, each speaking multiple languages and dialects (UNHCR, 2024), such as the Syrian refugees, who can speak various regional dialects of Arabic (e.g., Damascus, Aleppine) and other languages like Circassian, Turkmen, Kurdish, Kurmanji, and Domari, each with its own respective dialects (Abu Kwaik et al., 2018; Ain, 2016).

Previous attempts to use observational coding systems among displaced people have not only been hindered by the need to understand the participants' language but also faced additional issues. These include high cost due to the need for specialised software or training, inadequate documentation, low reliability, small overall sample size (or small sample size used to check interrater reliability), or high risk for bias as a result of relying on one single coder or a culturally homogeneous team (Dalgaard et al., 2020; Hein et al., 2020; Lebiger-Vogel et al., 2022; Lembcke et al., 2020; Scharpf et al., 2023; van Ee et al., 2013; Wiepkema, 2013). Earlier systems have also operated under the assumption of intradomain coherence, where different domains in a parent-child interaction (e.g., sensitivity) are measured through sub-domains aligned with that domain (e.g., Sensitivity is only defined by sensitive interactions), an approach that overlooks the often multidimensional nature of parent-child interactions (e.g., insensitive behaviours may also surface during sensitive interactions; Beebe et al., 2010; Fogel & Thelen, 1987; Lewis, 2000).

Perhaps most importantly, longitudinal observational studies of parent-child relationship, which are generally sparse (Bergmeier et al., 2015), remain strikingly absent from the displacement literature (while Lebigier-Vogel et al. 2022 briefly discussed some of their findings in relation to displacement and collected video data at two timepoints, they referred to their participants as immigrants and never specified if any were actually forcibly displaced); consequently, no observational coding system to date has undergone validation across multiple time points in displacement contexts. Without such longitudinal validation, it can be impossible to determine if an observational system is capturing enduring patterns in parent-child dynamics (Dishion et al., 2017).

1.3 Current study

In light of these gaps and the growing need for observational video data and coding systems necessary to analyse this data in the context of displacement, we conducted the first longitudinal behavioural observational study of displaced populations and developed a low-cost, paralinguistic-kinesic video coding system for ‘refugee parenting.’ Our aim is to test the reliability and validity of this coding system with displaced Syrians living in Jordan. At the time of data collection, Syrians represented the largest refugee population globally (UNHCR, 2021a), a substantial proportion of whom resided in Jordan (approximately 660,000 in 2021; UNHCR, 2021b). While substantially contributing to the rapidly growing field of ‘refugee parenting’ in specific and displaced persons more broadly, this study also offers valuable additions to the general parenting literature, particularly research dealing with the paralinguistic-kinesic aspects of parental communications, as well as parenting literature in low-resourced settings and among non-English speaking populations (Bozicevic et al., 2024; Colegrove & Havighurst, 2017; Mesman, 2021).

2. Methods

Before conducting this study, we obtained formal ethical approval from the Prime Minister's Office of Jordan and Trinity College Dublin (01E/2020/10). Recruitment and other procedural data were pre-registered (Hadfield, 2020), and all the other study materials pertaining to the observational coding system are available online on Open Science Framework (OSF; link, <https://osf.io/2q538/>), as is the analysis pre-registration for this manuscript. A summary of data collection details is presented in Figure 12, while a detailed description of the methodology is provided in each of the following sub-sections.

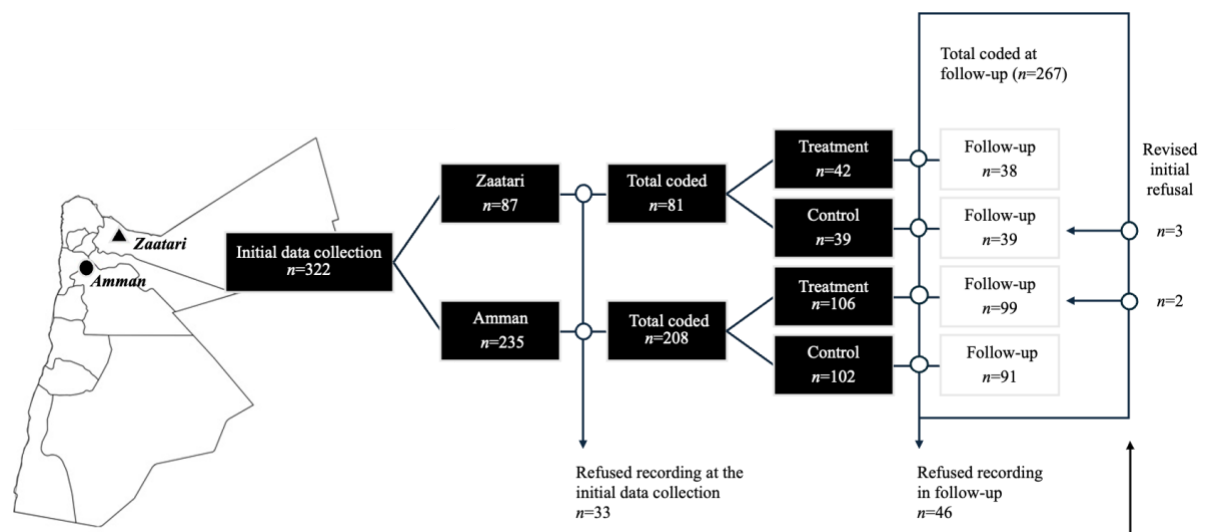


Figure 12. Overview of the Initial and Follow-up Data Collection Phases.

Note. 35 We coded videos using the Observational Behavioural System Evaluating Sensitivity and Synchrony (OBSESS). We followed up the same sample after four months. There is $n = 27$ who initially agreed to be recorded at the initial data collection then refused recording at follow-up.

* This is the same sample at four month follow-up data collection

2.1. Participants

For this video observational study, we draw our sample from a cohort study that includes Syrian refugee mothers ($n = 322$, $M_{age} = 32.61$ years, $SD = 7.02$) and their 4-8-year-old children ($n = 322$, 50% female, $M_{age} = 6.32$ years, $SD = 1.18$). Almost all mothers fled to Jordan between 2011 and 2014 (Hadfield et al., 2024), and as such, have been in displacement for about 10 years. Most lived in Amman ($n = 235$, 73.3%), with a smaller percentage living in the Zaatari refugee camp ($n = 87$, 26.7%), in line with broader trends for living situations among Syrian refugees in Jordan during the time of data collection (UNHCR, 2021b).

2.2. Study design

The main objective of the cohort study from which our video observational data was drawn was to evaluate ‘We Love Reading,’ a shared book-reading intervention by Taghyeer (Dajani, 2019), a nonprofit organisation founded in Jordan in 2010. This programme aims to improve child literacy and attitudes toward reading, with secondary goals of enhancing family relationships. To test the effectiveness of ‘We Love Reading,’ the cohort study employed a cluster randomised control trial (c-RCT) design: participants were assigned to either a ‘treatment group’ immediately receiving the intervention ($n = 154$) or a wait-listed control ($n = 168$) group. None had prior experience with the program.

Crucially, once the training was delivered, no statistically significant changes in mother or child behaviours for any primary or secondary outcomes were reported

(Hadfield et al., 2024)—note that the c-RCT had > 90% power to detect small-to-medium effects ($f = .20$, in line with Dowdall et al., 2020, meaning that the intervention's true effects were either vanishingly small or there were no effects at all; Hadfield et al., 2024). Because of this, our current study shifts focus from the intervention's impact to using the video data for a broader understanding of refugee family relationships. It may be also important to mention that the video recordings were separate from the 'We Love Reading' programme activities (the actual intervention took place on another day).

2.3. Eligibility criteria

For a family to be invited to the cohort study, the parents had to be Syrian, and the primary caregiver of a child aged 4 to 8 years (if multiple children fit the range, the child closest to 6 was selected). This age range offers both practical and conceptual advantages. Children of this age can generally engage in unstructured tasks, such as the toys used in our study, allowing for meaningful observation of parent-child interactions (Hein et al., 2020). This period also represents a significant milestone in cognitive, social, and emotional development, largely shaped by children's interactions with their caregivers and peers (Piaget, 2007). Exceptions to the sampling criteria in the cohort study included three children who were slightly outside the age range (all within one year of our criteria); we decided to keep this data in our analysis. All parent participants were biological mothers, with one exception—one grandmother was the primary caregiver because the child's biological mother had passed away; to simplify the reporting of our analysis, we treated this grandmother as we would a mother. With this in mind, everyone who participated in the cohort study, regardless of their assignment group, was invited to participate in these video observations. Most families agreed, with 287 (89% of the total)

consenting to be recorded, which constitutes a high participation rate in video observational research (Parry et al., 2016).

2.4. Timeline and recruitment procedure

The initial data collection period, which incorporated the first wave of dyadic video recording sessions, began in February 2021 and was completed in May 2021. Follow-up data collection, containing the second wave of dyadic video recording sessions, started in May 2021 and concluded in August 2021. Attrition was low, with only $n = 9$ mother-child dyads not participating in the follow-up four months later. To recruit Syrian refugee mothers, we relied on institutional knowledge and community networking; Taghyeer initially provided a list of organisations that support Syrian refugee mothers in Amman, whom we contacted to provide us with a list of eligible refugee mothers. The organisations which we drew our participants from included Sanabel Al-Khair ($n = 45$); Al-Yadoudeh Society ($n = 32$); Husun Al-Eman ($n = 35$); Yarub Charitiny ($n = 27$); and Nasamat Al-Farah ($n = 96$). In the Zaatari refugee camp, we relied on word of mouth; two ‘reading ambassadors’ who had previously volunteered with the ‘We Love Reading’ intervention provided a contact list of mothers who fit the criteria ($n = 47$ and $n = 40$, respectively). All our fieldworkers responsible for conducting video recording sessions spoke Arabic fluently and, in addition to having previous experience collecting data, underwent on-site data collection training, which included the piloting of data collection procedures with 10 dyads.

2.5. Home-visit protocol

Mothers who expressed interest in learning more about the project received another call from our fieldworkers, who provided more detail and an explanatory sheet via WhatsApp (including the research background, what their participation would entail, etc.).

Household visits were arranged for those who agreed to participate. On the day of the visit, two fieldworkers would introduce themselves and set up the environment, ensuring privacy and optimal conditions for recording. This involved preparing books and setting up toys for the video recording session, which involved play and book activities. All participants were intended to do both book and play activities. Recording of mother-child during play and book activities was done subsequently to the survey data collection and the experiments (these were not part of this study and included computerised tasks designed to measure emotion and cognition). Doing so provided the participants with more time to become familiar with the project, therefore promoting a more natural mother-child interaction to be captured by the video recordings (Gardner, 2000). At the start of every recording, one fieldworker explained the activity instructions to both the mother and child, while another handled the technical aspects, such as setting up the camera and tripod, displaying the identification paper to the camera, and starting the timer (play activity ~5 minutes, and book activity ~10 minutes). In most cases, fieldworkers left the room; however, when that was not possible due to the lack of available rooms ($n = 2$), fieldworkers stayed without interfering in the interaction. Fathers were only present rarely ($n = 3$), and most sessions went uninterrupted (only $n = 16$ were interrupted, e.g., another family member walking into the room). Our protocol on the OSF page included specific instructions for handling interruptions.

2.5.1. Materials and activity structure

The mother and child were given options for books and toys in each activity. In the book reading task, the mother and the child were asked to choose one of two books—either “Questions in a Traveling Bag,” about a child dealing with refugee experiences, or “The Words Cook,” about a child named Mayyada who navigates loneliness in kindergarten and social interactions with classmates (Dawood, 2020; Hayat, 2015). The mother was

instructed to use the book with the child in any way for up to 10 minutes, which could have included only looking and discussing the pictures in the book. In the play activity, the mother and child were given a set of age-appropriate toys, including a Duplo® and a Meadow Farm set, and were instructed to play as they usually would for 5 minutes. Our fieldworkers and Taghyeer staff, who regularly work with Syrian refugee families, selected both the toys and the books to ensure they were culturally appropriate (e.g., many Syrian participants come from rural backgrounds, making the farming theme and tractor in the Meadow Farm highly relatable).

2.6. Observational coding system

We have developed a coding system called OBSESS (Observational Behavioural System for Evaluating Sensitivity and Synchrony) to assess the quality of interactions between refugee mothers and their children (Figure 13). Our system draws inspiration from the Coding of Attachment-Related Parenting (CARP; Matias, 2006) and its short version, the S-CARP (Bennetts et al., 2021), both of which offer openly available training materials and do not require specialised software or incur training fees. Other systems, including ones used with refugee or displaced populations (see Table 4), were often limited by restricted access to training materials, the need for specialised software, or costly training requirements, among other limitations. While CARP and S-CARP provided a helpful starting point, we eventually made substantial adaptations to better suit the context of displacement, given that these systems were originally designed for English-speaking mother-child dyads in high-income countries.

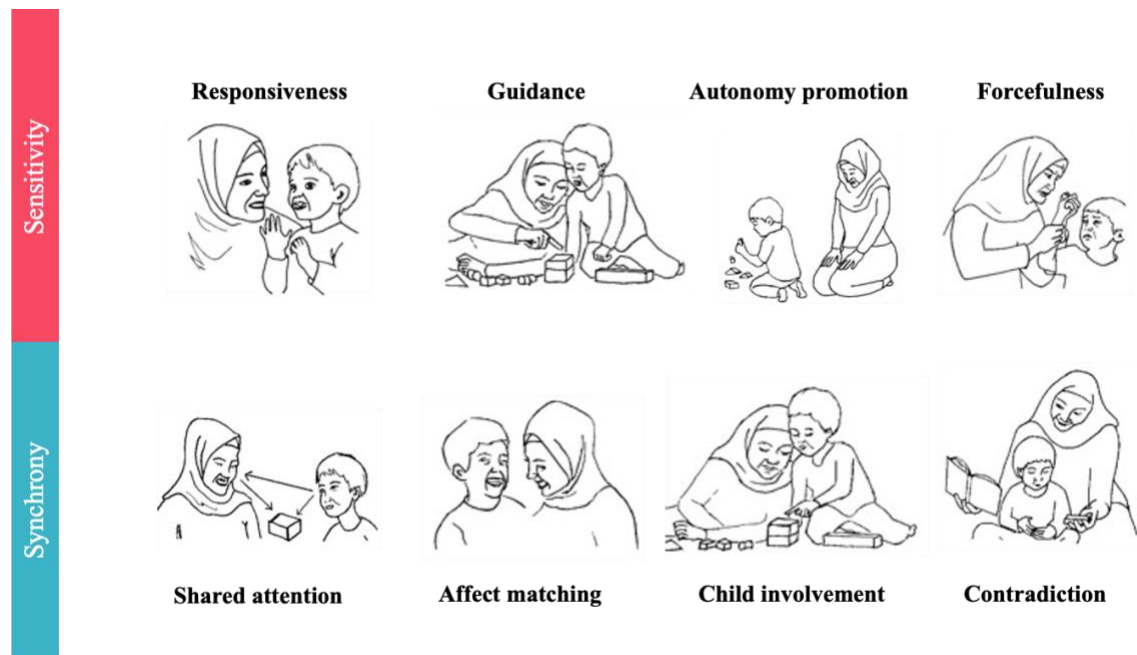


Figure 13. The Observational Behavioural System Evaluating Sensitivity and Synchrony (OBSESS).

Note. The total score for each domain (sensitivity and synchrony) can range from 1 to 9; ‘forcefulness’ and ‘contradiction’ are reverse scored.

Table 4

OBSESS Compared to Other Systems Validated on Displaced Mother-Child Dyads

Study authors	Coding system	Key measures	Sample details	Reliability	Language dependency	OBSSSES advantages
Dalgaard et al., 2020	Marschak Interaction Method Rating System (MIMRS)	Structure, challenge, engagement, nurture	Refugee dyads from 7 countries in Denmark ($n = 21$)	5 videos double coded	Dependent, required translation	Open-source training; measures synchrony; no translation needed

Hein et al., 2020	Combined criteria from two different sources (Bernard et al., 2013; Matas et al., 1978)	Sensitivity, intrusiveness, positive regard, disengagement, synchrony	Refugee dyads from Syria in Germany ($n = 104$)	104 videos double coded	Dependent, required translation	Open-source training; no translation needed
Lebiger-Vogel et al., 2022	Emotional Availability Scales (EAS)	Sensitivity, structuring, non-intrusiveness, non-hostility	Unspecified who is displaced; total sample originate from 37 countries residing in Germany ($n = 118$)	Unspecified	Unclear*	Open-source training; measures synchrony; no translation needed
Scharpf et al., 2023	Emotional Availability Scales (EAS)	Sensitivity, structuring, non-intrusiveness, non-hostility	Refugee dyads from Burundi in Tanzania ($n = 222$)	20 videos double coded	Independent due to methodological constraints**	Open-source training; measures synchrony; validated specifically for nonverbal communications

Note: Training protocols for all observational coding systems here either did not clearly explain their training material and sources or explained that the training was centralised by an institute or persons acting as the lead or master coder.

* Sensitivity is explained to include verbal expressions but does not explicitly state whether translations were needed for coding or if the same language as the participants was used.

** Coders focused on nonverbal aspects of the interactions due to the noisy environment of the refugee camps and the generally low level of verbal communication observed in the mother-child dyads.

van Ee et al. (2013) and Wiepkema (2013) both used the EAS observational coding on refugee samples, although they coded fathers, not mothers; hence why they were not included in this table.

2.6.1. Sub-domains and scoring

To evaluate the sensitivity domain of refugee mother-child interactions, OBSESS measures four key aspects: responsiveness, guidance, autonomy promotion, and forcefulness. Responsiveness captures how well the mother follows the child's lead, such as turning a page when the child attempts to do so. Guidance assesses the mother's efforts to structure the interaction, regardless of the child's response. For instance, setting up toys or showing something to the child counts as guidance even if the child does not react. Autonomy promotion focuses on how the mother encourages independent behaviour and decision-making in the child. Forcefulness, which was reverse scored, gauges control that might hinder the child's independence and responsiveness, such as preventing the child from turning a page or taking away an item the child is still interested in. These sub-domains align closely with established parenting elements in the broader literature (i.e., forcefulness/intrusiveness, guidance/structure, autonomy promotion/support; Aspland & Gardner, 2003; Matias, 2006).

The synchrony domain is also measured across four sub-domains: shared attention, affect matching, child involvement of the mother, and contradiction. Shared attention measures mutual engagement in activities. Affect matching observes the alignment of positive emotional states between mother and child. Child involvement of

the mother assesses the child's openness to the mother in interactions, focusing on opportunities the child provides for the mother to join the task, regardless of the child's intent. Contradiction, which was reverse scored, identifies disruptions in the natural interaction flow from the mother's side, such as shuffling between toys or indicating they will give their child a toy but then keeping it. Similar to above, these sub-domains map onto widely recognised parenting constructs in the literature (i.e., shared attention/ mutual attention, affect matching/emotional attunement, child involvement/reciprocity, contradiction/interactive mismatch; Davis et al., 2017; Leclère et al., 2014).

Our coding process quantifies these behaviours using a frequency-based system. Behaviours are categorised as 'none' (0 occurrences), 'rarely' (1-3 occurrences), 'about half of the time' (4-6 occurrences), or 'most of the time' (more than 6 occurrences). This frequency system was calibrated based on observed norms within our specific sample from the initial 100 videos we coded (to be described more in the following section), as standard norms do not exist for these behaviours. In the final scoring, 'none' is scored as 0, 'rarely' as 1, 'about half of the time' as 2, and 'most of the time' as 3. The total score for each domain (sensitivity or synchrony) can range from 0 to 9, calculated from these sub-domain scores. For the majority of families who were coded on both play and book activities, their scores were averaged—this was done in recognition of the fact that observational data is 'situationally sensitive' (Krijnen et al., 2023; Mathers et al., 2024); as such, by considering multiple contexts (both play and reading), we can get a more representative picture of a given parent-child dynamic.

2.6.2. *Design and training*

Some participating mothers were not comfortable having male researchers view their videos. This meant our team lead QK, who was male, could not watch the recordings. As

a result, we could not use the ‘master coder’ approach common in some observational research, where one experienced researcher watches most videos to develop training protocols and ensure consistent scoring across the study. Instead, two volunteer female research assistants (BV and CO, who do not understand Arabic) with backgrounds in developmental psychology were trained to code $n = 100$ recorded observations of refugee mother-child interactions, while QK, who did not view the videos, structured their observational feedback into a codebook that involved detailed instructions and descriptive examples on which behaviours constituted a particular sub-domain, and which behaviours did not. The codebook material can be accessed via the web application, Notion (<https://qusaikhraisha.notion.site/4894e2631d6e4f209ec8f60d0fdef16d?v=037885c3eca94d60966137cd84dd6901>).

To detail the training process of coders, BV and CO first received a reading list of academic papers relating to parenting and displacement’s impact on parent-child relationships (Bronfenbrenner, 1996; Bryant et al., 2018a; Hein et al., 2020; Hodes & Vostanis, 2019; Kuppens & Ceulemans, 2019). Building on his experience at the United Nations High Commissioner for Refugees (UNHCR) as a cognitive interviewer for highly sensitive populations, QK designed and delivered an extensive training on cross-cultural non-verbal behavioural analysis. The training program emphasised awareness of cognitive and emotional biases that could influence perception of behaviour (Azriel & Bar-Haim, 2020; Harvey, 2007). Coders were introduced to the cognitive bias codex (Manoogian & Benson, 2018) and were advised to focus on objective, observable behaviours rather than subjective interpretations of parental beliefs or intentions behind a behaviour (Gardner, 2000; Putnick, 2019). One exercise involved analysing a zoomed-in image of a tense face, which was interpreted as angry by most coders in training; when shown the full context of the image as a winning moment in a sport match, coders

understood how ambiguous information can lead to misinterpretation and inaccurate labelling.

After this initial training, BV and CO began independent coding sessions and discussed any divergent interpretations with QK, with this process leading to refinement in coding guidelines and instructions (this iterative process is illustrated in Figure 14). Within this iterative process, we also introduced new categories for ‘forcefulness’ and ‘contradiction’ to capture interactions that were previously unaccounted for. We established a frequency scale (rarely, about half the time, most of the time) for these and all other behaviours and added options for ‘none’ to categorise unobserved actions. By the time the 100 videos were coded, the coders were agreeing with each other almost every time (we did not perform a statistical test at this stage but rather relied on consensus through discussions), and they generally found the coding manual to be appropriate in capturing the observed behaviour. From here, QK codified the set of instructions that yielded high agreement between BV and CO into a codebook, which involved a description of different coding scenarios.

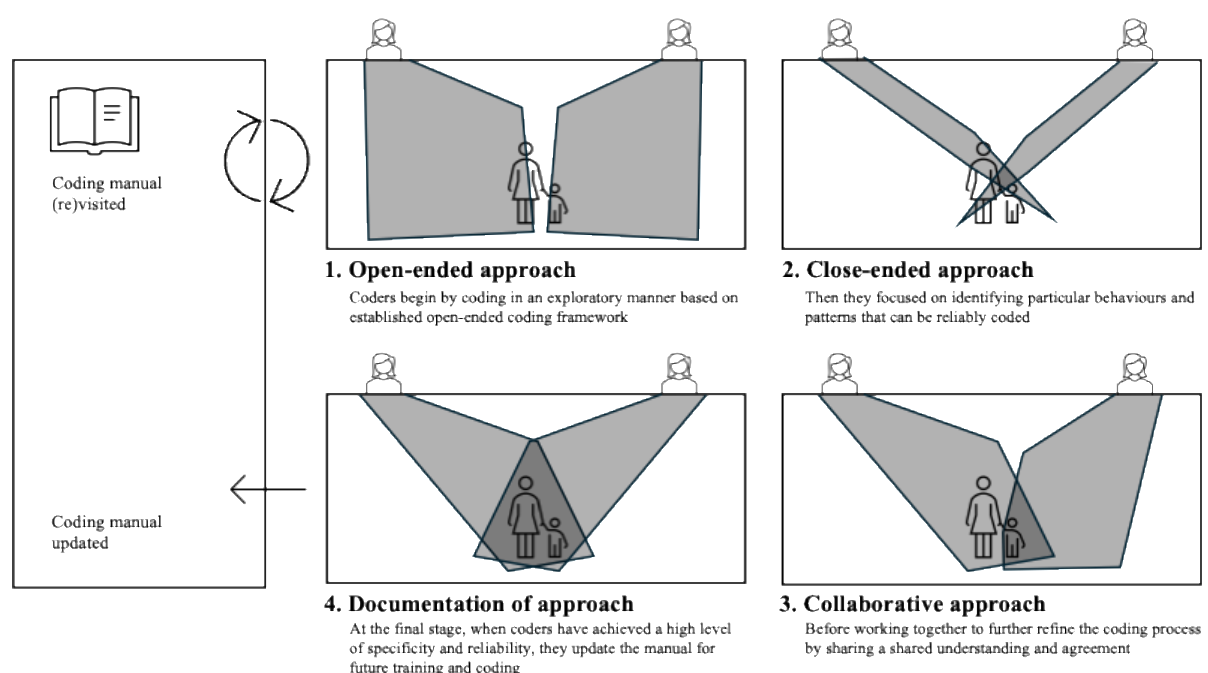


Figure 14. Decentralised Coding of Sensitivity and Synchrony in Refugee Mothers: A Visual Framework

We subsequently recruited and trained 18 female international coders (two understood Arabic, the language of the participants). Most were undergraduate students, and none had any previous knowledge or experience in coding or observational research. New coders completed the prescribed background readings on displacement and parenting (the same completed by BV and CO), along with other documents on research ethics and how to handle sensitive data. They were also given a similar training on behavioural assessment and cross-cultural considerations, followed by at least 20 hours of hands-on training sessions conducted over four weeks. The training protocol required coding 20-60 practice videos which had already been completed by experienced coders, with advancement contingent upon specific error thresholds. Examining a set of five videos, coders demonstrating errors across three or more sub-domains (e.g., two videos had inaccurate assessment of guidance, autonomy, and forcefulness) continued practicing with complete videos until meeting the threshold. Those with errors in one or two sub-domains engaged in targeted practice within those specific sub-domains until achieving an error rate of ≤ 1 per sub-domain per video. An error was defined as any coding discrepancy from the ‘gold-standard’ benchmarks established by multiple coders who finished training.

Once training was completed, coders met weekly to review scores and address discrepancies to reduce “coder drift” or the gradual change in coding from the standard agreed upon in training (Forrer et al., 2018). By the time of writing, we had coders from the Bahamas, China, Germany, India, Ireland, Italy, Jordan, Luxembourg, Poland, Spain, Turkey, the United Kingdom, and the United States. All coding was done online using a

simple survey on Qualtrics. Further details on our training process, material, and manual are available online on our OSF page.

It is important to clarify that linguistic background, specifically Arabic language proficiency, was not used as a coder selection criterion or analytical covariate in the coding process. All coders were able to listen to the audio while coding but were instructed not to consider language content in the coding. While two of our coders understood Arabic, during the training program, they were required to demonstrate inter-rater reliability that met stringent matching criteria with their peers (who had minimal to no exposure to Arabic language), and that included perfect or near-perfect agreement with non-Arabic speaking coders (error rate of ≤ 1 per sub-domain per video).

2.6.3. Reliability and validity

We pre-registered our analysis to evaluate the reliability and validity of the OBSESS coding system ([link, osf.io/gcv5z/?view_only=8cb81ef7dd754217a36212082018d324](https://osf.io/gcv5z/?view_only=8cb81ef7dd754217a36212082018d324)). Systematic variability was first assessed using a paired-sample *t*-test to compare single coder ratings of Sensitivity or Synchrony against a benchmark set by multiple coders (we supplemented this analysis with Bland-Altman plots; a visual method that can reveal any systematic patterns of disagreement between ratings that the *t*-test may not highlight). Second, we assessed for random variability by examining through Cohen's and Weighted Kappa, with a target agreement of .53 or higher (Cohen, 1960; Sun, 2011). Third, to measure coder variability (subjectivity), we used the Intraclass Correlation Coefficient (ICC), expecting an ICC above .5, indicating low subjectivity in coding.

Fourth, we investigated within-coder variability (accuracy), the extent to which the same coders agreed with themselves, by conducting a paired-sample *t*-test for coders who accidentally coded the same videos twice. Fifth, we analysed the time variability

(stability) of the coding scheme by comparing videos coded at the initial data collection and follow-up through a paired-sample *t*-test (to account for intraindividual stability of scores, we conducted a Bland-Altman test, while a Pearson's correlational analysis was undertaken to understand rank order differences; these two analyses were not pre-registered). Note that we first coded the initial videos and then the follow-up ones (where possible, the same coder was used to code both).

To validate the system, we set out to examine its correlation with parenting and parent-child relationship surveys, including the Pianta Child-Parent Relationship Scale (measured at both initial and follow-up stages; Pianta, 1992; Driscoll & Pianta, 2011), which consists of 15 items on a 5-point Likert scale. This scale includes measures of mother-child closeness (7 positive items, scores ranging from 7 to 35) and conflict (8 negative items, scores ranging from 8 to 40). While data for the Pianta scale was collected from participating mothers at two time points, only at follow-up we administered the Alabama Parenting Questionnaire (short-form APQ; Elgar et al., 2007) to measure positive parenting, poor supervision, and inconsistent discipline, which was done on a 5-point Likert scale (range of scores between 3 and 15). We hypothesised that the OBSESS measures of Sensitivity and Synchrony would positively correlate with positive parenting and closeness and negatively correlate with conflict, poor supervision, and inconsistent discipline.

Correlational analyses were conducted using IBM SPSS Statistics for Mac, version 29.0.2.0 (IBM Corp., Armonk, NY). Data visualisations and interrater tests were performed using R version 4.4.2 on RStudio, version 2024.04.2+764 (Posit, PBC, Boston, MA).

3. Results

3.1 Reliability

We analysed 1,092 videos using OBSESS across two time points ($\approx 10\%$ of our sample refused to be recorded, see Supplementary Table 4 for an overview of their characteristics). The initial phase included 564 videos (281 play activities, 283 book-reading activities), while the follow-up phase four months later comprised 528 videos (261 play activities, 267 book-reading activities). When comparing each coder scores (coding done individually) against the ‘benchmark’ (videos coded by multiple experienced coders), we found no significant differences ($t = -.11$, $df = 229$, $p = .90$). This suggests weak evidence for systematic deviation in the coding process (a finding consistent with the visual Bland-Altman analysis, see Supplementary Figure 1a).

Disagreement between coders did occur (Cohen’s Kappa $\kappa = .41$), but their magnitude was low, with most disagreements being limited to single-point differences (Supplementary Figure 1b). This was confirmed by the Weighted Kappa ($\kappa = .74$), which uses squared weights to account for the ordinal nature of the ratings and provide partial credit for near-adjacent ratings (the ICC score was .75, further confirming this finding). These intercoder agreement scores are higher than seen elsewhere in the literature (e.g., Lembcke et al., 2020), reflecting strong inter-rater reliability. Additional reliability checks came from instances where coders coded the same videos twice. Our analysis indicated low within-coder variability and minimal impact of potential coder fatigue or drift for subdomain coding ($t = -0.350$, $df = 175$, $p = 0.726$; note the sample size had excellent power to detect medium effects $d = 0.5$, power $> 99.9\%$ and large effects $d = 0.8$, power = 100%). Controlling for chance agreement, within-coder agreement was $\kappa = .76$, while weighted kappa was even higher at $\kappa = .88$, indicating excellent within-domain consistency.

Furthermore, OBSESS yielded robust stability (Figure 15 and Figure 16). Sensitivity scores averaged 6.91 ($SD = .98$) at initial collection and 7.30 ($SD = 1.05$) at follow-up. While this difference was statistically significant ($t(261) = -5.09, p < .001$, 95% $CI [-.54, -.24]$), it represented only a 4.33% change on our 9-point scale. These Sensitivity scores align with those reported in other non-Western samples using 9-point scales (7.33; Fourment et al., 2021) and are comparable to scores observed among securely attached children in Western samples (Thompson, 2021). Synchrony demonstrated even greater stability, with no significant change between initial ($M = 5.07, SD = 1.22$) and follow-up phases ($M = 4.98, SD = 1.05; t(261) = 1.15, p = .252$, 95% $CI [-.06, .24]$), mirroring a similar scoring pattern found in the 9-point Synchrony Global Coding System (3.49; Skuban et al., 2006).

In terms of longitudinal stability at the individual level (as opposed to the group-level averages reported above), our results show that mothers who scored higher in Sensitivity and Synchrony at the initial assessment maintained their relative standing at follow-up both for Synchrony ($r = .388, p < .001$) and for Sensitivity ($r = .251, p < .001$). A visual inspection of the Bland-Altman plots, which is recommended when assessing reliability over time (Aldridge et al., 2017), demonstrate that the scores for each family stayed fairly consistent—the majority of parents showed smaller changes of less than ± 2.5 points, an amount of variation smaller than what other researchers typically find in similar studies (Bennetts et al., 2016).

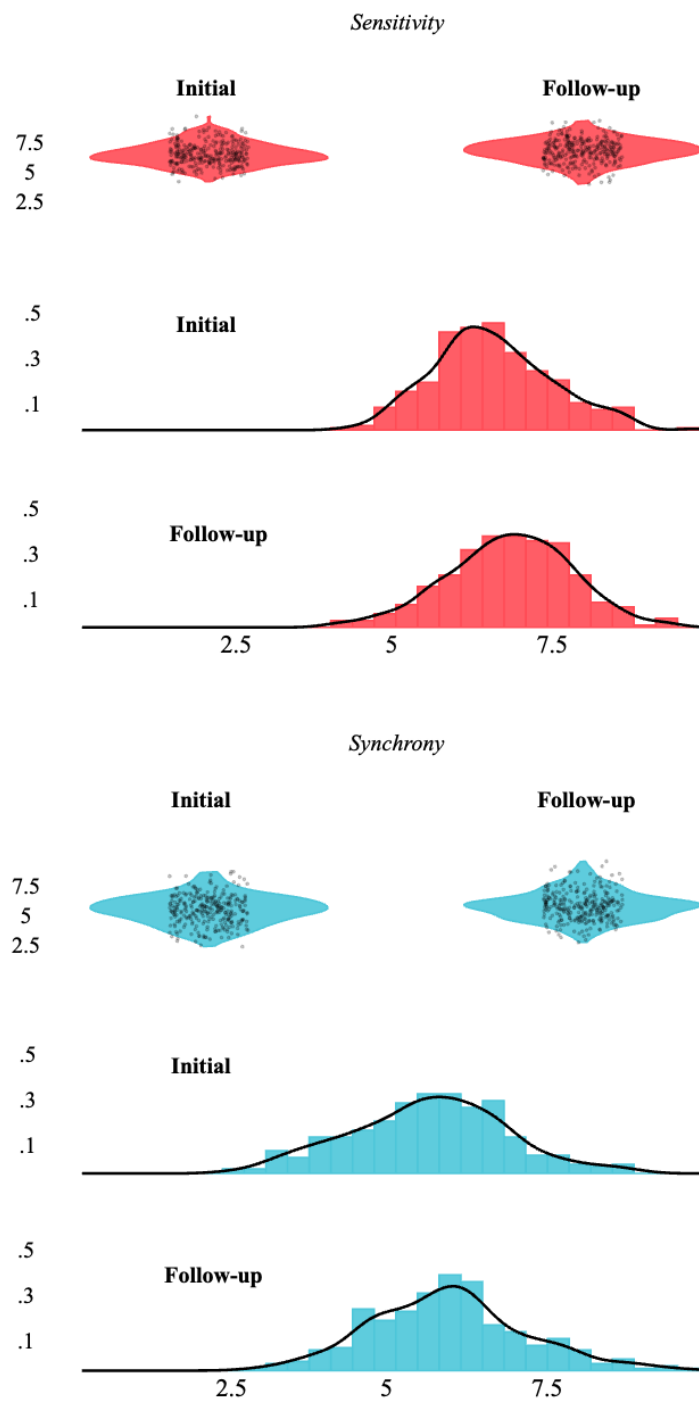


Figure 15. Distribution of Sensitivity and Synchrony at Initial and Follow-up Data Collection Phases.

Note. The total score for each domain (sensitivity or synchrony) can range from 0 to 9.

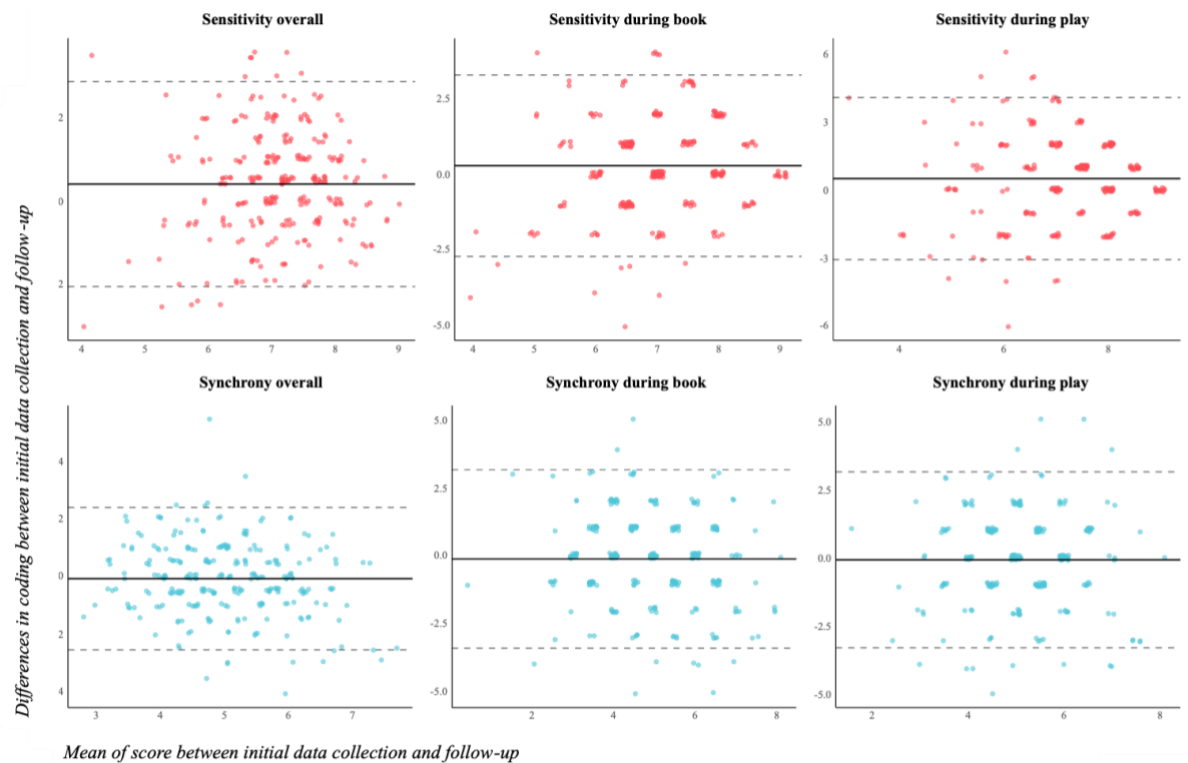


Figure 16. A Bland-Altman Plot Visually Representing Systematic Variability in Coding.

Note. The dashed red lines represent the upper and lower limits of agreement, calculated as the mean difference ± 1.96 times the standard deviation of the differences: points falling outside these limits may indicate significant disagreements between the raters, reflecting potential outliers or areas of systematic bias in the observational process. The x-axis, representing the average score between the coders, provides a reasonable estimate of each item's 'true' value.

3.2 Validity

When self-reporting, nearly all mothers gave themselves maximum or near-maximum scores on closeness with their children (averaging 33 out of 35 points). Except for the conflict subscale administered at follow-up ($\alpha = .717$), all the other Pianta subscales demonstrated low internal consistency ($\alpha < .70$). Similarly, the Alabama parenting questionnaires had low internal consistency: positive parenting subscale had the highest internal consistency that still was subpar ($\alpha = .629$) compared to inconsistent discipline

domain ($\alpha = .418$) and poor supervision domain ($\alpha = .580$). However, even this ‘better’ measure suffered from the same ceiling effect problem: most mothers scored at or near the maximum (most scores clustered together between 13-15). Given these substantial psychometric limitations, we focused our analysis on the Pianta conflict subscale at follow-up, which demonstrated adequate reliability (results from the other measures are provided in the Supplementary materials, Table 5). While the Pianta conflict did not relate to observed Synchrony ($r = -.096, p = .129$), the subscale showed an association with Sensitivity ($r = -.154, p = .014$).

In contrast to the problematic self-report measures, our observational assessments of maternal Sensitivity and Synchrony demonstrated strong psychometric properties and meaningful patterns of association. Within the same context, Sensitivity and Synchrony predicted each other strongly (book initial: $r = .37, p < .001$; follow-up: $r = .39, p < .001$; play initial: $r = .35, p < .001$; follow-up: $r = .48, p < .001$). Both Sensitivity (initial: $r = .13, p = .026$; follow-up: $r = .15, p = .015$) and Synchrony (initial: $r = .26, p < .001$; follow-up: $r = .059, p = .342$) also showed modest correlations across play and book-reading contexts (Figure 17). When averaging across both play and book-reading activities (Figure 18), Sensitivity strongly predicted Synchrony at both initial ($r = .40, p < .001$) and follow-up assessments ($r = .44, p < .001$). These findings reveal an important insight: the strong within-context correlations between Sensitivity and Synchrony, contrasted with weaker cross-context correlations, suggest how parenting behaviours are situation-specific—an important aspect of parenting that was overlooked in our self-report.

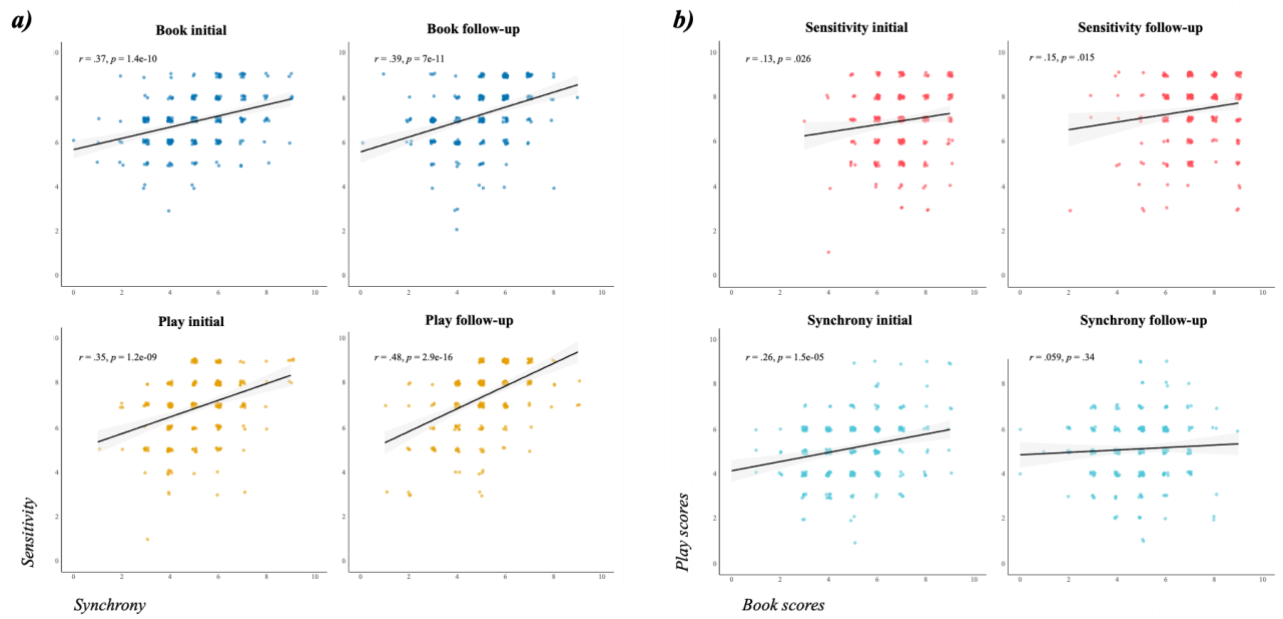


Figure 17. Sensitivity and Synchrony Correlations Across the Same (a) and Different Contexts (b).

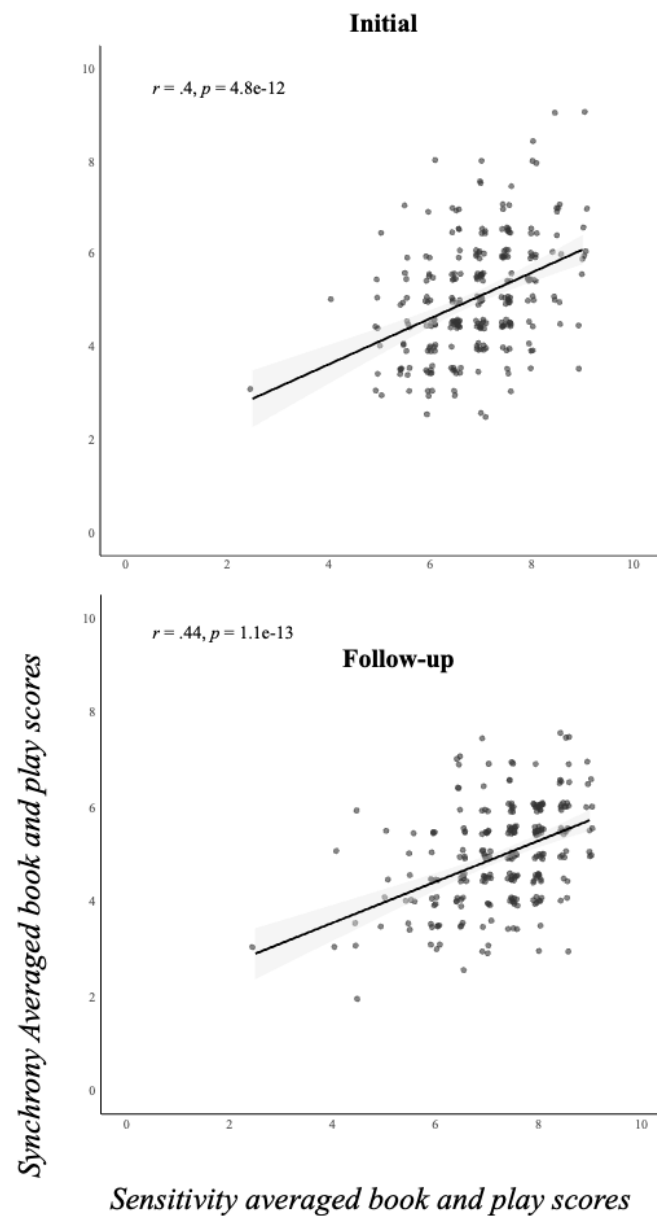


Figure 18. Correlations of Sensitivity and Synchrony Averaged Across Contexts at Initial Data Collection and Follow-up

4. Discussion

Our study introduces OBSESS, a novel observational coding system designed to assess refugee parent-child paralinguistic-kinesic interactions in low- to middle-income countries, where most displacement occurs (UNHCR, 2024). The linguistic diversity of

displaced populations makes coding lexical aspects of words prohibitively resource-intensive, requiring extensive teams of multilingual coders. This barrier makes minority refugee groups particularly difficult to observe and code. In this study, we demonstrate that coding paralinguistic-kinesic elements can circumvent language barriers while capturing meaningful refugee caregiving dynamics. Paralinguistic communication represents a rich dimension of parental interaction across many societies (Abdullah et al., 2021), while kinesiological elements of parenting constitute a substantial portion of caregiving interactions in low- to middle-income countries (Mesman et al., 2021).

However, we recognise that parent-child communications are inherently multimodal (verbal and non-verbal) and multidimensional (lexical and non-lexical), with *all* elements working in concert together to shape a parental interaction. Our goal in developing OBSESS, therefore, is not to undervalue the lexical elements of parental interactions in displacement contexts, but to spotlight the remarkably understudied but important paralinguistic-kinesic parental behaviours that surround verbal communications; unlike verbal content, these aspects are less conscious and thus less easily captured through surveys or interviews (Ghandour et al., 2024).

By validating OBSESS, we form a critical bridge to the rich developmental, clinical, and psychobiological scholarship on nonverbal behaviours conducted among non-displaced people to parenting research with displaced populations (Colegrove & Havighurst, 2017; Bennett et al., 2024). This literature may be especially useful in understanding how war-related stressors affect automatic, less conscious aspects of parenting (Dalgaard & Montgomery, 2015; Fredman et al., 2017; Measham & Rousseau, 2010). It could lead to a greater understanding of intergenerational trauma transmission in protracted displacement (Scharpf, Kaltenbach, et al., 2021), and may also enhance our

understanding of how the persistent 'macro' and 'daily' stressors influence 'refugee parenting' behaviours (Khraisha, Abujaber, et al., 2024; Miller & Rasmussen, 2017).

2.1 Strengths and limitations

Our study represents the only longitudinal study, as well as the most extensive cross-sectional analysis of video-based observational methodology to analyse parent-child interactions within a forcibly displaced population (Dalgaard et al., 2020; Hein et al., 2020; Lebiger-Vogel et al., 2022; Lembcke et al., 2020; Scharpf et al., 2023; van Ee et al., 2013; Wiepkema, 2013). Beyond scope, we demonstrate two important methodological advances. At the data collection stage, we show how large-scale video recordings and observational coding systems can be executed successfully among refugee communities that have reportedly been uncomfortable with cameras in other studies (e.g., Lembcke et al., 2020). When it came to coding the videos, we successfully 'decentralised' the coding approach, which was beneficial in reducing the common cultural bias arising from having one viewpoint dominate the coding system (Bozicevic et al., 2024; Colegrove & Havighurst, 2017).

As a result, our coding reliability was robust: there was minimal coder bias, substantial agreement when accounting for the degree of disagreement, and high consistency, accuracy, and stability over time. These findings echo those of well-established observational measures like the Emotional Availability Scales (Biringen et al., 2014) and the CARE-Index (Crittenden, 1981), with the crucial difference of OBSESS being the only validated, paralinguistic-kinesic observational coding system. This distinction offers important practical advantages: words are expensive to code and difficult to compare meaningfully across populations. Meanwhile, OBSESS's paralinguistic-kinesic design allows for relatively inexpensive implementation and

standardisation across diverse populations, bypassing the cultural homogeneity fundamental in other traditional coding systems (consequently enabling more straightforward cross-population comparisons; Dalgaard et al., 2020; Hein et al., 2020; Lebiger-Vogel et al., 2022; Lembcke et al., 2020; van Ee et al., 2013; Wiepkema, 2013).

In terms of validity, OBSESS measured meaningful aspects of ‘refugee parenting,’ with construct and concurrent validity being supported by multiple lines of evidence. This includes our scoring distributions and theoretical rationales, which show considerable consistency with established research in the field (Fourment et al., 2021; Skuban et al., 2006; Thompson, 2021). Furthermore, inter-correlations between Sensitivity and Synchrony (strong within-context, weak across-context) provide evidence for discriminant and convergent validity. Additionally, the parenting behaviours OBSESS captures are context-specific, which aligns with ‘real-life’ everyday parenting (Mathers et al., 2024), suggesting evidence for ecological validity. Lastly, as hypothesised, the Sensitivity component in OBSESS held a negative association with mother-reported, mother-child conflict.

A notable pattern emerged in our psychometric analyses: instruments measuring positive parenting (such as mother-reported child closeness and positive parenting) consistently exhibited ceiling effects, whereas measures of negative concepts (mother-reported child conflict and inconsistent discipline) showed normal skewness levels. This is a systematic response bias rather than random measurement error. While social desirability bias and Dunning-Kruger effects partially explain such patterns (Scott, 2020; Van de Mortel, 2008), they cannot fully account for why mothers avoided ‘flooring effects’ when reporting conflict (why mothers did not overwhelmingly report low conflict with their child). We propose that the discrepancies we observed may reflect two distinct phenomena. First, some refugee mothers may indeed be employing appropriate parenting

strategies—yet still experience significant parent-child conflict due to child difficulties that are unrelated to parenting quality, such as trauma-related behavioural challenges or developmental disruptions from displacement experiences. In these cases, high conflict reports may coexist with genuinely positive parenting behaviours. Second, other refugee mothers may be parenting in ways that are less effective but still respond positively on surveys to preserve their dignity. For these mothers, ceiling effects on positive parenting measures may mask caregiving difficulties.

The dignity-preserving responding style may be explained in our refugee sample using linguistic and ecological factors. Arabic has a much richer vocabulary than English (Abu-Errub et al., 2014; Hussein, 2015), and all languages have a bias for positivity (Boucher & Osgood, 1969; Dodds et al., 2015). When combined, these linguistic features create significant interpretive challenges for survey research translated into Arabic: positive parenting terms become particularly ambiguous in Arabic, allowing for more meanings. As parenting surveys are administered in refugee contexts, where maintaining family dignity sits at the heart of parenting well in displacement (Elias et al., 2025; Khraisha, Sawalha, et al., 2024), some Arab refugee mothers may interpret these positively-worded parenting questions not as neutral assessments, but as fundamental challenges to their identity as good parents. This observation is theoretically significant because it suggests that binary positive-negative frameworks embedded in ‘refugee parenting’ research become internalised standards that participants feel compelled to defend (Yashika & and Prakash, 2025). It is also why OBSESS can be especially valuable for the study of ‘refugee parenting.’ It circumvents issues with reliable reporting of positive parenting, which is particularly valuable given that ‘refugee parenting’ literature in protracted situations severely lacks documentation of positive parenting (Khraisha, AbuJaber, et al., 2024).

2.2 Future directions

OBSESS paralinguistic-kinesic coding system opens up avenues for new research and intervention that could cut through cultural and linguistic barriers among diverse refugee populations, including the majority living in protracted refugee situations. As of today, advancements in artificial technology (AI), specifically in the areas of neural deep learning open exciting possibilities for scalability; there are already AI software designed to detect basic non-verbal cues (Baur et al., 2013; Terhürne et al., 2022). Such automation could make large-scale video observational studies using OBSESS in the context of war and displacement in resource-constrained settings even more feasible.

5. Conclusion

OBSESS represents a valuable addition to the toolkit available for researchers studying families affected by war and forced displacement. Through its focus on the paralinguistic-kinesic elements of parental interactions, OBSESS offers a novel approach to understanding mother-child dynamics in situations where language barriers or cultural differences may impede traditional assessment methods. This focus also eliminates the need for costly and time-consuming translation processes or the recruitment of multilingual coders when coding linguistically diverse displaced populations. Its free-to-use nature democratises access to observational coding systems, enabling a broader range of researchers, including those in resource-constrained settings, to contribute meaningfully to the field. Taken together, OBSESS has the potential to offer a new window for understanding parent-child interactions in context of war and displacement, perhaps showing that the most revealing forms of communication can happen around the meaning of words.

Chapter 5: Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages

Khraisha, Q., Put, S., Kappenberg, J., Warraitch, A., & Hadfield, K. (2024). Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. *Research Synthesis Methods*, 15(4), 616–626.
<https://doi.org/10.1002/jrsm.1715>

As Chapter 2 highlighted the voluminous nature of research related to ‘refugee parenting,’ Chapter 4 discussed the issue of expensive proprietary observation systems as a significant reason for the limited applicability of video observation in the ‘refugee parenting’ literature, which is why OBSESS is made free and open-access. In this Chapter, I address the issue of cost associated with screening and extracting data from voluminous literatures such as that relating to ‘refugee parenting.’

This chapter is an exact copy of the published journal paper.

Abstract

Systematic reviews are vital for guiding practice, research, and policy, though they are often slow and labour-intensive. Large language models (LLMs) could speed up and automate systematic reviews, but their performance in such tasks has yet to be comprehensively evaluated against humans, and no study has tested GPT-4, the biggest LLM so far. This pre-registered study uses a ‘human-out-of-the-loop’ approach to evaluate GPT-4’s capability in title/abstract screening, full-text review, and data extraction across various literature types and languages. Although GPT-4 had accuracy on par with human performance in some tasks, results were skewed by chance agreement and dataset imbalance. Adjusting for these caused performance scores to drop across all stages: for data extraction, performance was moderate, and for screening, it ranged from none in highly balanced literature datasets ($\sim 1:1$) and moderate in those datasets where the ratio of inclusion to exclusion in studies was imbalanced ($\sim 1:3$). When screening full-text literature using highly reliable prompts, GPT-4’s performance was more robust, reaching ‘human-like’ levels. While our findings indicate that, currently, substantial caution should be exercised if LLMs are being used to conduct systematic reviews, they also offer preliminary evidence that, for certain review tasks delivered under specific conditions, LLMs can rival human performance.

1. Introduction

Systematic reviews are crucial in advancing practice, research, and policy (Aromataris et al., 2015). However, the current approach to systematic reviews is laborious and can be slow to the point that the resulting synthesis of knowledge may no longer be up to date when it is completed (Borah et al., 2017; Michelson & Reuter, 2019). The explosion of scientific literature, coupled with the complexity and specificity of many research questions, further adds to these challenges (Fiorini et al., 2018). Artificial intelligence (AI) has emerged as a potential solution, with recent studies and evaluations suggesting its capability to enhance the quality and efficiency of systematic reviews (Blaizot et al., 2022; Dijk et al., 2023; Kebede et al., 2023; Mahuli et al., 2023; Moreno-Garcia et al., 2023; Nugroho et al., 2023; Santos et al., 2023).

Some examples of AI tools that have been used in systematic reviews include Rayyan and Abstracker, which help with screening titles and abstracts (Giummarra et al., 2020; Rogers et al., 2020); trialStreamer, which helps with data extraction from full-text articles (Marshall et al., 2020); and RobotReviewer, which helps with assessing study quality and bias (Goldkuhle et al., 2018). The major shortcoming of these tools is that their performance significantly deteriorates without looping in a human in the decision-making process, as shown by Blaizot and colleagues' (2022) meta-analysis. One possible reason for this limitation is that they split text into fixed segments, which may hinder their ability to understand context, especially in longer texts (Guo et al., 2023). Newer large language models (LLMs) based on the transformer technology (Vaswani et al., 2017), such as Generative Pre-Trained Transformer (GPT) and Bidirectional Encoder Representations from Transformers (BERT), may overcome this problem since they capture more contextual information. This was demonstrated in a recent study, which

highlighted GPT's superior performance on systematic review tasks compared to older AI methods (Syriani et al., 2023).

Whether AI tools can match or surpass human performance in conducting systematic reviews is a question that carries profound implications for the future of scientific research. It holds the potential to radically transform knowledge synthesis, turning systematic reviews from static literature summaries into dynamic, continually updated resources – potentially altering the very way we approach science. Given these significant implications, it is imperative to acknowledge the current uncertain state of AI in this domain, especially regarding the most substantial LLM model, GPT-4. This model has been reported to significantly surpass all other LLMs, including previous versions of GPT, in various natural language processing tasks across both English and other languages (OpenAI, 2023). Yet, as of now, nothing is documented about GPT-4's performance in conducting systematic reviews.

Research on using other LLMs in systematic reviews (mostly earlier versions of GPT) is not as comprehensive or systematic as it could be, with much of the work containing contaminated datasets and inadequate metrics. No study, for instance, has tested grey literature and non-English literature, which can constitute a large proportion of the evidence base for some topics (Lawrence et al., 2014). Most studies focused on narrow aspects of systematic reviews, such as Boolean queries (S. Wang et al., 2023), or only evaluated performance for title and abstract screening (Alshami et al., 2023; Guo et al., 2023; Syriani et al., 2023). Some have methodological shortcomings, such as Mahuli et al. (2023), who did not provide an objective evaluation of GPT's performance, or Alshami et al. (2023), who did not test GPT's autonomous performance, instead relying on a 'human-in-the-loop' approach. A few may have included contaminated data, such as Guo et al. (2023) and Syriani et al. (2023), who used datasets for systematic reviews that

published their results before or in 2021, when GPT was trained, which may bias the results in favour of GPT. Other studies, such as Guo et al. (2023), did not consider imbalance nor incorporate the chance agreement in their interpretation of the results, potentially inflating GPT's accuracy. Syriani et al. (2023) addressed this issue but focused on investigating GPT's performance against other AI tools, not human reviewers. Our pre-registered study is the first to evaluate GPT-4's autonomous performance across several systematic review processes, including title/abstract screening, full-text screening, and data extraction. It is also the first to test an LLM model in reviewing grey literature and literature in other languages.

2. Methods

This study assessed GPT-4's performance in screening and extracting data from documents for an ongoing systematic review on parenting in protracted refugee situations. We used the ChatGPT interface to access the GPT-4 model between May and September 2023 for all screening stages and extraction. We manually copied and pasted the text of articles directly into the chat boxes within the constraints of a maximum character limit of 15,000 characters per interaction. We tested documents that were reviewed using four inclusion/exclusion criteria: containing empirical data, parenting behaviour, refugee status, and protracted refugee situation. The criteria were defined in detail in our GPT-4 prompts and outputs, which can be viewed together with the R code used for analysis on the Open Science Foundation (OSF) page (https://osf.io/bc3gj/?view_only=163ea151b0e24f748329444ae3b33752). We registered this study protocol on the OSF, while the details of the review can be found on both OSF (https://osf.io/tuv6q/?view_only=163ea151b0e24f748329444ae3b33752) and PROSPERO (CRD42022289404).

We screened 300 titles/abstracts and 150 full texts and extracted data from 30 full texts (see Figures 16 and 17). We tested for 13 non-English languages (see Figure 19 for breakdown). GPT-4 can comprehend non-English languages (OpenAI, 2023), so we did not perform extra training or translations for the non-English studies. Sample size was primarily based on studies reviewed by at least two humans for screening, which was the case for English language documents at all stages and those written in other languages in the title/abstract stage. However, due to time and resource constraints, we only used one human reviewer for non-English studies at the full-text level and data extraction from all studies. While we ensured a mix of decisions through a random selection of relevant and irrelevant studies written in English to gain deeper insight into inclusion performance, this was generally not possible for non-English studies due to the large number of excluded studies. Nevertheless, a random sample has been drawn from the excluded non-English studies.

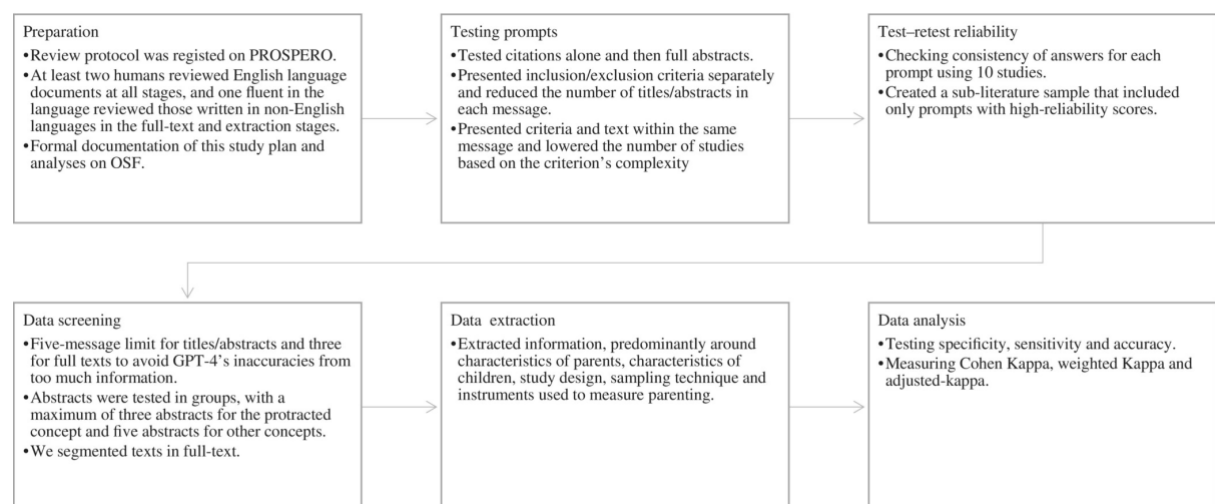


Figure 19. Flowchart of study steps.

Note. This flowchart shows the steps we took from preregistration to data analysis. OSF, Open Science Foundation.

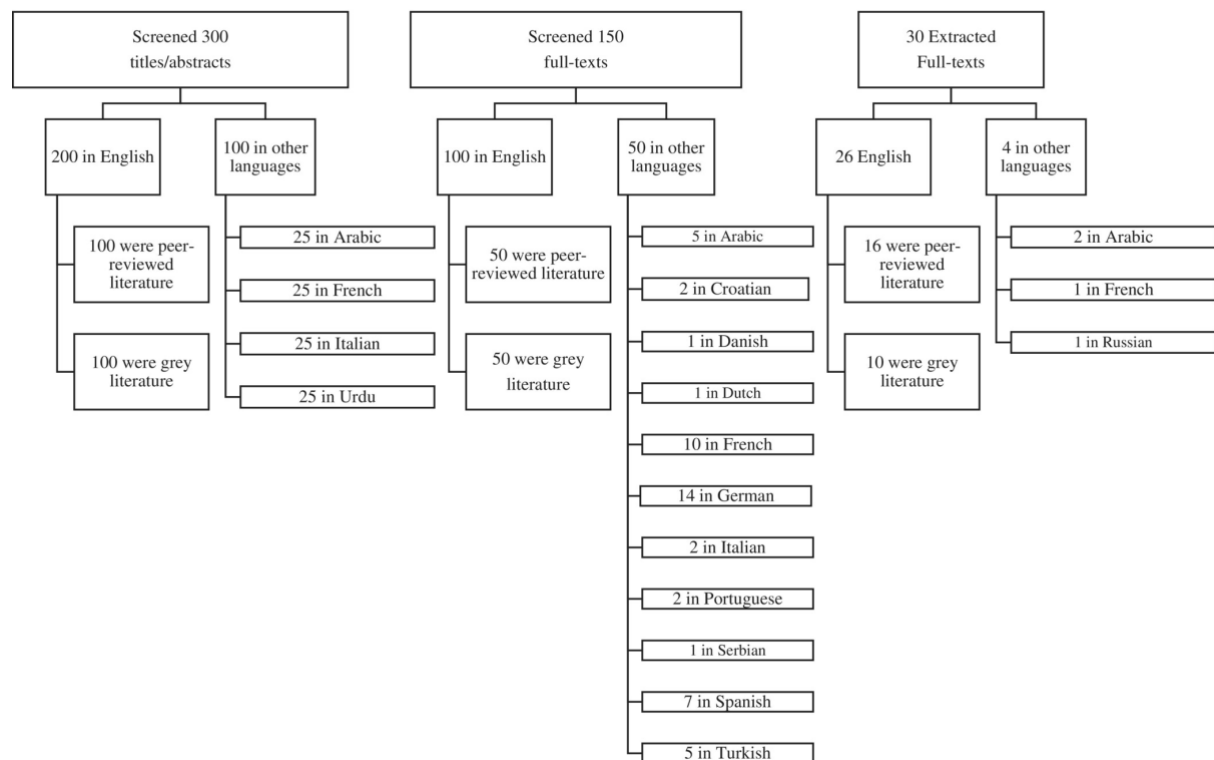


Figure 20. Distribution documents by language and type.

Note. This diagram shows the number and type of documents evaluated according to their language. The literature in other languages at the title and abstract screening stage was split into four groups, but this even split was not maintained at the full-text screening and data extraction stages because the available literature in other languages needed to match the same distribution. One Spanish study had its data extracted but was subsequently removed and randomly replaced with another study because it was excluded at the full-text stage.

2.1. Prompt engineering approach

Our prompt engineering method for GPT-4 involved an iterative approach at three stages: title/abstract screening, full-text screening, and data extraction. We refined our approach at each stage based on feedback and lessons learned from previous iterations. Each

iteration included three components: input modifications (e.g., prompt-citation separation, full abstract inclusion, simplified criteria), output restrictions (limiting GPT-4 to final decisions), and face validity monitoring (e.g., checking if the model excluded studies when keywords were mentioned). Within this process, we observed that the model sometimes generated ‘hallucinations:’ confident but inaccurate explanations (Beutel et al., 2023). Given this ambiguity surrounding the model’s internal reasoning, we focused solely on its final output, avoiding any subjective interpretations of its justification process. Basing our decision on GPT-4’s final outcome only, we assessed test-retest reliability using ten studies on each of our four criteria. Each criterion corresponds to a prompt. We tested the four prompts five times and generated five decisions per criterion per study. We then recorded each time GPT-4 output was inconsistent from the original answer. These scores were used to assess the impact of prompt reliability on accuracy scores. Note that some studies included in our prompt development approach were part of the included sample of studies in our main analysis.

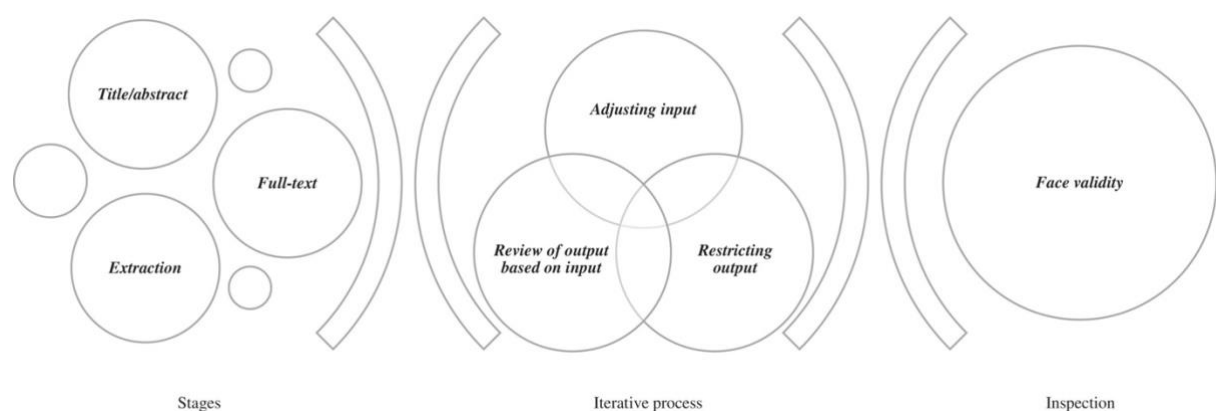


Figure 21. Optimising Generative Pre-Trained Transformer (GPT)-4 prompts for systematic review tasks for different stages using an iterative approach.

Note. This illustration conceptualises our method for identifying and addressing GPT-4 prompt limitations. We started with a first draft for a prompt for GPT-4 to perform a

specific task (e.g., screening titles/abstracts) before we analysed the outputs in relation to the input to identify the particular weaknesses of the initial prompt. To address the shortcomings, we removed, edited or added new instructions and repeated them until we achieved face validity between the input and output.

2.2. Screening and extraction

Each abstract batch (a maximum of three abstracts for the protracted concept – given that it contained a longer prompt – and five abstracts for other concepts) underwent four chat tests based on inclusion/exclusion criteria. We proceeded to the next criterion in another chat if GPT-4 responded ‘yes/maybe’. A ‘no’ meant exclusion from further review. If an abstract was excluded for not meeting a criterion (the exact reason did not matter at this stage), we removed it and introduced a new one for subsequent tests, meaning that the number of abstracts remained the same every time a criterion was tested. We set a five-message limit for titles/abstracts and three for full texts to avoid GPT-4’s inaccuracies arising from too much information. Lengthy abstracts in grey literature were divided for separate evaluations until GPT-4 finalised a decision. During grey literature screening, we modified the query to fit source formats, switching ‘titles/abstracts’ to ‘websites/reports’ and ‘texts.’ Of the 300 titles/abstracts tested for screening, a third were drawn from the peer-reviewed literature, a third were drawn from the grey literature, and a third were mixed literature in non-English.

For full texts, if GPT-4 said ‘Include’ for a segment, we viewed that criterion as met and began a new chat with the next criterion from the first segment. An ‘Exclude’ led us to present subsequent segments of that full text. If GPT-4 never indicated ‘Include’ by the last segment, the study was marked as excluded, and no further criteria were tested.

Where humans had excluded a study, we would start by testing that study using the

criterion the human raters used to exclude it. We did not indicate to GPT-4 that the humans had excluded the study or on what basis; we just started on the human exclusion criterion because we had to start on one of the four criteria. If GPT-4 shared the decision made by humans to exclude based on the first criterion, no further tests were needed. Otherwise, all remaining criteria were checked until we received an exclusion decision. As a result, 72% of the English peer-reviewed literature needed testing on the ‘parenting’ and ‘protracted refugee situations’ prompt, compared to 44% and 18% for grey and non-English literature, respectively. Peer-reviewed literature, grey literature, and non-English mixed sources made up the sample at this stage, each comprised one-third of the total 150 full-text documents tested.

Extracted information mainly revolved around characteristics of parents (e.g., number of parents, gender, and education distribution), characteristics of children (e.g., number of children, gender, and education distribution), study design (by duration, data collection method and group allocation), sampling technique, and instruments used to measure parenting (e.g., instrument type and name, target respondent and the main focus of the instrument). Only text from the Methods and Results sections of the papers were inputted into GPT-4. Similar to the above, we segmented the text into pieces. Only the first response to a prompt was deemed final, except if the response was incomplete; the subsequent responses were recorded as part of the answer. The extraction sample was mostly English studies, with only four non-English.

2.3. Analysis metrics

We applied four metrics: True Positives (*TP*), True Negatives (*TN*), False Positives (*FP*), and False Negatives (*FN*). *TP* and *TN* are the cases where GPT-4 agreed with human reviewers, meaning it made correct decisions. *FP* and *FN* are the cases where

GPT-4 disagreed with human reviewers, meaning it made wrong decisions. We calculated the imbalance of the dataset by dividing true cases by false cases.

$$Imbalance = \frac{TP + FP}{TN + FN}$$

For performance, we calculated sensitivity, specificity and accuracy, which are commonly used metrics (e.g., Frömke et al., 2022; Patel et al., 2021). Sensitivity (also known as recall) shows how well GPT-4 identified positive cases by taking (TP) and dividing them by the sum of TP and FN . Specificity indicates how well GPT-4 identified negative cases by taking TN and dividing them by the sum of TN and FP . Accuracy, which evaluates the overall correctness of GPT-4, was calculated by adding TP and TN and dividing by the total number of cases. There is no consensus on how to interpret these scores, as they largely depend on the context (e.g., Shreffler & Huecker, 2023). Previous studies have reported that human error rates in systematic review screening range from 5% to 20%, implying that a score of 100% is ‘superior’ to humans, while an accuracy score of 80% to 95% for GPT-4 could be regarded as ‘on par’ (Wang et al., 2020). In the worst possible documented prediction, human error rates reached up to 40% (Wang et al., 2020). This suggests that accuracy scores between 60% and 80% could be regarded as ‘near-par,’ and scores below 60% could be considered ‘subpar’.

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Cohen's kappa (κ) was calculated to compare the actual and expected agreement between GPT-4 and human reviewers (J. Cohen, 1960). This is important for systematic reviews, where inclusion is rare, and exclusion is common, which can make humans and GPT-4 seem more agreeable than they are. The actual agreement (P_o) is the proportion of cases humans and GPT-4 gave the same rating, positive or negative. The expected agreement (P_e) is the probability that humans and GPT-4 gave the same rating by chance, based on their rating frequencies. We subtracted the expected agreement from the actual agreement and divided it by the maximum possible agreement (1 minus the expected agreement). This gives a score between -1 and 1. We followed what McHugh (2012) suggested for the classification of agreement scores, which are more stringent on interpreting values than those suggested by Cohen (1960): values .0 – .20 as no agreement, .21 – .40 as minimal, .41 – .59 as weak, .60 – .79 as moderate, and 0.80 – .90 as almost perfect agreement.

Cohen's kappa has been shown to produce a false agreement rate in imbalanced datasets, so we used PABAK (prevalence-adjusted bias-adjusted kappa) to account kappa for the effects of prevalence and bias in the data set (Byrt et al., 1993). PABAK corrects this by accounting for the distribution of the categories in the denominator and adjusting the counts of agreements and disagreements in the formula. We also used weighted kappa to better capture the severity of false rejections by GPT-4, which are the most serious errors it can make, because this would exclude a study which meets the inclusion criteria. Weighted kappa ($\omega\kappa$) assigns higher weights to greater disagreements, with 0 for complete agreement. No standard test can combine weights with PABAK, so we could not account for the effects of data imbalance in the weighted Cohen's kappa scores. Sampson et al. (2011) found that the median search precision for systematic reviews is

around 3%; that is, there are about 30 times more excluded studies than included ones.

Therefore, we added a weight of 30 for false rejections in calculating weighted kappa.

$$\kappa = \frac{1 - Pe}{Po - Pe}$$

$$\omega\kappa = 1 - \frac{1 - Po\omega}{1 - Pe\omega}$$

$$PABAK = \frac{1 - 0.5}{2Po - 1}$$

3. Results

This study evaluated GPT-4's performance in screening and extracting peer-reviewed, grey (non-peer reviewed), and non-English literature. Across all stages and categories, the specificity was on par with human performance ($>.80$, except for English peer-reviewed full-text screening), indicating a robust ability of GPT-4 to identify irrelevant studies (Table 5) correctly. This was especially true in literature containing non-English studies ($>.90$). Sensitivity, indicating how effectively GPT-4 identified relevant studies, was highest in the extraction stage for both peer-reviewed (English: $.75$) and grey literature ($.65$), although note that for non-English studies the sensitivity was only $.36$ for data extraction. For non-English studies, perfect sensitivity was achieved during the full-text stage. Accuracy was somewhat higher in the data extraction stage than in the title/abstract screening phase, ranging between near-par and on par with human performance, except for the English peer-reviewed literature.

Table 5

Performance Evaluation of GPT-4 versus Human Reviewers in Screening and Extraction

	Balanc e	Sensitivit y	Specificit y	Accurac y	Cohen Kappa *	Weighte d Kappa	Adjuste d Kappa* *
Title and abstract screening							
English peer-reviewed	1	.42	.92	.67	.34	.23	.34
English grey	1	.48	.84	.66	.32	.24	.32
Other languages	.05	.50	.89	.88	.21	.40	.75
Full text screening							
English peer-reviewed	.92	.38	.69	.54	.07	.05	.08
English grey	.11	.60	.80	.78	.24	.44	.55
Other languages	.09	1	.95	.96	-.10	-.11	.64
High-reliability prompt group	.05	.36	.94	.85	.65	.97	.91
Data extraction							
English peer-reviewed	.03	.75	.84	.82	.54	.63	.63
English grey	.24	.65	.85	.81	.45	.53	.62
Other languages	.20	.36	.94	.85	.35	.29	.69

Performance Evaluation of GPT-4 versus Human Reviewers in Screening and Extraction

Note. We used PABAK (Prevalence-Adjusted Bias-Adjusted Kappa) to adjust Kappa for the effects of prevalence and bias in the data set (Byrt et al., 1993) and a weight of 30 for

false rejections in the calculation of weighted Kappa based on previous studies which have shown that the median search precision for systematic reviews is around 3% (Sampson et al., 2011).

* The inter-rater reliability between human reviewers for the full-text data was a Cohen Kappa coefficient of .77. However, the human review is not yet completed, so the final value may vary slightly. The inter-rater reliability between human reviewers was calculated using data from the peer-reviewed literature, but it also includes a small portion of grey literature and non-English studies.

** The human reviewers achieved an adjusted Cohen Kappa of .89 for the same literature sample Cohen Kappa was calculated for above.

To better understand the effect of prompts on accuracy, we analysed GPT-4's response consistency to the same prompt. This means that GPT-4 gave the same answer every time without any variation. For instance, if GPT-4 gave the same answer ten times in a row for the exact text and prompt, it scored 100%, but if it gave the same initial answer only seven times out of 10, it scored 70%. GPT-4 performed best when assessing empirical data and refugees (100% reliability) and struggled with the concepts of parenting behaviour (50% reliability) and protracted refugee situations (70% reliability). With these findings in mind, we created a sub-literature sample that included only prompts relating to refugee status and empirical data called the 'high-reliability prompt group,' given that these prompts had the highest reliability scores. Ultimately, this sub-sample included 23 studies: a third were English peer-reviewed studies, a third were English grey literature studies, and another third were non-English studies. Overall, highly reliable prompts achieved better accuracy (Table 1).

We subsequently looked at the balance of data, the ratio of relevant to irrelevant studies, for each literature type, language, and stage (for the extraction stage, this means the presence or absence of data). Peer-reviewed studies in English were fully balanced, as intended by our design, except in the extraction stage (.03; that is, one included for every 30 excluded). Unlike non-English studies (.05), the grey literature was fully balanced at the title/abstract stage, but then both grey literature and non-English studies were skewed towards irrelevance in the full-text screening (.11 and .09, respectively) and extraction stages (.24 and .20, respectively). It is important to note that while balance aids in understanding all aspects of performance, it does not reflect the inherent imbalances in real-world datasets of systematic review.

According to a meta-analysis by Sampson et al. (2011), the typical proportion of relevant studies in systematic reviews is around 3%, while proportions above 40% are outside of the observed range. Although this makes highly balanced studies atypical, it also makes them valuable for detecting any gaps in performance on the inclusion of relevant studies. Our balanced dataset of English peer-reviewed literature had lower accuracy (title/abstract: .67, full-text: .69) than the more unbalanced non-English literature dataset (title/abstract: .88, full-text: .96). In extraction – which was the only time the dataset of English peer-reviewed literature was unbalanced – accuracy was the highest (.84). The associated adjusted kappa scores were low, ranging from ‘none’ to ‘moderate’ as categorised by McHugh (2012). When aggregated, highly balanced samples had an average adjusted kappa score of 26%, while highly imbalanced samples scored an average of 76%. This led to an overall average adjusted kappa score of 55% across all datasets, as shown in Figure 4. When we exclusively examined responses from highly reliable prompts — which incidentally had a high imbalance — the adjusted kappa cores were ‘almost perfect’ agreement (.91). When we weighted kappa to emphasise false

rejections, in a way penalising GPT-4 for missing key studies, scores for the highly reliable prompt group improved (.97).

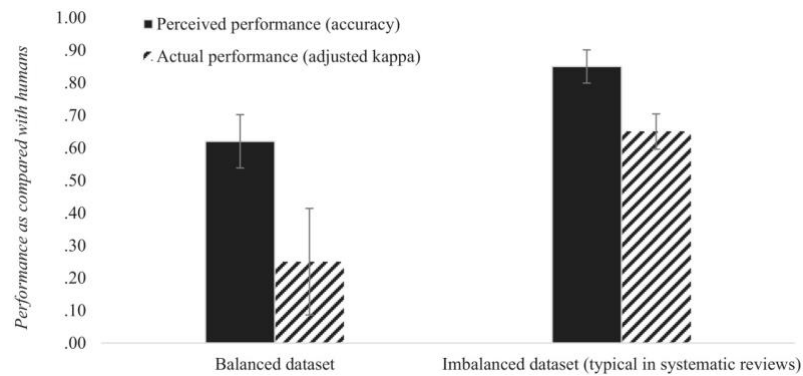


Figure 22. Generative Pre-Trained Transformer (GPT)-4 performance in systematic review screening with balanced and imbalanced datasets.

Note. This figure illustrates the aggregated performance of GPT-4 in screening and data extraction, measured by accuracy and adjusted kappa across two dataset types: balanced and imbalanced. Balanced datasets have a nearly equal distribution of included and excluded studies (approximately 1:1), which includes all datasets with a balance (see Table 1) of 0.90 and higher. In contrast, imbalanced datasets have a disproportionate ratio, with a significant majority of studies being excluded (approximately 1:3), which includes all datasets with a balance percentage of 0.40 or lower. To avoid study duplication, the high-reliability prompt group was not part of this analysis. The cut-off scores are based on Sampson et al.'s (2011) meta-analysis results of systematic review datasets, which demonstrate that nearly all systematic reviews are typically imbalanced (0.40). The confidence intervals display the 95% confidence level, calculated using a margin of error for the z -value.

4. Discussion

We found mixed results on the efficacy of GPT-4 compared to human reviewers across various systematic review tasks, languages, and literature types. GPT-4's accuracy was influenced by chance agreement and dataset imbalance, and when these factors were considered, GPT-4 tended to underperform humans substantially. Nevertheless, under specific conditions – namely, when given highly reliable prompts in full-text screening – GPT-4 demonstrated an ‘almost perfect’ performance, on par with humans. While our findings indicate the need for caution in assuming uniform proficiency across tasks, they also suggest that, under certain conditions, LLMs have the potential to revolutionise how we synthesise knowledge.

Our study was the first to report on the novel application of an LLM in conducting full-text screening and extraction. Automation techniques have been hailed as a way of increasing reproducibility (Ivimey-Cook et al., 2023). However, previous endeavours employing AI tools for these tasks pinpointed several challenges. These included the need for continuous human intervention in full-text screening (Beller et al., 2018), extensive pre-screening of training datasets (Halamoda-Kenzaoui et al., 2022), and the incapability to review full-texts (Clark et al., 2020; Nye et al., 2018) or to extract data deviating from a predetermined structure (Summerscales et al., 2011; Wallace et al., 2016). Our findings indicate that GPT-4 has limited potential in full-text screening and data extraction, with moderate performance in non-English and grey literature and a very poor ability with English peer-reviewed texts – which often comprise the bulk of included studies in reviews. These results contradicted our initial expectations shaped by preliminary experimentation with ChatGPT, highlighting the pitfalls of anecdotal testing noted by other researchers (Mahuli et al., 2023; Qureshi et al., 2023), an issue that likely arises

because GPT's human-like text generation is usually equated with human-level understanding (Jakesch et al., 2023).

Our results for sensitivity and specificity for title and abstract screening were consistent with previous studies. In line with Guo et al. (2023) and Alshami et al. (2023), we found that specificity was the most robust metric for title/abstract screening. Their specificity score was close to ours (90% and 93%, respectively, vs our 92%). Their sensitivity score was slightly higher than ours (76% and 84%, respectively, vs our 67%; although note that Alshami et al., 2023 used a human-in-the-loop method). Correspondingly, our accuracy score for the peer-reviewed literature (67%) was lower than in previous work (Alshami et al., 2023; Guo et al., 2023), possibly because our study was the only one that artificially balanced its dataset. As seen in Figure 4, GPT-4 significantly performed better in imbalanced datasets, where exclusion studies outnumbered inclusion studies. This observation is consistent with GPT-4's proclivity towards over-exclusion of studies, reported by us and others (Alshami et al., 2023; Guo et al., 2023). While this bias is essential to recognise, it is also unclear to what extent it changes performance in 'real life,' considering that we artificially balanced our dataset and most systematic literature review datasets have a low number of relevant studies (around 3% on average; Sampson et al., 2011). The adjusted kappa metric proves valuable here, providing a chance-corrected measure of GPT-4's performance that also accounts for the imbalance in the dataset. When employed, it significantly shifted chance-corrected performance scores from minimal to moderate, culminating in an average score of .65 (Figure 4). This performance was closer to human reviewers than initially indicated by unadjusted Cohen's kappa, as reported in a meta-analysis on kappa scores (.53; Sun, 2011).

Our sampling and analytical approach thus brings forth a nuanced picture of the ‘balance bias’ that appears to be a fundamental limitation of GPT-4 in the literature review process. It first indicates that GPT-4 may not be the best fit for systematic reviews, as it tends to over-exclude (e.g. false negatives). However, as the adjusted kappa result indicates, this characteristic does not entirely diminish its usefulness. GPT-4 may be used with caution as a ‘stringent’ secondary reviewer in the review process, whereby it takes the role of the more conservative reviewer – leaning towards rejection – complementing a primary human reviewer. Yet the use of GPT-4 as a secondary reviewer must also involve careful crafting of clear, specific, and concise prompts, given how our data suggested that prompt reliability can be affected by both word count and prompt complexity.

There are five main strengths and weaknesses to this study. First, we used a variety of metrics to measure GPT-4’s accuracy, consistency, and agreement with human reviewers. However, our relatively small sample size of evaluated papers might limit the generalisability of these findings for other systematic reviews. Second, though we registered and documented our research protocol and analysis plan to ensure its validity and replicability, it would also have been beneficial to create a detailed plan for prompt engineering in advance. This is particularly relevant considering the surge in research interest in prompt engineering since the commencement of our study. A Google Scholar search using the terms “prompt engineering” and “GPT” now yields over 6,550 results, more than half of which were published in the first month of 2024. Refer to Syriani (2023) for a systematic approach to prompt engineering in systematic reviews and Akinseloyin (2023) for using GPT in prompt generation for systematic reviews. Third, while focusing on GPT-4’s final decision allowed us to provide a precise measure of outcome without introducing subjective bias, we note that understanding the model’s reasoning process and reliability may be valuable for diagnosing errors, improving model

behaviour, and enhancing transparency. Fourth, we note that segmenting texts into chat boxes within a 15,000-character limit per interaction may have impacted results due to the potential loss of coherence in segmented texts. Lastly, while we used GPT-4's public interactive 'Chat' feature, considering its ease of use and popularity among researchers for screening, we recognise that the GPT Application Programming Interface (API) could have been a more streamlined option. Using APIs may have enabled us to circumvent some of the extra time it took to screen due to character limits imposed on our queries. Over time, as computing power improves and these limits loosen, this barrier may become less relevant.

5. Conclusion

Over a hundred years ago, audiences worldwide were captivated by a horse named Hans, who, with a confident tap of his hoof, appeared to solve mathematical problems. They believed Hans possessed an extraordinary cognitive ability, almost human-like in nature. Yet, beneath this illusion lay a simpler truth: Hans was reading his handler. He picked up on the faintest of cues – a twitch of a muscle, a barely noticeable nod, or even an unconscious sigh of anticipation. This phenomenon mirrors the workings of LLMs like GPT-4. While not deciphering human cues per se, they are heavily influenced by the input they receive, much like Hans needed clear guidance from his handler. Our findings underscore this: while GPT-4 has substantial limitations for many systematic review tasks when given a reliable prompt, GPT-4 can be effective at screening studies. In harnessing this potential, LLMs might pave the way for a transformative era in systematic reviews.

Chapter 6: Discussion

This thesis advances the study of ‘refugee parenting’ through empirical, conceptual, and methodological contributions. It does so through three distinct but related contributions: (i) redirecting scholarly attention to the understudied 78% of displaced people living in protracted refugee situations, whose long-term parenting experiences remain largely overlooked in current literature; (ii) expanding conceptual frameworks to incorporate ‘refugee coparenting’ dynamics that acknowledge the collaborative and relational nature of childrearing during displacement; and (iii) optimising methodological approaches to video data collection and literature synthesis through tools that enhance accessibility.

1. Linking temporality with relationality in ‘refugee parenting’

As argued in the Introduction, three distinct research domains—parenting studies, forced migration literature, and mental health research—have converged to shape research approaches to ‘refugee parenting.’ While this convergence provided valuable foundations, it has also imported some assumptions that, however valid in their origin context, may not appropriately capture the everyday reality for the 99% of refugees who never reach resettlement (UNHCR 2024; only 1% of refugees are ever resettled). Ample evidence supports this position: studies primarily conceptualising displacement trauma as a discrete event bounded in the past have demonstrated significant predictive power on parenting in contexts like Australia (Bryant et al., 2018), where refugee families access relatively stable support systems and institutional resources (which only 1% of refugees ever access; UNHCR, 2024). Yet these paradigms fall short in contexts such as Lebanon, Nepal, or Jordan, where daily survival challenges and immediate stressors emerge as

more powerful predictors of parenting behaviour (Jordans et al., 2012; A. Sim, Bowes, et al., 2018; A. Sim et al., 2019; A. Sim, Fazel, et al., 2018).

Encouragingly, the field has not been static in addressing these shortcomings, most notably with the widely applied ecological model that has accelerated progress toward more contextually sensitive understandings of refugees' behaviour and mental health, including parenting in displacement (K. E. Miller & Rasmussen, 2024). This thesis therefore joins this ongoing effort and adds to it in several ways, among which is highlighting two important inter-connected dimensions relating to parenting in displacement: the relational dimension, which highlights how the environmental demands that influence 'refugee parenting' are mediated through interpersonal networks, and the temporal dimension, which emphasises how long-term displacement creates unique challenges on refugee parents as they and their children seek improved livelihoods in liminal circumstances.

Across the systematic review and the qualitative study in this thesis, there was clear evidence that the temporality and relationality of "refugee parenting" are interconnected: protracted displacement creates continuous *universal/culture-specific* pathways of distress (temporal) that cascade through entire refugee family systems by straining parental roles that complicate the *negotiation* process between the refugee parents (relational) across a series of parental interactions that unfold cyclically (temporal), with these emotions periodically getting *mirrored* onto the child and the community (relational).

Although not explicitly labelled as such, the understanding of refugee behaviour as either temporally or relationally embedded has gained traction. Notable studies exploring temporality include those by Huntley et al., (2021) and Eltanamly et al. (2022), which demonstrate that different phases of refuge distinctly affect parenting practices.

Relational perspectives are also increasingly being integrated into studies spanning refugees in Lebanon (Jain et al., 2025) to refugees in Uganda (Loutet et al., 2024). While this growing body of research demonstrate how refugee parents naturally adapt their parenting behaviours across both dimensions, it still treats time and relationships as separate analytical categories, missing their dynamic interplay—how relationship networks evolve and take on different meanings across various stages of displacement, and how temporal challenges are collectively experienced and addressed through these evolving relationships. This fragmentation of temporal and relational analyses may help partly explain why some notable ‘refugee parenting’ interventions in humanitarian settings were not able to show the desired impact (e.g., Miller et al., 2022)—these interventions may be overlooking the temporal dimension of relationality; or, in other words, how parental dynamics gets altered as children develop and as displacement needs evolve or change with time or contextual needs, such as when families were presented with the COVID-19 pandemic.

2. Understanding inertia and friction in protracted refugee situations

Protracted refugee situations are created by having a large number of refugees (25,000+) from a particular country who have been refugees in that country for many years (5+) due to *both* the ongoing conflict in their country of origin *and* the state (in)action that have allowed for their refugee situations to persist (rather than end with resettlement; UNHCR, 2009). In the country of asylum, a legal ‘friction’ force stemming from established state policies that systematically restrict refugee movement and opportunities result in depression and anxiety that shape parenting behaviour (as argued in the systematic review). Meanwhile, a psychological ‘inertia’ force reflects accumulated community patterns of response to these constraints, as argued in the qualitative study under the

anchoring theme (in other studies, anchoring was similarly described as “Community talks,” Betancourt et al., 2015). These two forces resemble the physics principle of damped oscillation (Yang, 2023), where community-based opposition (analogous to inertia) and structural impediments (corresponding to friction) are two primary forms of resistance oscillation (change in behaviour).

In relation to parenting but also other forms of behaviours captured in a protracted refugee situation, these forces model could be particularly useful for explaining why it is difficult to make lasting changes in refugee behaviour with interventions in protracted refugee situations (Hadfield, Dajani, et al., 2024), including those targeting parenting practices (Bryant et al., 2022; Jordans et al., 2023). Different forms of ‘friction’ (e.g., bureaucratic barriers, social discrimination, resource limitations) and ‘inertia’ (e.g., community networks, cultural practices) may influence refugees’ responses to changing circumstances and create a ‘stopping power’ for sustained behaviour change. These forces can be thought of as an extension to the ecological model. Metaphorically, the refugee family can be thought of as a car, where ‘refugee parenting’ is the ‘engine’ and parental mental health is the ‘cooling system.’ What the ecological model by Miller & Rasmussen (2017) have suggested is that it is not only important to pay attention to previous ‘accidents’ the car might had (war trauma), but it is also important to look at the current ‘road’ the car is being driven on (ongoing stressors). What is being proposed here is that we must also pay attention to the ‘traffic’ (inertia force) together with the ‘traffic controller’ (friction force).

3. OBSESS and LLMs: Towards more inclusive methodological approaches

At the time their original paper was published, Miller and Rasmussen (2010) aimed to bridge the gap between the trauma-focused framework, which originated during war

periods, and the psychosocial approaches developed during the interwar era (Hayward, 2012). Yet behind this theoretical divide therein lies a broader democratic deficit in knowledge production, where Global North researchers and institutions have historically held disproportionate power in defining—and debating—how mental health and behaviour should be conceptualised and studied (Castro Torres & Alburez-Gutierrez, 2022; Henrich et al., 2010; X. Li et al., 2021; Osseiran & Nimer, 2024; Sear, 2016; Talebi, 2024). While knowledge production inequities are well known and discussed in many literatures, they are yet to be highlighted and addressed in the ‘refugee parenting’ literature. This thesis does so in two fundamental aspects of the research process: literature synthesis and data gathering.

When it comes to observational data, the thesis provides the first open-access coding system that can be used in displacement context. This methodological innovation represents an attempt to make future ‘refugee parenting’ research more inclusive and less prone to the aforementioned representational imbalances in knowledge production. Because OBSESS can be implemented in adverse environments without many cost barriers associated with traditional coding systems (Cañas et al., 2020), it may be better positioned for use in low- and middle-income countries where observational research tools are critically lacking (Bozicevic et al., 2024), yet where most of the world’s displaced population lives (UNHCR, 2024a). This research effort counts towards a more inclusive ‘refugee parenting’ science, by providing accessible tools for *future* video data gathering. As for the LLM study, it also contributes to increasing inclusivity in ‘refugee parenting’ research, but through finding a more cost-efficient way to synthesise *past* knowledge (e.g., high cost of hiring specialised research assistants).

Using LLMs to aid research synthesis efforts could result in a significant cost reduction: consider that manually screening 200 abstracts requires 3+ hours of human

labour costing \$30-40+ at minimum wage in countries like UK or US (Brysbaert, 2019 meta-analysis estimates people average reading 238 words per minute), while LLM-based screening costs approximately \$7.10 maximum and takes only 10-20 minutes (M. Li et al., 2024). Though cost differences are smaller in countries hosting large amount of refugees like Jordan (\$6) and Turkey (\$0.81/hour), these calculations exclude training costs and specialised labour which typically commands higher wages (cost of using LLMs are likely to keep getting cheaper, unlike salaries, which tend to raise). Despite the advantages, most LLMs (with notable exceptions like the Chinese LLM DeepSeek) are trained on data reflecting Western epistemologies and designed by companies that follow specific ideologies (Buyl et al., 2024). Models are therefore biased about who they represent as a ‘refugee’ or a ‘parent’ and what is ‘parenting’ or where displacement often happens—this should be taken in consideration when designing prompts to use such tools (Navigli et al., 2023).

4. Limitations

One overarching limitation of this thesis was the absence of a participatory approach throughout the research process. The validation of an LLM for use in systematic reviews did not incorporate input from Global South researchers. Additionally, OBSESS was designed without the involvement of refugee parents in the research design or data interpretation. Neither the systematic review nor the qualitative study include refugee parents as co-analysts. As a result, the findings presented in this thesis may inadvertently reflect Western-centric assumptions or overlook cultural nuances that a participatory approach might have revealed. Future research in this area would benefit from embedding participatory elements such as iterative feedback loops with refugee communities and collaboration with scholars from underrepresented settings.

The thesis could have also been strengthened by using a broader and more cohesive conceptual framework of ‘refugee parenting.’ Specifically, the systematic review was limited in scope by focusing on parenting and mental health, rather than also including parenting in relation to protection, food, education, and health (essential components of the Nurturing Care framework). In turn, the qualitative investigation of coparenting could have been expanded to include extended family members such as grandparents, older siblings, cousins, and uncles, as well as broader community practices of coparenting and alloparenting. This would provide a more holistic understanding of how caregiving responsibilities are distributed within non-nuclear refugee families’ formations—in relation to caregiver responsiveness and mental health, but also to the other components of Nurturing Care. It would be also valuable for OBSESS to expand beyond mother-child dyads to include multiple caregivers and triadic interactions, while examining ‘refugee parenting’ behaviours across various caregiving contexts or tasks (e.g., interactions relating to protection or feeding).

5. Conclusion

This thesis makes important contextual and conceptual contributions to the field by focusing on (co)parenting in protracted refugee situations, where 78% of refugees reside. Meanwhile, it advances the field methodologically by introducing OBSESS and LLMs as new approaches to enhance parenting data reliability, and research synthesis capabilities, for institutions and scholars with lower resources. While the research did not directly evaluate interventions, its findings carry significant implications for supporting refugee families. The results point toward the value of multi-level, context-sensitive support strategies. At the macro level, there is a critical need for ‘science diplomacy’ as an intervention to increase resettlement quotas and ultimately end protracted refugee

situations. At the program level, parenting interventions in protracted refugee situations should integrate both temporal and relational elements of 'refugee parenting.' This means placing greater emphasis on strengthening family systems, particularly coparenting dynamics, while remaining sensitive to the adaptive utility that certain seemingly 'negative' parenting behaviours may serve in protracted refugee situations. When testing intervention outcomes, researchers may need third-person data for documenting 'refugee parenting,' especially the positive aspects of mother-child relationships.

References

- 5th ed.; DSM-5; American Psychiatric Association. (2013). *American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition: DSM-5* (4th ed., Vol. 1). American Psychiatric Association.
<http://www.psychiatryonline.com/resourceTOC.aspx?resourceID=1>
- Abdullah, R. M., Ameen, S. Y., Ahmed, D. M., Kak, S. F., Yasin, H. M., Ibrahim, I. M., Ahmed, A. M., Rashid, Z. N., Omar, N., & Salih, A. A. (2021). Paralinguistic Speech Processing: An Overview. *Asian Journal of Research in Computer Science*, 34–46. <https://doi.org/10.9734/ajrcos/2021/v10i130233>
- Abu Kwaik, K., Saad, M., Chatzikyriakidis, S., & Dobnik, S. (2018). Shami: A Corpus of Levantine Arabic Dialects. In N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis, & T. Tokunaga (Eds.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. European Language Resources Association (ELRA). <https://aclanthology.org/L18-1576>
- Abu Salih, M., Abargil, M., Badarneh, S., Klein Selle, N., Irani, M., & Atzil, S. (2023). Evidence for cultural differences in affect during mother–infant interactions. *Scientific Reports*, 13, 4831. <https://doi.org/10.1038/s41598-023-31907-y>
- Abu-Errub, A., Odeh, A., Shambour, Q., & Al-Haj, H. (2014). Arabic roots extraction using morphological analysis. *IJCSI International Journal of Computer Science Issues*, 11, 128–134.
- Acolin, J., & Fishman, P. (2023). Beyond the biomedical, towards the agentic: A paradigm shift for population health science. *Social Science & Medicine* (1982), 326, 115950. <https://doi.org/10.1016/j.socscimed.2023.115950>

Adfam. (2009). *Becoming visible: The Somali community and substance use in London*.

https://adfam.org.uk/files/docs/becoming_visible.pdf

Adugna, F., Rudolf, M., & Getachew, M. (2022). A matter of time and contacts: Trans-local networks and long-term mobility of Eritrean refugees. *Journal of Ethnic and Migration Studies*, 0(0), 1–19. <https://doi.org/10.1080/1369183X.2022.2090155>

Afifi, T. D., Afifi, W. A., Merrill, A. F., & Nimah, N. (2016). ‘Fractured communities’: Uncertainty, stress, and (a lack of) communal coping in Palestinian refugee camps. *Journal of Applied Communication Research*, 44(4), 343–361.

<https://doi.org/10.1080/00909882.2016.1225166>

Ahmad, I., & Smetana, J. (2021). Palestinian Refugee Youth in Jordan: Parental Practices, Neighborhood Cohesion and Assistance, and Adolescent Wellbeing. *International Journal of Environmental Research and Public Health*, 18(7).

<https://doi.org/10.3390/ijerph18073649>

Ain, N. A. (2016). *A Sociolinguistic Study in am, Northern Jordan*.

<https://repository.essex.ac.uk/19387/1/POSTVIVA4.pdf>

Ainsworth, M. D. S., Bell, S. M., & Stayton, D. F. (1974). Infant-mother attachment and social development: Socialization as a product of reciprocal responsiveness to signals. In *The integration of a child into a social world* (pp. 99–135). Cambridge University Press.

Akesson, B., & Sousa, C. (2020). Parental Suffering and Resilience Among Recently Displaced Syrian Refugees in Lebanon. *Journal of Child and Family Studies*, 29(5), 1264–1273. <https://doi.org/10.1007/s10826-019-01664-6>

Akhter, S., & Kusakabe, K. (2014). Gender-based Violence among Documented Rohingya Refugees in Bangladesh. *INDIAN JOURNAL OF GENDER STUDIES*, 21(2), 225–246. <https://doi.org/10.1177/0971521514525088>

- Aldridge, V. K., Dovey, T. M., & Wade, A. (2017). Assessing Test-Retest Reliability of Psychological Measures: Persistent Methodological Problems. *European Psychologist*, 22(4), 207–218. <https://doi.org/10.1027/1016-9040/a000298>
- Aleinikoff, T., & Poellot, S. (2014). The Responsibility to Solve: The International Community and Protracted Refugee Situations. *Georgetown Law Faculty Publications and Other Works*.
<https://scholarship.law.georgetown.edu/facpub/2090>
- Alisic, E., Zalta, A. K., Van Wesel, F., Larsen, S. E., Hafstad, G. S., Hassanpour, K., & Smid, G. E. (2014). Rates of post-traumatic stress disorder in trauma-exposed children and adolescents: Meta-analysis. *British Journal of Psychiatry*, 204(5), 335–340. <https://doi.org/10.1192/bjp.bp.113.131227>
- Alleyne-Green, B., Espinoza-Kulick, A., Grocher, K., McCutcheon, K., & Betancourt, T. (2019). The Impact of War Violence Exposure and Psychological Distress on Parenting Practices Among a Sample of Young Adults Affected by War Postconflict Sierra Leone. *Peace and Conflict: Journal of Peace Psychology*, 25, 325–334. <https://doi.org/10.1037/pac0000421>
- Al-Natour, A., Morris, E. J., & Mohammad Al-Ostaz, S. (2022). Through Her Eyes: The Impact of War on Syrian Refugee Families. *Journal of Transcultural Nursing : Official Journal of the Transcultural Nursing Society*, 33(1), 26–32.
<https://doi.org/10.1177/10436596211026367>
- Alnuaimi, D. (2018). *Healthcare Access for Syrian Refugees and the Jordanian Host Community in Ramtha and the Houran Plain, Jordan: A Qualitative Study*.
- Alshami, A., Elsayed, M., Ali, E., Eltoukhy, A. E. E., & Zayed, T. (2023). Harnessing the Power of ChatGPT for Automating Systematic Review Process: Methodology,

- Case Study, Limitations, and Future Directions. *Systems*, 11(7), 351.
<https://doi.org/10.3390/systems11070351>
- An, W., Beauville, R., & Rosche, B. (2022). Causal Network Analysis. *Annual Review of Sociology*, 48(Volume 48, 2022), 23–41. <https://doi.org/10.1146/annurev-soc-030320-102100>
- Arakelyan, S., & Ager, A. (2021). Annual Research Review: A multilevel bioecological analysis of factors influencing the mental health and psychosocial well-being of refugee children. *Journal of Child Psychology and Psychiatry*, 62(5), 484–509.
<https://doi.org/10.1111/jcpp.13355>
- Arnetz, B. B., Broadbridge, C. L., Jamil, H., Lumley, M. A., Pole, N., Barkho, E., Fakhouri, M., Talia, Y. R., & Arnetz, J. E. (2014). Specific Trauma Subtypes Improve the Predictive Validity of the Harvard Trauma Questionnaire in Iraqi Refugees. *Journal of Immigrant and Minority Health*, 16(6), 1055–1061.
<https://doi.org/10.1007/s10903-014-9995-9>
- Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. *JBIM Evidence Implementation*, 13(3), 132. <https://doi.org/10.1097/XEB.0000000000000055>
- Asaf, Y. (2017). Syrian Women and the Refugee Crisis: Surviving the Conflict, Building Peace, and Taking New Gender Roles. *Social Sciences*, 6(3), 110.
<https://doi.org/10.3390/socsci6030110>
- Aspland, H., & Gardner, F. (2003). Observational Measures of Parent-Child Interaction: An Introductory Review. *Child and Adolescent Mental Health*, 8(3), 136–143.
<https://doi.org/10.1111/1475-3588.00061>

- Assali, A., Younis, M., Sager, N., Dakis, M., & Young, D. (2022). Stories of Struggle and Resilience: Parenting in Three Refugee Contexts. In S. Ali Samadi (Ed.), *Parenting—Challenges of Child Rearing in a Changing Society*. IntechOpen. <https://doi.org/10.5772/intechopen.103827>
- Atallah, D. G. (2017). A community-based qualitative study of intergenerational resilience with Palestinian refugee families facing structural violence and historical trauma. *Transcultural Psychiatry*, 54(3), 357–383. <https://doi.org/10.1177/1363461517706287>
- Azriel, O., & Bar-Haim, Y. (2020). Attention bias. In *Clinical handbook of fear and anxiety: Maintenance processes and treatment mechanisms* (pp. 203–218). American Psychological Association. <https://doi.org/10.1037/0000150-012>
- Babcock Fenerci, R. L., Chu, A. T., & DePrince, A. P. (2016). Intergenerational Transmission of Trauma-Related Distress: Maternal Betrayal Trauma, Parenting Attitudes, and Behaviors. *Journal of Aggression, Maltreatment & Trauma*, 25(4), 382–399. <https://doi.org/10.1080/10926771.2015.1129655>
- Babuç, Z. T. (2021). A Relational Sociological Analysis on the Impact of COVID-19 Pandemic Lockdown on Syrian Migrants’ Lives in Turkey: The Case of Mersin Province. *Journal of International Migration and Integration*, 1–22. <https://doi.org/10.1007/s12134-021-00907-0>
- Babuç, Z. T. (2022). A Relational Sociological Analysis on the Impact of COVID-19 Pandemic Lockdown on Syrian Migrants’ Lives in Turkey: The Case of Mersin Province. *Journal of International Migration and Integration*, 23(4), 1645–1666. <https://doi.org/10.1007/s12134-021-00907-0>
- Baranov, V., Frost, A., Hagaman, A., Simmons, J. G., Manzoor, M. S., Biroli, P., Bhalotra, S., Rahman, A., Sikander, S., & Maselko, J. (2022). Effects of a maternal

- psychosocial intervention on hair derived biomarkers of HPA axis function in mothers and children in rural Pakistan. *SSM - Mental Health*, 2, 100082.
<https://doi.org/10.1016/j.ssmmh.2022.100082>
- Barber, B. K. (2001). Political violence, social integration, and youth functioning: Palestinian youth from the Intifada. *Journal of Community Psychology*, 29(3), 259–280. <https://doi.org/10.1002/jcop.1017>
- Barber, B. K., Spellings, C., McNeely, C., Page, P. D., Giacaman, R., Arafat, C., Daher, M., El Sarraj, E., & Mallouh, M. A. (2014). Politics drives human functioning, dignity, and quality of life. *Social Science & Medicine*, 122, 90–102.
<https://doi.org/10.1016/j.socscimed.2014.09.055>
- Barel, E., Van IJzendoorn, M. H., Sagi-Schwartz, A., & Bakermans-Kranenburg, M. J. (2010). Surviving the Holocaust: A meta-analysis of the long-term sequelae of a genocide. *Psychological Bulletin*, 136(5), 677–698.
<https://doi.org/10.1037/a0020339>
- Baumrind, D. (1966). Effects of Authoritative Parental Control on Child Behavior. *Child Development*, 37(4), 887. <https://doi.org/10.2307/1126611>
- Baumrind, D. (1971). Current patterns of parental authority. *Developmental Psychology*, 4(1, Pt.2), 1–103. <https://doi.org/10.1037/h0030372>
- Baur, T., Damian, I., Lingenfelser, F., Wagner, J., & André, E. (2013). NovA: Automated Analysis of Nonverbal Signals in Social Interactions. In A. A. Salah, H. Hung, O. Aran, & H. Gunes (Eds.), *Human Behavior Understanding* (pp. 160–171). Springer International Publishing. https://doi.org/10.1007/978-3-319-02714-2_14
- Bedaso, A., & Duko, B. (2022). Epidemiology of depression among displaced people: A systematic review and meta-analysis. *Psychiatry Research*, 311, 114493.
<https://doi.org/10.1016/j.psychres.2022.114493>

- Beebe, B., Jaffe, J., Markese, S., Buck, K., Chen, H., Cohen, P., Bahrick, L., Andrews, H., & Feldstein, S. (2010). The origins of 12-month attachment: A microanalysis of 4-month mother-infant interaction. *Attachment & Human Development, 12*(1–2), 3–141. <https://doi.org/10.1080/14616730903338985>
- Bellamy, Haysom, Wake, & Barbelet. (2017). *The lives and livelihoods of Syrian refugees*.
- Beller, E., Clark, J., Tsafnat, G., Adams, C., Diehl, H., Lund, H., Ouzzani, M., Thayer, K., Thomas, J., Turner, T., Xia, J., Robinson, K., Glasziou, P., Adams, C., Ahtirski, O., Beller, E., Clark, J., Christensen, R., Diehl, H., ... On behalf of the founding members of the ICASR group. (2018). Making progress with the automation of systematic reviews: Principles of the International Collaboration for the Automation of Systematic Reviews (ICASR). *Systematic Reviews, 7*(1), 77. <https://doi.org/10.1186/s13643-018-0740-7>
- Bennetts, S. K., Love, J., Westrupp, E. M., Hackworth, N. J., Mensah, F. K., Nicholson, J. M., & Levickis, P. (2021). Protocol for the Adaptation of a Direct Observational Measure of Parent-Child Interaction for Use With 7–8-Year-Old Children. *Frontiers in Psychology, 11*. <https://www.frontiersin.org/articles/10.3389/fpsyg.2020.619336>
- Bennetts, S. K., Mensah, F. K., Westrupp, E. M., Hackworth, N. J., & Reilly, S. (2016). The Agreement between Parent-Reported and Directly Measured Child Language and Parenting Behaviors. *Frontiers in Psychology, 7*. <https://doi.org/10.3389/fpsyg.2016.01710>
- Bentley, G. E., Thomas-Jackson, S. C., Goff, B. S. N., & Piland, N. (2022). Ambiguous Loss and Coping Strategies in Couples Raising a Child With Down Syndrome: A

- Qualitative Directed Content Analysis Study. *Medical Research Archives*, 10(11).
<https://doi.org/10.18103/mra.v10i11.3267>
- Bergmeier, H., Skouteris, H., & Hetherington, M. (2015). Systematic research review of observational approaches used to evaluate mother-child mealtime interactions during preschool years²³. *The American Journal of Clinical Nutrition*, 101(1), 7–15. <https://doi.org/10.3945/ajcn.114.092114>
- Bergnehr, D. (2022). Adapted fathering for new times: Refugee men's narratives on caring for home and children. *Journal of Family Studies*, 28(3), 934–949.
<https://doi.org/10.1080/13229400.2020.1769708>
- Bergset, K., & Ulvik, O. S. (2021). Parenting in exile: Refugee parents' multivoiced narratives. *International Social Work*, 64(3), 412–424.
<https://doi.org/10.1177/0020872819825778>
- Bernard, K., Meade, E. B., & Dozier, M. (2013). Parental synchrony and nurturance as targets in an attachment based intervention: Building upon Mary Ainsworth's insights about mother-infant interaction. *Attachment & Human Development*, 15(5–6), 507–523. <https://doi.org/10.1080/14616734.2013.820920>
- Betancourt, T. S., Abdi, S., Ito, B. S., Lilienthal, G. M., Agalab, N., & Ellis, H. (2015). We left one war and came to another: Resource loss, acculturative stress, and caregiver-child relationships in somali refugee families. *Cultural Diversity and Ethnic Minority Psychology*, 21(1), 114–125. <https://doi.org/10.1037/a0037538>
- Betancourt, T. S., & Khan, K. T. (2008). The mental health of children affected by armed conflict: Protective processes and pathways to resilience. *International Review of Psychiatry*, 20(3), 317–328. <https://doi.org/10.1080/09540260802090363>
- Betts, Chaara, Omata, & Sterck. (2019). *Uganda's Self-Reliance Model: Does it Work?*

- Beutel, G., Geerits, E., & Kielstein, J. T. (2023). Artificial hallucination: GPT on LSD? *Critical Care*, 27(1), 148. <https://doi.org/10.1186/s13054-023-04425-6>
- Biringen, Z., Derscheid, D., Vliegen, N., Closson, L., & Easterbrooks, M. A. (2014). Emotional availability (EA): Theoretical background, empirical research using the EA Scales, and clinical applications. *Developmental Review*, 34(2), 114–167. <https://doi.org/10.1016/j.dr.2014.01.002>
- Bjørknes, R., Kjøbli, J., Manger, T., & Jakobsen, R. (2012). Parent training among ethnic minorities: Parenting practices as mediators of change in child conduct problems. *Family Relations*, 61(1), 101–114. <https://doi.org/10.1111/j.1741-3729.2011.00683.x>
- Black, M. M., Aboud, F., Billah, S. M., Cerna-Turoff, I., Dandona, R., Horiuchi, S., Lazzerini, M., Muñiz, M., & Strong, K. (2024). Responsive caregiving: Conceptual clarity and the need for indicators. *The Lancet Child & Adolescent Health*, 8(10), 713–715. [https://doi.org/10.1016/S2352-4642\(24\)00194-9](https://doi.org/10.1016/S2352-4642(24)00194-9)
- Blackmore, R., Boyle, J. A., Fazel, M., Ranasinha, S., Gray, K. M., Fitzgerald, G., Misso, M., & Gibson-Helm, M. (2020). The prevalence of mental illness in refugees and asylum seekers: A systematic review and meta-analysis. *PLOS Medicine*, 17(9), e1003337. <https://doi.org/10.1371/journal.pmed.1003337>
- Blackmore, R., Gray, K. M., Boyle, J. A., Fazel, M., Ranasinha, S., Fitzgerald, G., Misso, M., & Gibson-Helm, M. (2020). Systematic Review and Meta-analysis: The Prevalence of Mental Illness in Child and Adolescent Refugees and Asylum Seekers. *Journal of the American Academy of Child and Adolescent Psychiatry*, 59(6), 705–714. <https://doi.org/10.1016/j.jaac.2019.11.011>
- Blaizot, A., Veettil, S. K., Saidoung, P., Moreno-Garcia, C. F., Wiratunga, N., Aceves-Martins, M., Lai, N. M., & Chaiyakunapruk, N. (2022). Using artificial

- intelligence methods for systematic review in health sciences: A systematic review. *Research Synthesis Methods*, 13(3), 353–362.
<https://doi.org/10.1002/jrsm.1553>
- Bogic, M., Njoku, A., & Priebe, S. (2015). Long-term mental health of war-refugees: A systematic literature review. *BMC International Health and Human Rights*, 15(1).
<https://doi.org/10.1186/s12914-015-0064-9>
- Borah, R., Brown, A. W., Capers, P. L., & Kaiser, K. A. (2017). Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open*, 7(2), e012545.
<https://doi.org/10.1136/bmjopen-2016-012545>
- Borairi, S., Deneault, A.-A., Madigan, S., Fearon, P., Devereux, C., Geer, M., Jeyanayagam, B., Martini, J., & Jenkins, J. (2024). A meta-analytic examination of sensitive responsiveness as a mediator between depression in mothers and psychopathology in children. *Attachment & Human Development*, 26(4), 273–300.
<https://doi.org/10.1080/14616734.2024.2359689>
- Boruszak-Kiziukiewicz, J., & Kmita, G. (2020). Parenting Self-Efficacy in Immigrant Families—A Systematic Review. *Frontiers in Psychology*, 11, 985.
<https://doi.org/10.3389/fpsyg.2020.00985>
- Boucher, J., & Osgood, C. E. (1969). The pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior*, 8(1), 1–8. [https://doi.org/10.1016/S0022-5371\(69\)80002-2](https://doi.org/10.1016/S0022-5371(69)80002-2)
- Bowen, M. (1966). The use of family theory in clinical practice. *Comprehensive Psychiatry*, 7(5), 345–374. [https://doi.org/10.1016/S0010-440X\(66\)80065-2](https://doi.org/10.1016/S0010-440X(66)80065-2)
- Bozicevic, L., Lucas, C., Magai, D. N., Ooi, Y., Maliwichi, L., Sharp, H., & Gladstone, M. (2024). Evaluating caregiver-child interactions in low- and middle-income

- countries: A systematic review of tools and methods. *Journal of Reproductive and Infant Psychology*, 0(0), 1–36. <https://doi.org/10.1080/02646838.2024.2321615>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352. <https://doi.org/10.1080/14780887.2020.1769238>
- Brenner, P. S., & DeLamater, J. (2016). Lies, Damned Lies, and Survey Self-Reports? Identity as a Cause of Measurement Bias. *Social Psychology Quarterly*, 79(4), 333–354. <https://doi.org/10.1177/0190272516628298>
- Bronfenbrenner, U. (1974). Developmental Research, Public Policy, and the Ecology of Childhood. *Child Development*, 45(1), 1. <https://doi.org/10.2307/1127743>
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, 32, 513–531.
- Bronfenbrenner, U. (1979). Contexts of child rearing: Problems and prospects. *American Psychologist*, 34(10), 844–850. <https://doi.org/10.1037/0003-066X.34.10.844>
- Bronfenbrenner, U. (1996). *The ecology of human development: Experiments by nature and design*. Harvard University Press.
- Brown, F. L., Bosqui, T., Elias, J., Farah, S., Mayya, A., Abo Nakkoul, D., Walsh, B., Chreif, S., Einein, A., Meksassi, B., Abi Saad, R., Naal, H., Ghossainy, M. E., Donnelly, M., Betancourt, T. S., Carr, A., Puffer, E., El Chammay, R., & Jordans, M. J. D. (2022). Family systemic psychosocial support for at-risk adolescents in Lebanon: Study protocol for a multi-site randomised controlled trial. *Trials*, 23(1), 327. <https://doi.org/10.1186/s13063-022-06284-y>

- Bryant, R. A., Bawaneh, A., Awwad, M., Al-Hayek, H., Giardinelli, L., Whitney, C., Jordans, M. J. D., Cuijpers, P., Sijbrandij, M., Ventevogel, P., Dawson, K., Akhtar, A., & STRENGTHS Consortium. (2022). Twelve-month follow-up of a randomised clinical trial of a brief group psychological intervention for common mental disorders in Syrian refugees in Jordan. *Epidemiology and Psychiatric Sciences*, 31, e81. <https://doi.org/10.1017/S2045796022000658>
- Bryant, R. A., Edwards, B., Creamer, M., O'Donnell, M., Forbes, D., Felmingham, K. L., Silove, D., Steel, Z., McFarlane, A. C., Van Hooff, M., Nickerson, A., & Hadzi-Pavlovic, D. (2021). Prolonged grief in refugees, parenting behaviour and children's mental health. *Australian & New Zealand Journal of Psychiatry*, 55(9), 863–873. <https://doi.org/10.1177/0004867420967420>
- Bryant, R. A., Edwards, B., Creamer, M., O'Donnell, M., Forbes, D., Felmingham, K. L., Silove, D., Steel, Z., Nickerson, A., McFarlane, A. C., Van Hooff, M., & Hadzi-Pavlovic, D. (2018a). The effect of post-traumatic stress disorder on refugees' parenting and their children's mental health: A cohort study. *The Lancet Public Health*, 3(5), e249–e258. [https://doi.org/10.1016/S2468-2667\(18\)30051-3](https://doi.org/10.1016/S2468-2667(18)30051-3)
- Bryant, R. A., Edwards, B., Creamer, M., O'Donnell, M., Forbes, D., Felmingham, K. L., Silove, D., Steel, Z., Nickerson, A., McFarlane, A. C., Van Hooff, M., & Hadzi-Pavlovic, D. (2018b). The effect of post-traumatic stress disorder on refugees' parenting and their children's mental health: A cohort study. *The Lancet Public Health*, 3(5), e249–e258. [https://doi.org/10.1016/S2468-2667\(18\)30051-3](https://doi.org/10.1016/S2468-2667(18)30051-3)
- Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language*, 109, 104047. <https://doi.org/10.1016/j.jml.2019.104047>

- Bulford, E., Fogarty, A., Giallo, R., Brown, S., Szwarc, J., & Riggs, E. (2022). Risk and Protective Factors Experienced by Fathers of Refugee Background during the Early Years of Parenting: A Qualitative Study. *International Journal of Environmental Research and Public Health*, 19(11), 6940. <https://doi.org/10.3390/ijerph19116940>
- Bunn, M. (2019). *Restoring Social Bonds: Group-Based Treatment and the Social Resources of Syrian Refugees*. University of Chicago.
- Bunn, M., Zolman, N., Smith, C. P., Khanna, D., Hanneke, R., Betancourt, T. S., & Weine, S. (2022). Family-based mental health interventions for refugees across the migration continuum: A systematic review. *SSM - Mental Health*, 2, 100153. <https://doi.org/10.1016/j.ssmmh.2022.100153>
- Burkitt, I. (2016). Relational agency: Relational sociology, agency and interaction. *European Journal of Social Theory*, 19(3), 322–339. <https://doi.org/10.1177/1368431015591426>
- Buyl, M., Rogiers, A., Noels, S., Dominguez-Catena, I., Heiter, E., Romero, R., Johary, I., Mara, A., Lijffijt, J., & Bie, T. (2024). *Large Language Models Reflect the Ideology of their Creators*. <https://doi.org/10.48550/arXiv.2410.18417>
- Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology*, 46(5), 423–429. [https://doi.org/10.1016/0895-4356\(93\)90018-v](https://doi.org/10.1016/0895-4356(93)90018-v)
- Cabrera, N. J., Tamis-LeMonda, C. S., Bradley, R. H., Hofferth, S., & Lamb, M. E. (2000). Fatherhood in the twenty-first century. *Child Development*, 71(1), 127–136. <https://doi.org/10.1111/1467-8624.00126>

- Cabrera, N. J., Volling, B. L., & Barr, R. (2018). Fathers Are Parents, Too! Widening the Lens on Parenting for Children's Development. *Child Development Perspectives*, 12(3), 152–157. <https://doi.org/10.1111/cdep.12275>
- Campbell, C. G. (2022). Two Decades of Coparenting Research: A Scoping Review. *Marriage & Family Review*, 1–33. <https://doi.org/10.1080/01494929.2022.2152520>
- Campos, B., Ullman, J. B., Aguilera, A., & Dunkel Schetter, C. (2014). Familism and psychological health: The intervening role of closeness and social support. *Cultural Diversity and Ethnic Minority Psychology*, 20(2), 191–201. <https://doi.org/10.1037/a0034094>
- Cañas, M., Ibabe, I., & De Paúl, J. (2020). Promising observational instruments of parent-child (0–12 years) interaction within the child protection system: A systematic review. *Child Abuse & Neglect*, 109, 104713. <https://doi.org/10.1016/j.chiabu.2020.104713>
- Capps, L. L. (1999). Fright illness in Hmong children. *Pediatric Nursing*, 25(4), 378–394.
- Carey, W. (1982). Validity of Parental Assessments of Development and Behavior. *American Journal of Diseases of Children*, 136(2), 97–99. <https://doi.org/10.1001/archpedi.1982.03970380009001>
- Carpiniello, B. (2023). The Mental Health Costs of Armed Conflicts—A Review of Systematic Reviews Conducted on Refugees, Asylum-Seekers and People Living in War Zones. *International Journal of Environmental Research and Public Health*, 20(4), Article 4. <https://doi.org/10.3390/ijerph20042840>
- Carrilio, T. E., & Walter, C. A. (1984). Mirroring and autonomy: The dual tasks of mothers. *Child and Adolescent Social Work Journal*, 1(3), 143–152. <https://doi.org/10.1007/BF00798430>

Castro Torres, A. F., & Alburez-Gutierrez, D. (2022). North and South: Naming practices and the hidden dimension of global disparities in knowledge production.

Proceedings of the National Academy of Sciences, 119(10), e2119373119.

<https://doi.org/10.1073/pnas.2119373119>

CDC. (2023, April 28). *About Mental Health*.

<https://www.cdc.gov/mentalhealth/learn/index.htm>

Çelik, Ç., & İçduygu, A. (2019). Schools and Refugee Children: The Case of Syrians in Turkey. *International Migration*, 57(2), 253–267.

<https://doi.org/10.1111/imig.12488>

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, 76(6), 893–910. <https://doi.org/10.1037/0022-3514.76.6.893>

Chaudhuri, N. (2004). Interventions to Improve Children’s Health by Improving the Housing Environment. *Reviews on Environmental Health*, 19(3–4), 197–222.

<https://doi.org/10.1515/reveh-2004-19-3-404>

Christie, H., Hamilton-Giachritsis, C., Alves-Costa, F., Tomlinson, M., & Halligan, S. L. (2019). The impact of parental posttraumatic stress disorder on parenting: A systematic review. *European Journal of Psychotraumatology*, 10(1), 1550345.

<https://doi.org/10.1080/20008198.2018.1550345>

Christie, H., Hamilton-Giachritsis, C., Alves-Costa, F., Tomlinson, M., Stewart, J., Skeen, S., Notholi, V., Gqwaka, P., Sambudla, A., & Halligan, S. (2020). Associations between parental trauma, mental health, and parenting: A qualitative study in a high-adversity South African community. *Social Science & Medicine*, 265, 113474. <https://doi.org/10.1016/j.socscimed.2020.113474>

- Clark, H., Coll-Seck, A. M., Banerjee, A., Peterson, S., Dalglish, S. L., Ameratunga, S., Balabanova, D., Bhan, M. K., Bhutta, Z. A., Borrazzo, J., Claeson, M., Doherty, T., El-Jardali, F., George, A. S., Gichaga, A., Gram, L., Hipgrave, D. B., Kwamie, A., Meng, Q., ... Costello, A. (2020). A future for the world's children? A WHO–UNICEF–Lancet Commission. *The Lancet*, 395(10224), 605–658.
[https://doi.org/10.1016/S0140-6736\(19\)32540-1](https://doi.org/10.1016/S0140-6736(19)32540-1)
- Clark, J., Glasziou, P., Del Mar, C., Bannach-Brown, A., Stehlik, P., & Scott, A. M. (2020). A full systematic review was completed in 2 weeks using automation tools: A case study. *Journal of Clinical Epidemiology*, 121, 81–90.
<https://doi.org/10.1016/j.jclinepi.2020.01.008>
- Cluver, L., Shenderovich, Y., Meinck, F., Berezin, M. N., Doubt, J., Ward, C. L., Parra-Cardona, J., Lombard, C., Lachman, J. M., Wittesaele, C., Wessels, I., Gardner, F., & Steinert, J. I. (2020). Parenting, mental health and economic pathways to prevention of violence against children in South Africa. *Social Science & Medicine*, 262, 113194. <https://doi.org/10.1016/j.socscimed.2020.113194>
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), 37–46.
<https://doi.org/10.1177/001316446002000104>
- Cohen, L. R., Hien, D. A., & Batchelder, S. (2008). The Impact of Cumulative Maternal Trauma and Diagnosis on Parenting Behavior. *Child Maltreatment*, 13(1), 27–38.
<https://doi.org/10.1177/1077559507310045>
- Colegrove, V. M., & Havighurst, S. S. (2017). Review of Nonverbal Communication in Parent–Child Relationships: Assessment and Intervention. *Journal of Child and Family Studies*, 26(2), 574–590. <https://doi.org/10.1007/s10826-016-0563-x>

- Conger, R. D., & Donnellan, M. B. (2007). *An interactionist perspective on the socioeconomic context of human development* (Vol. 58, p. 199).
<https://doi.org/10.1146/annurev.psych.58.110405.085551>
- Conger, R. D., Ge, X., Elder, G. H., Jr., Lorenz, F. O., & Simons, R. L. (1994). Economic Stress, Coercive Family Process, and Developmental Problems of Adolescents. *Child Development*, 65(2), 541–561. <https://doi.org/10.1111/j.1467-8624.1994.tb00768.x>
- Convention on the Rights of the Child (1989).
<https://www.ohchr.org/en/professionalinterest/pages/crc.aspx>
- Côté-Olijnyk, M., Perry, J. C., Paré, M.-È., & Kronick, R. (2024). The mental health of migrants living in limbo: A mixed-methods systematic review with meta-analysis. *Psychiatry Research*, 337, 115931. <https://doi.org/10.1016/j.psychres.2024.115931>
- Creech, S. K., & Misca, G. (2017). Parenting with PTSD: A Review of Research on the Influence of PTSD on Parent-Child Functioning in Military and Veteran Families. *Frontiers in Psychology*, 8.
<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.01101>
- Crisp, J. (2003). Refugees and the Global Politics of Asylum. *The Political Quarterly*, 74(s1), 75–87. <https://doi.org/10.1111/j.1467-923X.2003.00582.x>
- Crittenden, P. (1981). Abusing, Neglecting, Problematic, and Adequate Dyads: Differentiating by Patterns of Interaction. *Merrill-Palmer Quarterly of Behavior and Development*, 27, 201–218. <https://doi.org/10.2307/23083982>
- Cuellar, J., Jones, D. J., & Sterrett, E. (2015). Examining Parenting in the Neighborhood Context: A Review. *Journal of Child and Family Studies*, 24(1), 195–219.
<https://doi.org/10.1007/s10826-013-9826-y>

- Dajani, R. (2019). *What is the secret sauce? The story of the We Love Reading program*.
- Dalgaard, N. T., & Montgomery, E. (2015). Disclosure and silencing: A systematic review of the literature on patterns of trauma communication in refugee families. *Transcultural Psychiatry*, 52(5), 579–593.
<https://doi.org/10.1177/1363461514568442>
- Dalgaard, N. T., Thøgersen, M. H., Væver, M. S., & Montgomery, E. (2020). Family violence in traumatized refugee families: A mixed methods study of mother/child dyadic functioning, parental symptom levels and children's psychosocial adjustment. *Nordic Psychology*, 72(2), 83–104.
<https://doi.org/10.1080/19012276.2019.1653221>
- Darling, N., & Steinberg, L. (1993). Parenting style as context: An integrative model. *Psychological Bulletin*, 113(3), 487–496. <https://doi.org/10.1037/0033-2909.113.3.487>
- Dattani, S., Rodés-Guirao, L., Ritchie, H., & Roser, M. (2023). Mental Health. *Our World in Data*. <https://ourworldindata.org/mental-health>
- Davis, M., Bilms, J., & Suveg, C. (2017). In Sync and in Control: A Meta-Analysis of Parent–Child Positive Behavioral Synchrony and Youth Self-Regulation. *Family Process*, 56(4), 962–980. <https://doi.org/10.1111/famp.12259>
- Dawood, N. (2020). *The Word's Cook*. <https://welovereading.org/product/the-words-cook/>
- Deans, C. L. (2020). Maternal sensitivity, its relationship with child outcomes, and interventions that address it: A systematic literature review. *Early Child Development and Care*, 190(2), 252–275.
<https://doi.org/10.1080/03004430.2018.1465415>
- DeepL. (2024). *DeepL Translate*. <https://www.deepl.com/translator>

- Delilovic, S., Hagström, A., Shedrawy, J., Hollander, A. C., Lönnroth, K., & Hasson, H. (2023). Is legal status associated with mental illness among newly arrived refugees in Sweden: An epidemiological study. *BMC Psychiatry*, 23(1), 197. <https://doi.org/10.1186/s12888-023-04679-y>
- Deneault, A.-A., Cabrera, N. J., & Bureau, J.-F. (2022). A meta-analysis on observed paternal and maternal sensitivity. *Child Development*, n/a(n/a). <https://doi.org/10.1111/cdev.13832>
- Deslippe, D. A. (1996). Organized Labor, National Politics, and Second-Wave Feminism in the United States, 1965–1975. *International Labor and Working-Class History*, 49, 143–165. <https://doi.org/10.1017/S0147547900001745>
- Dijk, S. H. B. van, Brusse-Keizer, M. G. J., Bucsán, C. C., Palen, J. van der, Doggen, C. J. M., & Lenferink, A. (2023). Artificial intelligence in systematic reviews: Promising when appropriately used. *BMJ Open*, 13(7), e072254. <https://doi.org/10.1136/bmjopen-2023-072254>
- Dishion, T. J., Mun, C. J., Tein, J.-Y., Kim, H., Shaw, D. S., Gardner, F., Wilson, M. N., & Peterson, J. (2017). The Validation of Macro and Micro Observations of Parent–Child Dynamics Using the Relationship Affect Coding System in Early Childhood. *Prevention Science : The Official Journal of the Society for Prevention Research*, 18(3), 268–280. <https://doi.org/10.1007/s11121-016-0697-5>
- Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., Mitchell, L., Harris, K. D., Kloumann, I. M., Bagrow, J. P., Megerdooian, K., McMahon, M. T., Tivnan, B. F., & Danforth, C. M. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences*, 112(8), 2389–2394. <https://doi.org/10.1073/pnas.1411678112>
- Doktorlari. (2019). *Multi-Sectoral Needs Assessment of Syrian Refugees in Turkey*.

- Dowdall, N., Melendez-Torres, G. J., Murray, L., Gardner, F., Hartford, L., & Cooper, P. J. (2020). Shared Picture Book Reading Interventions for Child Language Development: A Systematic Review and Meta-Analysis. *Child Development*, 91(2), e383–e399. <https://doi.org/10.1111/cdev.13225>
- Dybdahl, R. (2001). Children and Mothers in War: An Outcome Study of a Psychosocial Intervention Program. *Child Development*, 72(4), 1214–1230. <https://doi.org/10.1111/1467-8624.00343>
- Eagle Prof., G., & Kaminer, D. (2013). Continuous traumatic stress: Expanding the lexicon of traumatic stress. *Peace and Conflict*, 19(2), 85–99. <https://doi.org/10.1037/a0032485>
- Eggerman, M., & Panter-Brick, C. (2010). Suffering, hope, and entrapment: Resilience and cultural values in Afghanistan. *Social Science & Medicine* (1982), 71(1–2), 71–83. <https://doi.org/10.1016/j.socscimed.2010.03.023>
- Eisenberg, N., Cumberland, A., & Spinrad, T. L. (1998). Parental Socialization of Emotion. *Psychological Inquiry*, 9(4), 241–273.
- El Arab, R., & Sagbakken, M. (2019). Child marriage of female Syrian refugees in Jordan and Lebanon: A literature review. *Global Health Action*, 12(1), 1585709. <https://doi.org/10.1080/16549716.2019.1585709>
- Elias, J., Meksassi, B., Brown, F. L., Masri, R. E., Ali, R., Chaar, S., Roberts, B., McKee, M., Asmar, M. K., Chammay, R. E., & Singh, N. S. (2025). Lebanese and Syrian refugee parents' experiences of accessing mental health care for their children in Lebanon: Findings from a qualitative study. *PLOS Mental Health*, 2(4), e0000305. <https://doi.org/10.1371/journal.pmen.0000305>
- Eltanamy, H., Leijten, P., Jak, S., & Overbeek, G. (2021). Parenting in Times of War: A Meta-Analysis and Qualitative Synthesis of War Exposure, Parenting, and Child

- Adjustment. *Trauma, Violence, & Abuse*, 22(1), 147–160.
<https://doi.org/10.1177/1524838019833001>
- Eltanamly, H., Leijten, P., van Rooij, F., & Overbeek, G. (2022). Parenting in times of refuge: A qualitative investigation. *Family Process*, 61(3), 1248–1263.
<https://doi.org/10.1111/famp.12717>
- Etzold, B., & Fechter, A.-M. (2022). Unsettling protracted displacement: Connectivity and mobility beyond ‘Limbo.’ *Journal of Ethnic and Migration Studies*, 0(0), 1–18. <https://doi.org/10.1080/1369183X.2022.2090153>
- European External Action Service (EEAS). (2022). *What is Science Diplomacy?*
https://www.eeas.europa.eu/eeas/what-science-diplomacy_en
- Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion (Washington, D.C.)*, 7(2), 336–353. <https://doi.org/10.1037/1528-3542.7.2.336>
- Fazel, M., & Betancourt, T. S. (2018). Preventive mental health interventions for refugee children and adolescents in high-income settings. *The Lancet Child and Adolescent Health*, 2(2), 121. [https://doi.org/10.1016/S2352-4642\(17\)30147-5](https://doi.org/10.1016/S2352-4642(17)30147-5)
- Feldman, H. M., Campbell, T. F., Kurs-Lasky, M., Rockette, H. E., Dale, P. S., Colborn, D. K., & Paradise, J. L. (2005). Concurrent and Predictive Validity of Parent Reports of Child Language at Ages 2 and 3 Years. *Child Development*, 76(4), 856–868. <https://doi.org/10.1111/j.1467-8624.2005.00882.x>
- Feldman, R. (2023). Father contribution to human resilience. *Development and Psychopathology*, 35(5), 2402–2419. <https://doi.org/10.1017/S0954579423000354>
- Fellmeth, G., Plugge, E. H., Nosten, S., Oo, M. M., Fazel, M., Charunwatthana, P., Nosten, F., Fitzpatrick, R., & McGready, R. (2018). Living with severe perinatal depression: A qualitative study of the experiences of labour migrant and refugee

- women on the Thai-Myanmar border. *BMC Psychiatry*, 18(1), 229.
<https://doi.org/10.1186/s12888-018-1815-7>
- Fiese, B. H., Tomcho, T. J., Douglas, M., Josephs, K., Poltrock, S., & Baker, T. (2002). A review of 50 years of research on naturally occurring family routines and rituals: Cause for celebration? *Journal of Family Psychology: JFP: Journal of the Division of Family Psychology of the American Psychological Association (Division 43)*, 16(4), 381–390. <https://doi.org/10.1037//0893-3200.16.4.381>
- Fiorini, N., Canese, K., Starchenko, G., Kireev, E., Kim, W., Miller, V., Osipov, M., Kholodov, M., Ismagilov, R., Mohan, S., Ostell, J., & Lu, Z. (2018). Best Match: New relevance search for PubMed. *PLOS Biology*, 16(8), e2005343.
<https://doi.org/10.1371/journal.pbio.2005343>
- Fogel, A., & Thelen, E. (1987). Development of early expressive and communicative action: Reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology*, 23(6), 747–761. <https://doi.org/10.1037/0012-1649.23.6.747>
- Fogle, D. O. (1978). Learned helplessness and learned restlessness. *Psychotherapy: Theory, Research & Practice*, 15(1), 39–47. <https://doi.org/10.1037/h0085839>
- Fourment, K., Nóbrega, M., Conde, G., del Prado, J. N., & Mesman, J. (2021). Maternal sensitivity in rural Andean and Amazonian Peru. *Attachment & Human Development*, 23(2), 134–149. <https://doi.org/10.1080/14616734.2020.1828515>
- Fredman, S. J., Beck, J. G., Shnaider, P., Le, Y., Pukay-Martin, N. D., Pentel, K. Z., Monson, C. M., Simon, N. M., & Marques, L. (2017). Longitudinal Associations Between PTSD Symptoms and Dyadic Conflict Communication Following a Severe Motor Vehicle Accident. *Behavior Therapy*, 48(2), 235–246.
<https://doi.org/10.1016/j.beth.2016.05.001>

- Frömke, C., Kirstein, M., & Zapf, A. (2022). A semiparametric approach for meta-analysis of diagnostic accuracy studies with multiple cut-offs. *Research Synthesis Methods*, 13(5), 612–621. <https://doi.org/10.1002/jrsm.1579>
- Gardner, F. (2000). Methodological Issues in the Direct Observation of Parent–Child Interaction: Do Observational Findings Reflect the Natural Behavior of Participants? *Clinical Child and Family Psychology Review*, 3(3), 185–198. <https://doi.org/10.1023/A:1009503409699>
- Gare, A. (2000). Aleksandr Bogdanov and Systems Theory. *Democracy & Nature*, 6(3), 341–359. <https://doi.org/10.1080/10855660020020230>
- Gentzler, A. L., Contreras-Grau, J. M., Kerns, K. A., & Weimer, B. L. (2005). Parent–Child Emotional Communication and Children’s Coping in Middle Childhood. *Social Development*, 14(4), 591–612. <https://doi.org/10.1111/j.1467-9507.2005.00319.x>
- Ghandour, L. A., Anouti, S., Lotfi, T., Meho, L., Kashash, R., Al-Akkawi, A., Majed, A., Akl, E., & Afifi, R. A. (2024). Parenting a High and Growing Population of Youth in the Arab Region: A Scoping Review for an Evidence-Informed Research Agenda. *Journal of Adolescent Health*, 75(1), 16–25. <https://doi.org/10.1016/j.jadohealth.2024.02.033>
- Gillespie, S., Banegas, J., Maxwell, J., Chan, A. C. Y., Darawshy, N. A.-S., Wasil, A. R., Marsalis, S., & Gewirtz, A. (2022). Parenting Interventions for Refugees and Forcibly Displaced Families: A Systematic Review. *Clinical Child and Family Psychology Review*, 25(2), 395–412. <https://doi.org/10.1007/s10567-021-00375-z>
- Giummarra, M. J., Lau, G., Grant, G., & Gabbe, B. J. (2020). A systematic review of the association between fault or blame-related attributions and procedures after

- transport injury and health and work-related outcomes. *Accident; Analysis and Prevention*, 135, 105333. <https://doi.org/10.1016/j.aap.2019.105333>
- Goldkuhle, M., Dimaki, M., Gartlehner, G., Monsef, I., Dahm, P., Glossmann, J.-P., Engert, A., von Tresckow, B., & Skoetz, N. (2018). Nivolumab for adults with Hodgkin's lymphoma (a rapid review using the software RobotReviewer). *The Cochrane Database of Systematic Reviews*, 7(7), CD012556. <https://doi.org/10.1002/14651858.CD012556.pub2>
- Goodman, S. H., Rouse, M. H., Connell, A. M., Broth, M. R., Hall, C. M., & Heyward, D. (2011). Maternal Depression and Child Psychopathology: A Meta-Analytic Review. *Clinical Child and Family Psychology Review*, 14(1), 1–27. <https://doi.org/10.1007/s10567-010-0080-1>
- Gopalkrishnan, N. (2018). Cultural Diversity and Mental Health: Considerations for Policy and Practice. *Frontiers in Public Health*, 6. <https://doi.org/10.3389/fpubh.2018.00179>
- Greene, C. A., Haisley, L., Wallace, C., & Ford, J. D. (2020). Intergenerational effects of childhood maltreatment: A systematic review of the parenting practices of adult survivors of childhood abuse, neglect, and violence. *Clinical Psychology Review*, 80, 101891. <https://doi.org/10.1016/j.cpr.2020.101891>
- Greene, M. C., Jordans, M. J. D., Kohrt, B. A., Ventevogel, P., Kirmayer, L. J., Hassan, G., Chiumento, A., van Ommeren, M., & Tol, W. A. (2017). Addressing culture and context in humanitarian response: Preparing desk reviews to inform mental health and psychosocial support. *Conflict and Health*, 11(1), 21. <https://doi.org/10.1186/s13031-017-0123-z>

- Grzeškowiak, M. (2024). When Legal Inclusion is not Enough: The “Uganda Model” of Refugee Protection on the Brink of Failure. *Refugee Survey Quarterly*, 43(1), 95–112. <https://doi.org/10.1093/rsq/hdad022>
- Grzymala-Kazłowska, A. (2018). From connecting to social anchoring: Adaptation and ‘settlement’ of Polish migrants in the UK. *Journal of Ethnic and Migration Studies*, 44(2), 252–269. <https://doi.org/10.1080/1369183X.2017.1341713>
- Guo, E., Gupta, M., Deng, J., Park, Y.-J., Paget, M., & Naugler, C. (2023). *Automated Paper Screening for Clinical Reviews Using Large Language Models* (arXiv:2305.00844). arXiv. <http://arxiv.org/abs/2305.00844>
- Haas, B. M. (2017). Citizens-in-Waiting, Deportees-in-Waiting: Power, Temporality, and Suffering in the U.S. Asylum System. *Ethos*, 45(1), 75–97. <https://doi.org/10.1111/etho.12150>
- Hadfield, K. (2020). *Family Intervention for Empowerment through Reading and Education (FIERCE)* [Dataset]. <https://doi.org/10.1257/rct.6523-1.20000000000000002>
- Hadfield, K., Al-Hamad, M., Bakhti, R., Dajani, R., El Kharouf, A., Michalek, J., Mukunzi, J., Qtaishat, L., Sethi, T., von Stumm, S., & Mareschal, I. (2022). Predictors of Literacy and Attitudes Toward Reading Among Syrian Refugee Children in Jordan. *International Journal of Early Childhood*, 1–21. <https://doi.org/10.1007/s13158-022-00334-x>
- Hadfield, K., Al-Soleiti, M., Dajani, R., Mareschal, I., & Panter-Brick, C. (2024). Father Involvement and Child Development: A Prospective Study of Syrian Refugee Families. *Journal of Child and Family Studies*, 33, 1–14. <https://doi.org/10.1007/s10826-024-02809-y>

- Hadfield, K., Dajani, R., Kharouf, A. E., Michalek, J., Qtaishat, L., Stumm, S. von, & Mareschal, I. (2024). *Effectiveness of a community-led shared book reading intervention in Syrian refugee children: A randomised controlled trial*. OSF. <https://doi.org/10.31234/osf.io/5u2tj>
- Hadfield, K., & Ungar, M. (2018). Family resilience: Emerging trends in theory and practice. *Journal of Family Social Work, 21*(2), 81–84. <https://doi.org/10.1080/10522158.2018.1424426>
- Halamoda-Kenzaoui, B., Rolland, E., Piovesan, J., Puertas Gallardo, A., & Bremer-Hoffmann, S. (2022). Toxic effects of nanomaterials for health applications: How automation can support a systematic review of the literature? *Journal of Applied Toxicology, 42*(1), 41–51. <https://doi.org/10.1002/jat.4204>
- Hall, J., & Werner, K. (2022). Trauma and Trust: How War Exposure Shapes Social and Institutional Trust Among Refugees. *Frontiers in Psychology, 13*, 786838. <https://doi.org/10.3389/fpsyg.2022.786838>
- Harrison, R., Jones, B., Gardner, P., & Lawton, R. (2021). Quality assessment with diverse studies (QuADS): An appraisal tool for methodological and reporting quality in systematic reviews of mixed- or multi-method studies. *BMC Health Services Research, 21*(1), 144. <https://doi.org/10.1186/s12913-021-06122-y>
- Harvey, N. (2007). Use of heuristics: Insights from forecasting research. In *Judgement and Choice: Perspectives on the Work of Daniel Kahneman*. Psychology Press.
- Hawes, D. J., & Dadds, M. R. (2006). Assessing Parenting Practices Through Parent-Report and Direct Observation During Parent-Training. *Journal of Child and Family Studies, 15*(5), 555–568. <https://doi.org/10.1007/s10826-006-9029-x>
- Hayat, M. (2015). *Questions in a Traveling Bag*. <https://welovereading.org/product/questions-in-a-traveling-bag/>

- Hayden, E. P., Durbin, C. E., Klein, D. N., & Olino, T. M. (2010). Maternal personality influences the relationship between maternal reports and laboratory measures of child temperament. *Journal of Personality Assessment*, 92(6), 586–593.
<https://doi.org/10.1080/00223891.2010.513308>
- Hayward, R. (2012). The invention of the psychosocial: An introduction. *History of the Human Sciences*, 25(5), 3–12. <https://doi.org/10.1177/0952695112471658>
- Hecker, T., Kyaruzi, E., Borchardt, J., & Scharpf, F. (2022). Factors Contributing to Violence Against Children: Insights From a Multi-informant Study Among Family-Triads From Three East-African Refugee Camps. *Journal of Interpersonal Violence*, 37(15–16), NP14507–NP14537.
<https://doi.org/10.1177/08862605211013960>
- Hein, S., Bick, J., Issa, G., Aoude, L., Maalouf, C., Awar, A., Nourallah, S., Zonderman, A. L., Householder, S., Katsovich, L., Khoshnood, K., Moore, C., Salah, R., Britto, P. R., Leckman, J. F., & Ponguta, L. A. (2020). Maternal perceptions of father involvement among refugee and disadvantaged families in Beirut, Lebanon. *PloS One*, 15(3), e0229670. <https://doi.org/10.1371/journal.pone.0229670>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83.
<https://doi.org/10.1017/S0140525X0999152X>
- Hickman. (2008). *Immigration and social cohesion in the UK* [§].
<https://www.jrf.org.uk/sites/default/files/jrf/migrated/files/2230-deprivation-cohesion-immigration.pdf>
- Hinchey, L. M. E., Chammaa, M., Ruvoilo Grasser, L., Saad, B., Gorski, K., & Javanbakht, A. (2023). What happened matters: Trauma type and cumulative trauma exposure in refugee youth psychopathology. *Psychological Trauma*:

Theory, Research, Practice, and Policy, No Pagination Specified-No Pagination Specified. <https://doi.org/10.1037/tra0001618>

- Hinchey, L. M. E., Khalil, D., & Javanbakht, A. (2023). Practical approaches to conducting biopsychosocial research with refugee and internally displaced communities. *Comprehensive Psychoneuroendocrinology*, *16*, 100217. <https://doi.org/10.1016/j.cpnec.2023.100217>
- Hodes, M., & Vostanis, P. (2019). Practitioner Review: Mental health problems of refugee children and adolescents and their management. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *60*(7), 716–731. <https://doi.org/10.1111/jcpp.13002>
- Hughes, H. M., & Haynes, S. N. (1978). Structured laboratory observation in the behavioral assessment of parent-child interactions: A methodological critique. *Behavior Therapy*, *9*(3), 428–447. [https://doi.org/10.1016/S0005-7894\(78\)80088-4](https://doi.org/10.1016/S0005-7894(78)80088-4)
- Huntley, F., Ulph, F., El-Khani, A., & Calam, R. (2021). The parenting process in Syrian refugees: A grounded theory study. *Peace and Conflict: Journal of Peace Psychology*, *27*(1), 67–80. <https://doi.org/10.1037/pac0000533>
- Hussein, K. S. (2015). *Measuring Lexical Richness through Type-Token Curve: A Corpus-Based Analysis of Arabic and English Texts*.
- Hvidtfeldt, C., Petersen, J. H., & Norredam, M. (2022). Waiting for family reunification and the risk of mental disorders among refugee fathers: A 24-year longitudinal cohort study from Denmark. *Social Psychiatry and Psychiatric Epidemiology*, *57*(5), 1061–1072. <https://doi.org/10.1007/s00127-021-02170-1>
- Ingabire, C. M., & Richters, A. (2020). Suicidal Ideation and Behavior Among Congolese Refugees in Rwanda: Contributing Factors, Consequences, and Support

Mechanisms in the Context of Culture. *Frontiers in Psychiatry*, 11, 299.

<https://doi.org/10.3389/fpsy.2020.00299>

Irace, K. (2011). Coparenting in diverse family systems. In J. P. McHale (Ed.), *Coparenting: A conceptual and clinical examination of family systems* (pp. 15–37). American Psychological Association. <https://doi.org/10.1037/12328-001>

Itani, T., Jacobsen, K. H., & Kraemer, A. (2017). Suicidal ideation and planning among Palestinian middle school students living in Gaza Strip, West Bank, and United Nations Relief and Works Agency (UNRWA) camps. *International Journal of Pediatrics and Adolescent Medicine*, 4(2), 54–60.

<https://doi.org/10.1016/j.ijpam.2017.03.003>

Ivimey-Cook, E. R., Noble, D. W. A., Nakagawa, S., Lajeunesse, M. J., & Pick, J. L. (2023). Advice for improving the reproducibility of data extraction in meta-analysis. *Research Synthesis Methods*. <https://doi.org/10.1002/jrsm.1663>

Jabbar, S. A., & Zaza, H. I. (2019). Post-traumatic Stress and Depression (PTSD) and general anxiety among Iraqi refugee children: A case study from Jordan. *Early Child Development and Care*, 189(7), 1114–1134.

<https://doi.org/10.1080/03004430.2017.1369974>

Jacobs, C., Kyamusugulwa, P. M., Kubiha, S. L., Assumani, I., Ruhamy, J., & Katembera, R. S. (2022). Is translocality a hidden solution to overcome protracted displacement in the DR Congo? *Journal of Ethnic and Migration Studies*, 0(0), 1–15. <https://doi.org/10.1080/1369183X.2022.2090154>

Jain, T., Cooper, L., Khauli, N., Pike, K., Khoury, B., & Hermosilla, S. (2025). Social connections related to health and well-being needs identified by children and adolescents affected by displacement in Lebanon: A participatory research study.

Journal of Migration and Health, 11, 100284.

<https://doi.org/10.1016/j.jmh.2024.100284>

Jensen, S. K., Placencio-Castro, M., Murray, S. M., Brennan, R. T., Goshev, S., Farrar, J., Yousafzai, A., Rawlings, L. B., Wilson, B., Habyarimana, E., Sezibera, V., & Betancourt, T. S. (2021). Effect of a home-visiting parenting program to promote early childhood development and prevent violence: A cluster-randomized trial in Rwanda. *BMJ Global Health*, 6(1), e003508. <https://doi.org/10.1136/bmjgh-2020-003508>

Jones, E., & Wessely, S. (2005). *Shell Shock to PTSD: Military Psychiatry from 1900 to the Gulf War*. Psychology Press. <https://doi.org/10.4324/9780203086827>

Jordans, M. J. D., Bakolis, I., Arnous, M., Koppenol-Gonzalez, G. V., Tossyeh, F., Chen, A., & Miller, K. E. (2023). Effectiveness of the caregiver support intervention on child psychosocial wellbeing among Syrian refugees in Lebanon: Mediation and secondary analysis of a Randomized Controlled Trial. *Child Abuse & Neglect*, 106335. <https://doi.org/10.1016/j.chiabu.2023.106335>

Jordans, M. J. D., Semrau, M., Thornicroft, G., & Ommeren, M. van. (2012). Role of current perceived needs in explaining the association between past trauma exposure and distress in humanitarian settings in Jordan and Nepal. *The British Journal of Psychiatry*, 201(4), 276–281.

<https://doi.org/10.1192/bjp.bp.111.102137>

Kamal, F., & Killian, K. D. (2016). Invisible Lives and Hidden Realities of Undocumented Youth. *Refuge: Canada's Journal on Refugees*, 31(2), Article 2.

<https://doi.org/10.25071/1920-7336.40310>

Kamruzzaman, P., Williams, K., Wardak, A., Kabir, M. E., & Ayobi, Y. (2022). Exploring dignity in the context of displacement – evidence from Rohingyas in Bangladesh

and IDPs in Afghanistan. *Third World Quarterly*, 43(8), 1854–1874.

<https://doi.org/10.1080/01436597.2022.2074387>

Karam, E. G., Fayyad, J. A., Farhat, C., Pluess, M., Haddad, Y. C., Tabet, C. C., Farah, L., & Kessler, R. C. (2019). Role of childhood adversities and environmental sensitivity in the development of post-traumatic stress disorder in war-exposed Syrian refugee children and adolescents. *British Journal of Psychiatry*, 214(6), 354–360. <https://doi.org/10.1192/bjp.2018.272>

Kebede, M. M., Le Cornet, C., & Fortner, R. T. (2023). In-depth evaluation of machine learning methods for semi-automating article screening in a systematic review of mechanistic literature. *Research Synthesis Methods*, 14(2), 156–172. <https://doi.org/10.1002/jrsm.1589>

Khraisha, Q., Abujaber, N., Carpenter, S., Crossen, R. J., Kappenberg, J., Kelly, R., Murphy, C., Norton, O., Put, S. M., Schnoebelen, K., Warraitch, A., Roney, S., & Hadfield, K. (2024). Parenting and mental health in protracted refugee situations: A systematic review. *Comprehensive Psychiatry*, 152536. <https://doi.org/10.1016/j.comppsy.2024.152536>

Khraisha, Q., Put, S., Kappenberg, J., Warraitch, A., & Hadfield, K. (2024). Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. *Research Synthesis Methods*, 15(4), 616–626. <https://doi.org/10.1002/jrsm.1715>

Khraisha, Q., Sawalha, L., Hadfield, K., Al-Soleiti, M., Dajani, R., & Panter-Brick, C. (2024). Coparenting, mental health, and the pursuit of dignity: A systems-level analysis of refugee father-mother narratives. *Social Science & Medicine*, 340, 116452. <https://doi.org/10.1016/j.socscimed.2023.116452>

- Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher*, 42(8), 846–854.
<https://doi.org/10.1080/0142159X.2020.1755030>
- Kira, I., Ashby, J. S., Lewandowski, L., Alawneh, A. N., Mohanesh, J., & Odenat, L. (2013). Advances in continuous traumatic stress theory: Traumatogenic dynamics and consequences of intergroup conflict: The Palestinian adolescents case. *Psychology*, 4(4), 396–409.
- Kira, I., Lewandowski, L., Somers, C. L., Yoon, J. S., & Chiodo, L. (2012). The effects of trauma types, cumulative trauma, and PTSD on IQ in two highly traumatized adolescent groups. *Psychological Trauma: Theory, Research, Practice, and Policy*, 4(1), 128–139. <https://doi.org/10.1037/a0022121>
- Kite, M. E., Russo, N. F., Brehm, S. S., Fouad, N. A., Hall, C. C., Hyde, J. S., & Keita, G. P. (2001). Women psychologists in academe. Mixed progress, unwarranted complacency. *The American Psychologist*, 56(12), 1080–1098.
- Klein, E. K., Bond, L., McLean, K. E., Feika, M., Bah, A. J., & Betancourt, T. S. (2023). Navigating the tension between fatherhood ideals and realities of a post-conflict setting: A phenomenological study of former child soldiers in Sierra Leone. *SSM - Qualitative Research in Health*, 3, 100227.
<https://doi.org/10.1016/j.ssmqr.2023.100227>
- Koenen, K. C., Ratanatharathorn, A., Ng, L., McLaughlin, K. A., Bromet, E. J., Stein, D. J., Karam, E. G., Ruscio, A. M., Benjet, C., Scott, K., Atwoli, L., Petukhova, M., Lim, C. C. W., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Bunting, B., Ciutan, M., de Girolamo, G., ... Kessler, R. C. (2017). Posttraumatic stress disorder in the World Mental Health Surveys. *Psychological Medicine*, 47(13), 2260–2274. <https://doi.org/10.1017/S0033291717000708>

- Kouta, C., Velonaki, V. S., Lazzarino, R., Rousou, E., Apostolara, P., Doru, A., Dudau, V., Kalokairinou, A., Licciardello, O., Mauceri, M., Nikolaidou, E., Pezzella, A., Samartzi, C., Kuckert-Wöstheinrich, A., & Papadopoulos, I. (2022). Empowering the Migrant and Refugee Family's Parenting Skills: A Literature Review. *Journal of Community Health Research*. <https://doi.org/10.18502/jchr.v11i4.11735>
- Krijnen, L. J. G., Verhoeven, M., & van Baar, A. L. (2023). Observing mother-child interaction in a free-play vs. A structured task context and its relationship with preterm and term born toddlers' psychosocial outcomes. *Frontiers in Child and Adolescent Psychiatry*, 2. <https://doi.org/10.3389/frcha.2023.1176560>
- Kuppens, S., & Ceulemans, E. (2019). Parenting Styles: A Closer Look at a Well-Known Concept. *Journal of Child and Family Studies*, 28(1), 168–181. <https://doi.org/10.1007/s10826-018-1242-x>
- Kuru, Ni., & Ungar, M. (2021). Refugee Children's Resilience: A Qualitative Social Ecological Study of Life in a Camp. *Journal of Refugee Studies*, 34(4), 4207–4224. <https://doi.org/10.1093/jrs/feaa092>
- Lakkis, N. A., Osman, M. H., Aoude, L. C., Maalouf, C. J., Issa, H. G., & Issa, G. M. (2020). A Pilot Intervention to Promote Positive Parenting in Refugees from Syria in Lebanon and Jordan. *Frontiers in Psychiatry*, 11. <https://doi.org/10.3389/fpsy.2020.00257>
- Lamb, M. (2010). How do fathers influence children's development? Let me count the ways. *The Role of the Father in Child Development*, 1–26.
- Lamb, M. E. (1977). The development of mother-infant and father-infant attachments in the second year of life. *Developmental Psychology*, 13(6), 637–648. <https://doi.org/10.1037/0012-1649.13.6.637>

- Lamb, M. E. (2000). The History of Research on Father Involvement: An Overview. *Marriage & Family Review*, 29(2–3), 23–42.
https://doi.org/10.1300/J002v29n02_03
- Lamb, M. E., Pleck, J. H., Charnov, E. L., & Levine, J. A. (1985). Paternal Behavior in Humans. *American Zoologist*, 25(3), 883–894.
- Lamb, M., Sternberg, K., & Thompson, R. (1997). The effects of divorce and custody arrangements on children's behavior, development, and adjustment. *Family Court Review*, 35, 393–404. <https://doi.org/10.1111/j.174-1617.1997.tb00482.x>
- Lansford, J. E. (2022a). Annual Research Review: Cross-cultural similarities and differences in parenting. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 63(4), 466–479. <https://doi.org/10.1111/jcpp.13539>
- Lansford, J. E. (2022b). Cross-Cultural Similarities and Differences in Parenting. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 63(4), 466–479.
<https://doi.org/10.1111/jcpp.13539>
- Law, J., & Roy, P. (2008). Parental Report of Infant Language Skills: A Review of the Development and Application of the Communicative Development Inventories. *Child and Adolescent Mental Health*, 13(4), 198–206.
<https://doi.org/10.1111/j.1475-3588.2008.00503.x>
- Lawrence, A., Houghton, J., Thomas, J., & Weldon, P. (2014). *Where Is the Evidence? Realising the Value of Grey Literature for Public Policy & Practice, A Discussion Paper*.
- Lebiger-Vogel, J., Rickmeyer, C., Leuzinger-Bohleber, M., & Meurs, P. (2022). Fostering Emotional Availability in Mother-Child-Dyads With an Immigrant Background: A Randomized-Controlled-Trial on the Effects of the Early Prevention Program First

Steps. *Frontiers in Psychology*, 13, 790244.

<https://doi.org/10.3389/fpsyg.2022.790244>

Leckman, J. F., Ponguta, L. A., Pavarini, G., Hein, S. D., McCarthy, M. F., Staiti, H., Hanöz-Penney, S., Rubinstein, J., Pruett, K. D., Yazgan, M. Y., Fallon, N. S., Hartl, F. J., Ziv, M., Salah, R., Britto, P. R., Fitzpatrick, S., & Panter-Brick, C. (2021). Love and peace across generations: Biobehavioral systems and global partnerships. *Comprehensive Psychoneuroendocrinology*, 8, 100092.

<https://doi.org/10.1016/j.cpneec.2021.100092>

Leclère, C., Viaux, S., Avril, M., Achard, C., Chetouani, M., Missonnier, S., & Cohen, D. (2014). Why Synchrony Matters during Mother-Child Interactions: A Systematic Review. *PLoS ONE*, 9(12), e113571.

<https://doi.org/10.1371/journal.pone.0113571>

Lembcke, H., Buchmüller, T., & Leyendecker, B. (2020). Refugee mother-child dyads' hair cortisol, post-traumatic stress, and affectionate parenting.

Psychoneuroendocrinology, 111, 104470.

<https://doi.org/10.1016/j.psyneuen.2019.104470>

Lewando Hundt, G., Chatty, D., Thabet, A. A., & Abuatayas, H. (2004). Advocating Multi-disciplinarity in Studying Complex Emergencies: The Limitations of a Psychological Approach to Understanding How Young People Cope with Prolonged Conflict in Gaza. *Journal of Biosocial Science*, 36, 417–431.

<https://doi.org/10.1017/S0021932004006649>

Lewek, M., & Nawer, A. (2017). *Kindheit im Wartezustand: Studie zur Situation von Kindern und Jugendlichen in Flüchtlingsunterkünften in Deutschland 2017*.

Lewis, M. W. (2000). Exploring Paradox: Toward a More Comprehensive Guide. *The Academy of Management Review*, 25(4), 760–776. <https://doi.org/10.2307/259204>

- Li, M., Sun, J., & Tan, X. (2024). Evaluating the effectiveness of large language models in abstract screening: A comparative analysis. *Systematic Reviews*, 13(1), 219. <https://doi.org/10.1186/s13643-024-02609-x>
- Li, X., Hu, Y., Huang, C.-Y. S., & Chuang, S. S. (2021). Beyond W.E.I.R.D. (Western, educated, industrial, rich, democratic)-centric theories and perspectives: Masculinity and fathering in Chinese societies. *Journal of Family Theory & Review*, 13(3), 317–333. <https://doi.org/10.1111/jftr.12403>
- Loescher, G. (2001). The UNHCR and World Politics: State Interests vs. Institutional Autonomy. *The International Migration Review*, 35(1), 33–56.
- Loutet, M. G., Logie, C. H., Okumu, M., Coelho, M., Blondeel, K., McAlpine, A., Mackenzie, F., Lukone, S. O., Kisubi, N., Lukone, J. O., Isaac, A. M., Kyambadde, P., & Toskin, I. (2024). Qualitative insights on sexual health counselling from refugee youth in Bidi Bidi Refugee Settlement, Uganda: Advancing contextual considerations for brief sexuality-related communication in a humanitarian setting. *PLOS ONE*, 19(11), e0310682. <https://doi.org/10.1371/journal.pone.0310682>
- Lovejoy, M. C., Graczyk, P. A., O'Hare, E., & Neuman, G. (2000). Maternal depression and parenting behavior: A meta-analytic review. *Clinical Psychology Review*, 20(5), 561–592. [https://doi.org/10.1016/s0272-7358\(98\)00100-7](https://doi.org/10.1016/s0272-7358(98)00100-7)
- Mahuli, S. A., Rai, A., Mahuli, A. V., & Kumar, A. (2023). Application ChatGPT in conducting systematic reviews and meta-analyses. *British Dental Journal*, 235(2), Article 2. <https://doi.org/10.1038/s41415-023-6132-y>
- Manoogian & Benson. (2018). *Cognitive Bias Codex* [Graphic]. This file was derived from: The Cognitive Bias Codex - 180+ biases, designed by John Manoogian III

(jm3).png:

https://commons.wikimedia.org/wiki/File:Cognitive_bias_codex_en.svg

- Maphalala, M. C., & Mpofu, N. (2017). Are we there yet? A literature study of the challenges of women academics in institutions of higher education. *Gender and Behaviour, 15*(2), 9216–9224. <https://doi.org/10.10520/EJC-b41b93666>
- Marshall, I. J., Nye, B., Kuiper, J., Noel-Storr, A., Marshall, R., Maclean, R., Soboczenski, F., Nenkova, A., Thomas, J., & Wallace, B. C. (2020). Trialstreamer: A living, automatically updated database of clinical trial reports. *Journal of the American Medical Informatics Association, 27*(12), 1903–1912. <https://doi.org/10.1093/jamia/ocaa163>
- Martin, M. J., Sturge-Apple, M. L., Davies, P. T., Romero, C. V., & Buckholz, A. (2017). A process model of the implications of spillover from coparenting conflicts into the parent–child attachment relationship in adolescence. *Development and Psychopathology, 29*(2), 417–431. <https://doi.org/10.1017/S0954579417000086>
- Masarik, A. S., & Conger, R. D. (2017). Stress and child development: A review of the Family Stress Model. *Current Opinion in Psychology, 13*, 85–90. <https://doi.org/10.1016/j.copsyc.2016.05.008>
- Matas, L., Arend, R. A., & Sroufe, L. A. (1978). Continuity of Adaptation in the Second Year: The Relationship between Quality of Attachment and Later Competence. *Child Development, 49*(3), 547–556. <https://doi.org/10.2307/1128221>
- Mathers, S. J., Hodgkiss, A., Kolancali, P., Booton, S. A., Wang, Z., & Murphy, V. A. (2024). Comparing parent-child interaction during wordless book reading, print book reading and imaginative play. *Journal of Child Language, 1*–26. <https://doi.org/10.1017/S0305000924000072>

- Matias, C. (2006). Direct Observation of Parent-Child Interaction based on Attachment Theory. *Unpublished PhD Thesis - King's College London*.
https://www.academia.edu/3803251/Direct_Observation_of_Parent_Child_Interaction_based_on_Attachment_Theory
- McEwen, F. S., Biazoli, C. E., Popham, C. M., Moghames, P., Saab, D., Fayyad, J., Karam, E., Bosqui, T., & Pluess, M. (2023). Prevalence and predictors of mental health problems in refugee children living in informal settlements in Lebanon. *Nature Mental Health*, 1(2), Article 2. <https://doi.org/10.1038/s44220-023-00017-z>
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282.
- McLeod, B. D., Weisz, J. R., & Wood, J. J. (2007). Examining the association between parenting and childhood depression: A meta-analysis. *Clinical Psychology Review*, 27(8), 986–1003. <https://doi.org/10.1016/j.cpr.2007.03.001>
- Measham, T., & Rousseau, C. (2010). Family disclosure of war trauma to children. *Traumatology*, 16(4), 85–96. <https://doi.org/10.1177/1534765610395664>
- Media Education Foundation (Director). (1998). *Edward Said: The myth of the clash of civilizations* [Video recording]. https://youtu.be/ty_-zHrfEUY?si=O5ulRdcOj6v-uOyu&t=2767
- Meijer, L., Finkenauer, C., Blankers, M., de Gee, A., Kramer, J., Shields-Zeeman, L., & Thomaes, K. (2023). Study protocol: Development and randomized controlled trial of a preventive blended care parenting intervention for parents with PTSD. *BMC Psychiatry*, 23(1), 102. <https://doi.org/10.1186/s12888-023-04548-8>
- Meijer, L., Franz, M. R., Deković, M., van Ee, E., Finkenauer, C., Kleber, R. J., van de Putte, E. M., & Thomaes, K. (2023). Towards a more comprehensive

- understanding of PTSD and parenting. *Comprehensive Psychiatry*, 127, 152423.
<https://doi.org/10.1016/j.comppsy.2023.152423>
- Meisenberg, G., & Williams, A. (2008). Are acquiescent and extreme response styles related to low intelligence and education? *Personality and Individual Differences*, 44(7), 1539–1550. <https://doi.org/10.1016/j.paid.2008.01.010>
- Menjivar, C. (1997). Immigrant kinship networks: Vietnamese, Salvadoreans and Mexicans in comparative perspective. *Journal of Comparative Family Studies*, 28(1), 1–25. <https://doi.org/10.3138/jcfs.28.1.1>
- Merry, L., Pelaez, S., & Edwards, N. C. (2017). Refugees, asylum-seekers and undocumented migrants and the experience of parenthood: A synthesis of the qualitative literature. *Globalization and Health*, 13(1).
<https://doi.org/10.1186/s12992-017-0299-4>
- Mesa-Vieira, C., Haas, A. D., Buitrago-Garcia, D., Roa-Diaz, Z. M., Minder, B., Gamba, M., Salvador, D., Gomez, D., Lewis, M., Gonzalez-Jaramillo, W. C., Pahud de Mortanges, A., Buttia, C., Muka, T., Trujillo, N., & Franco, O. H. (2022). Mental health of migrants with pre-migration exposure to armed conflict: A systematic review and meta-analysis. *The Lancet Public Health*, 7(5), e469–e481.
[https://doi.org/10.1016/S2468-2667\(22\)00061-5](https://doi.org/10.1016/S2468-2667(22)00061-5)
- Mesman, J. (2021). Video observations of sensitive caregiving “off the beaten track”: Introduction to the special issue. *Attachment & Human Development*, 23(2), 115–123. <https://doi.org/10.1080/14616734.2020.1828511>
- Mesman, J., Basweti, N., & Misati, J. (2021). Sensitive infant caregiving among the rural Gusii in Kenya. *Attachment & Human Development*, 23(2), 124–133.
<https://doi.org/10.1080/14616734.2020.1828512>

- Meyer, S., Murray, L. K., Puffer, E. S., Larsen, J., & Bolton, P. (2013). The nature and impact of chronic stressors on refugee children in Ban Mai Nai Soi camp, Thailand. *Global Public Health*, 8(9), 1027–1047.
<https://doi.org/10.1080/17441692.2013.811531>
- Meyer, S. R., Lasater, M., & Tol, W. A. (2017). Migration and Mental Health in Low- and Middle-Income Countries: A Systematic Review. *Psychiatry*, 80(4), 374–381.
<https://doi.org/10.1080/00332747.2017.1354608>
- Meyer, S. R., Yu, G., Hermosilla, S., & Stark, L. (2017). Latent class analysis of violence against adolescents and psychosocial outcomes in refugee settings in Uganda and Rwanda. *Global Mental Health*, 4(e19), 1–12.
- Michalek, J., Lisi, M., Binetti, N., Ozkaya, S., Hadfield, K., Dajani, R., & Mareschal, I. (2022). War-related trauma linked to increased sustained attention to threat in children. *Child Development*, cdev.13739. <https://doi.org/10.1111/cdev.13739>
- Michelson, M., & Reuter, K. (2019). The significant cost of systematic reviews and meta-analyses: A call for greater involvement of machine learning to assess the promise of clinical trials. *Contemporary Clinical Trials Communications*, 16, 100443.
<https://doi.org/10.1016/j.conctc.2019.100443>
- Mielke, K., & Etzold, B. (2022). Afghans' narrowing mobility options in Pakistan and the right to transnational living: A figurational perspective. *Journal of Ethnic and Migration Studies*, 0(0), 1–18. <https://doi.org/10.1080/1369183X.2022.2090156>
- Miller, K. E. (1996). The Effects of State Terrorism and Exile on Indigenous Guatemalan Refugee Children: A Mental Health Assessment and an Analysis of Children's Narratives. *Child Development*, 67(1), 89–106. <https://doi.org/10.1111/j.1467-8624.1996.tb01721.x>

- Miller, K. E., Chen, A., Bakolis, I., Koppenol-Gonzalez, G. V., Arnous, M., Tossyeh, F., Al-Ogaily, D., & Jordans, M. J. D. (2025). Development and psychometric evaluation of the Brief Parenting Questionnaire. *BMC Psychology*, 13(1), 85. <https://doi.org/10.1186/s40359-025-02395-8>
- Miller, K. E., Chen, A., Koppenol-Gonzalez, G. V., Bakolis, I., Arnous, M., Tossyeh, F., El Hassan, A., Saleh, A., Saade, J., Nahas, N., Abboud, M., Jawad, L., & Jordans, M. J. D. (2022). Supporting parenting among Syrian refugees in Lebanon: A randomized controlled trial of the caregiver support intervention. *Journal of Child Psychology and Psychiatry*, n/a(n/a). <https://doi.org/10.1111/jcpp.13668>
- Miller, K. E., Chen, A., Koppenol-Gonzalez, G. V., Bakolis, I., Arnous, M., Tossyeh, F., El Hassan, A., Saleh, A., Saade, J., Nahas, N., Abboud, M., Jawad, L., & Jordans, M. J. D. (2023). Supporting parenting among Syrian refugees in Lebanon: A randomized controlled trial of the caregiver support intervention. *Journal of Child Psychology and Psychiatry*, 64(1), 71–82. <https://doi.org/10.1111/jcpp.13668>
- Miller, K. E., Ghalayini, H., Arnous, M., Tossyeh, F., Chen, A., van den Broek, M., Koppenol-Gonzalez, G. V., Saade, J., & Jordans, M. J. D. (2020). Strengthening parenting in conflict-affected communities: Development of the Caregiver Support Intervention. *Global Mental Health*, 7. <https://doi.org/10.1017/gmh.2020.8>
- Miller, K. E., & Jordans, M. J. D. (2016). Determinants of Children's Mental Health in War-Torn Settings: Translating Research Into Action. *Current Psychiatry Reports*, 18(6). <https://doi.org/10.1007/s11920-016-0692-3>
- Miller, K. E., Koppenol-Gonzalez, G. V., Arnous, M., Tossyeh, F., Chen, A., Nahas, N., & Jordans, M. J. D. (2020). Supporting Syrian families displaced by armed conflict:

- A pilot randomized controlled trial of the Caregiver Support Intervention. *Child Abuse and Neglect*, 106. <https://doi.org/10.1016/j.chiabu.2020.104512>
- Miller, K. E., Omidian, P., Quraishy, A. S., Quraishy, N., Nasiry, M. N., Nasiry, S., Karyar, N. M., & Yaqubi, A. A. (2006). The Afghan symptom checklist: A culturally grounded approach to mental health assessment in a conflict zone. *The American Journal of Orthopsychiatry*, 76(4), 423–433. <https://doi.org/10.1037/0002-9432.76.4.423>
- Miller, K. E., & Rasmussen, A. (2010). War exposure, daily stressors, and mental health in conflict and post-conflict settings: Bridging the divide between trauma-focused and psychosocial frameworks. *Social Science and Medicine*, 70(1), 7–16. <https://doi.org/10.1016/j.socscimed.2009.09.029>
- Miller, K. E., & Rasmussen, A. (2017). The mental health of civilians displaced by armed conflict: An ecological model of refugee distress. *Epidemiology and Psychiatric Sciences*, 26(2), 129–138. <https://doi.org/10.1017/S2045796016000172>
- Miller, K. E., & Rasmussen, A. (2024). War exposure, daily stressors, and mental health 15 years on: Implications of an ecological framework for addressing the mental health of conflict-affected populations. *Epidemiology and Psychiatric Sciences*, 33, e78. <https://doi.org/10.1017/S2045796024000830>
- Miller, & Rasmussen. (2017). The mental health of civilians displaced by armed conflict: An ecological model of refugee distress. *Epidemiology and Psychiatric Sciences*, 4(2), 1–10.
- Miller, W. R., & Seligman, M. E. (1975). Depression and learned helplessness in man. *Journal of Abnormal Psychology*, 84(3), 228–238. <https://doi.org/10.1037/h0076720>

- Miller-Graff, L. E., & Cummings, E. M. (2022). Supporting Youth and Families in Gaza: A Randomized Controlled Trial of a Family-Based Intervention Program. *International Journal of Environmental Research and Public Health*, 19(14), 8337. <https://doi.org/10.3390/ijerph19148337>
- Minuchin, S., & Fishman, H. C. (1981). *Family Therapy Techniques*. Harvard University Press.
- Monteith, Lwasa, Nsangi, Sseviri, & Byarugaba. (2017). *Upholding the rights of urban refugees in Uganda*.
- Moreno-Garcia, C. F., Jayne, C., Elyan, E., & Aceves-Martins, M. (2023). A novel application of machine learning and zero-shot classification methods for automated abstract screening in systematic reviews. *Decision Analytics Journal*, 6, 100162. <https://doi.org/10.1016/j.dajour.2023.100162>
- Morsbach, S. K., & Prinz, R. J. (2006). Understanding and Improving the Validity of Self-Report of Parenting. *Clinical Child and Family Psychology Review*, 9(1), 1–21. <https://doi.org/10.1007/s10567-006-0001-5>
- Nandy, A., Nixon, E., & Quigley, J. (2021). Observed and reported coparenting and toddlers' adaptive functioning. *Infant and Child Development*, 30(3), e2226. <https://doi.org/10.1002/icd.2226>
- Navigli, R., Conia, S., & Ross, B. (2023). Biases in Large Language Models: Origins, Inventory, and Discussion. *J. Data and Information Quality*, 15(2), 10:1-10:21. <https://doi.org/10.1145/3597307>
- NRC. (2015). *Gaza: The Impact of Conflict on Women*.
- Nugroho, P. A., Anna, N. E. V., & Ismail, N. (2023). The shift in research trends related to artificial intelligence in library repositories during the coronavirus pandemic.

Library Hi Tech, ahead-of-print(ahead-of-print). <https://doi.org/10.1108/LHT-07-2022-0326>

Nye, B., Li, J. J., Patel, R., Yang, Y., Marshall, I., Nenkova, A., & Wallace, B. (2018). A Corpus with Multi-Level Annotations of Patients, Interventions and Outcomes to Support Language Processing for Medical Literature. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 197–207. <https://doi.org/10.18653/v1/P18-1019>

Obsidian (Version Version 1.0.3). (2022). [macOS]. Dynalist Inc.

<https://help.obsidian.md/Plugins/Graph+view>

Ogburn, E. L., Sofrygin, O., Díaz, I., & van der Laan, M. J. (2024). Causal Inference for Social Network Data. *Journal of the American Statistical Association*, 119(545), 597–611. <https://doi.org/10.1080/01621459.2022.2131557>

OpenAI. (2023). *GPT-4 Technical Report* (arXiv:2303.08774). arXiv.

<http://arxiv.org/abs/2303.08774>

Oppenheimer, V. K. (1994). Women's Rising Employment and the Future of the Family in Industrial Societies. *Population and Development Review*, 20(2), 293–342. <https://doi.org/10.2307/2137521>

Orlansky, H. (1949). Infant care and personality. *Psychological Bulletin*, 46(1), 1–48. <https://doi.org/10.1037/h0058106>

Osseiran, S., & Nimer, M. (2024). Knowledge production in refugee studies from the South: Theorization of refugee labour in the literature on Turkey. *Migration Studies*, 12(3), mnae028. <https://doi.org/10.1093/migration/mnae028>

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>
- Panter-Brick, C. (2021). Resilience Humanitarianism and Peacebuilding. In M. Ungar (Ed.), *Multisystemic Resilience* (1st ed., pp. 361-C19.P73). Oxford University Press New York. <https://doi.org/10.1093/oso/9780190095888.003.0020>
- Panter-Brick, C. (2022). Energizing partnerships in research-to-policy projects. *American Anthropologist*, 124(4), 751–766. <https://doi.org/10.1111/aman.13776>
- Panter-Brick, C. (2023). Pathways to resilience and pathways to flourishing: Examining the added-value of multisystem research and intervention in contexts of war and forced displacement. *Development and Psychopathology*, 35(5), 2214–2225.
<https://doi.org/10.1017/S095457942300113X>
- Panter-Brick, C., Burgess, A., Eggerman, M., McAllister, F., Pruett, K., & Leckman, J. F. (2014). Practitioner Review: Engaging fathers – recommendations for a game change in parenting interventions based on a systematic review of the global evidence. *Journal of Child Psychology and Psychiatry*, 55(11), 1187–1212.
<https://doi.org/10.1111/jcpp.12280>
- Panter-Brick, C., Eggerman, M., Ager, A., Hadfield, K., & Dajani, R. (2020). Measuring the psychosocial, biological, and cognitive signatures of profound stress in humanitarian settings: Impacts, challenges, and strategies in the field. *Conflict and Health*, 14(1), 40. <https://doi.org/10.1186/s13031-020-00286-w>
- Panter-Brick, C., Eggerman, M., Gonzalez, V., & Safdar, S. (2009). Violence, suffering, and mental health in Afghanistan: A school-based survey. *The Lancet*, 374(9692), 807–816. [https://doi.org/10.1016/S0140-6736\(09\)61080-1](https://doi.org/10.1016/S0140-6736(09)61080-1)

- Panter-Brick, C., Grimon, M.-P., & Eggerman, M. (2014). Caregiver—Child mental health: A prospective study in conflict and refugee settings. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 55(4), 313–327.
<https://doi.org/10.1111/jcpp.12167>
- Papatzani, E. (Evangelia), Hatziprokopiou, P., Vlastou-Dimopoulou, F., & Siotou, A. (2022). On not staying put where they have put you: Mobilities disrupting the socio-spatial figurations of displacement in Greece. *Journal of Ethnic and Migration Studies*, 0(0), 1–19. <https://doi.org/10.1080/1369183X.2022.2090158>
- Parke, R. D., & Sawin, D. B. (1976). The Father's Role in Infancy: A Re-Evaluation. *The Family Coordinator*, 25(4), 365. <https://doi.org/10.2307/582848>
- Parry, R., Pino, M., Faull, C., & Feathers, L. (2016). Acceptability and design of video-based research on healthcare communication: Evidence and recommendations. *Patient Education and Counseling*, 99(8), 1271–1284.
<https://doi.org/10.1016/j.pec.2016.03.013>
- Patel, A., Cooper, N., Freeman, S., & Sutton, A. (2021). Graphical enhancements to summary receiver operating characteristic plots to facilitate the analysis and reporting of meta-analysis of diagnostic test accuracy data. *Research Synthesis Methods*, 12(1), 34–44. <https://doi.org/10.1002/jrsm.1439>
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. In *Handbook of research methods in personality psychology* (pp. 224–239). The Guilford Press.
- Pearlman, W. (2021). What makes the “refugee crisis” a crisis? Displaced Syrians’ reflections on dignity. *Digest of Middle East Studies*, 30(4), 278–283.
<https://doi.org/10.1111/dome.12249>
- Pedersen, G. A., Smallegange, E., Coetzee, A., Hartog, K., Turner, J., Jordans, M. J. D., & Brown, F. L. (2019a). A Systematic Review of the Evidence for Family and

- Parenting Interventions in Low- and Middle-Income Countries: Child and Youth Mental Health Outcomes. *Journal of Child and Family Studies*, 28(8), 2036–2055. <https://doi.org/10.1007/s10826-019-01399-4>
- Pedersen, G. A., Smallegange, E., Coetzee, A., Hartog, K., Turner, J., Jordans, M. J. D., & Brown, F. L. (2019b). A Systematic Review of the Evidence for Family and Parenting Interventions in Low- and Middle-Income Countries: Child and Youth Mental Health Outcomes. *Journal of Child and Family Studies*, 28(8), 2036–2055. <https://doi.org/10.1007/s10826-019-01399-4>
- Piaget, J. (2007). *The child's conception of the world*. Rowman & Littlefield.
- Pleck, J. H. (2010). Fatherhood and masculinity. In *The role of the father in child development*, 5th ed (pp. 27–57). John Wiley & Sons, Inc.
- Popham, C. M., McEwen, F. S., Karam, E., Fayyad, J., Karam, G., Saab, D., Moghames, P., & Pluess, M. (2022). The dynamic nature of refugee children's resilience: A cohort study of Syrian refugees in Lebanon. *Epidemiology and Psychiatric Sciences*, 31. <https://doi.org/10.1017/S2045796022000191>
- Popham, C. M., McEwen, F. S., Karam, E., Fayyad, J., Karam, G., Saab, D., Moghames, P., & Pluess, M. (2023). Predictors of psychological risk and resilience among Syrian refugee children. *Journal of Child Psychology and Psychiatry*, 64(1), 91–99. <https://doi.org/10.1111/jcpp.13670>
- Postler, K. B., Helms, H. M., & Anastopoulos, A. D. (2022). Examining the linkages between marital quality and anxiety: A meta-analytic review. *Family Process*, n/a(n/a), e12798. <https://doi.org/10.1111/famp.12798>
- Prime, H., Andrews, K., Markwell, A., Gonzalez, A., Janus, M., Tricco, A. C., Bennett, T., & Atkinson, L. (2023). Positive Parenting and Early Childhood Cognition: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Clinical*

Child and Family Psychology Review, 26(2), 362–400.

<https://doi.org/10.1007/s10567-022-00423-2>

Puffer, E. S., Annan, J., Sim, A. L., Salhi, C., & Betancourt, T. S. (2017). The impact of a family skills training intervention among Burmese migrant families in Thailand: A randomized controlled trial. *PLoS ONE*, 12(3).

<https://doi.org/10.1371/journal.pone.0172611>

Putnick, D. L. (2019). Measurement of Parenting. In *Handbook of Parenting* (3rd ed.). Routledge.

Qureshi, R., Han, G., Fapohunda, K., Abariga, S., Wilson, R., & Li, T. (2020). Authorship diversity among systematic reviews in eyes and vision. *Systematic Reviews*, 9, 192. <https://doi.org/10.1186/s13643-020-01451-1>

Rabo, A., Estrada Tun, P., & Jörum, E. (2021). Syrians in Sweden: Constructing Difference Regarding Gender and Family. *Journal of Refugee Studies*, 34(2), 1291–1306. <https://doi.org/10.1093/jrs/feab007>

Reed, R. V., Fazel, M., Jones, L., Panter-Brick, C., & Stein, A. (2012). Mental health of displaced and refugee children resettled in low-income and middle-income countries: Risk and protective factors. *The Lancet*, 379(9812), 250–265. [https://doi.org/10.1016/S0140-6736\(11\)60050-0](https://doi.org/10.1016/S0140-6736(11)60050-0)

Reese, E., & Read, S. (2000). Predictive validity of the New Zealand MacArthur Communicative Development Inventory: Words and Sentences. *Journal of Child Language*, 27(2), 255–266. <https://doi.org/10.1017/S0305000900004098>

Rhee, S., Chang, J., Berthold, S. M., & Mar, G. (2012). Child Maltreatment Among Immigrant Vietnamese Families: Characteristics and Implications for Practice. *Child and Adolescent Social Work Journal*, 29(2), 85–101.

<https://doi.org/10.1007/s10560-011-0253-x>

- Rodrigues, M., Sokolovic, N., Madigan, S., Luo, Y., Silva, V., Misra, S., & Jenkins, J. (2021). Paternal Sensitivity and Children's Cognitive and Socioemotional Outcomes: A Meta-Analytic Review. *Child Development*, 92(2), 554–577.
<https://doi.org/10.1111/cdev.13545>
- Rogers, C. R., Matthews, P., Xu, L., Boucher, K., Riley, C., Huntington, M., Le Duc, N., Okuyemi, K. S., & Foster, M. J. (2020). Interventions for increasing colorectal cancer screening uptake among African-American men: A systematic review and meta-analysis. *PloS One*, 15(9), e0238354.
<https://doi.org/10.1371/journal.pone.0238354>
- Rosenman, R., Tennekoon, V., & Hill, L. G. (2011). Measuring bias in self-reported data. *International Journal of Behavioural & Healthcare Research*, 2(4), 320–332.
<https://doi.org/10.1504/IJBHR.2011.043414>
- Rubenstein, B. L., Lu, L. Z. N., MacFarlane, M., & Stark, L. (2020). Predictors of Interpersonal Violence in the Household in Humanitarian Settings: A Systematic Review. *Trauma, Violence, & Abuse*, 21(1), 31–44.
<https://doi.org/10.1177/1524838017738724>
- Sadiq, A. (2022). Leading Literate Lives: Afghan Refugee Children in a First-Asylum Country. *Journal of Literacy Research*, 54(1), 28–50.
<https://doi.org/10.1177/1086296X221076431>
- Sager, A. A., Goodman, S. H., Jeong, J., Bain, P. A., & Ahun, M. N. (2024). Effects of multi-component parenting and parental mental health interventions on early childhood development and parent outcomes: A systematic review and meta-analysis. *The Lancet Child & Adolescent Health*, 8(9), 656–669.
[https://doi.org/10.1016/S2352-4642\(24\)00134-2](https://doi.org/10.1016/S2352-4642(24)00134-2)

- Sampson, M., Tetzlaff, J., & Urquhart, C. (2011). Precision of healthcare systematic review searches in a cross-sectional sample. *Research Synthesis Methods*, 2(2), 119–125. <https://doi.org/10.1002/jrsm.42>
- Sanders, M. R., Kirby, J. N., Tellegen, C. L., & Day, J. J. (2014). The Triple P-Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychology Review*, 34(4), 337–357. <https://doi.org/10.1016/j.cpr.2014.04.003>
- Sangalang, C. C., & Vang, C. (2017a). Intergenerational Trauma in Refugee Families: A Systematic Review. *Journal of Immigrant and Minority Health*, 19(3), 745–754. <https://doi.org/10.1007/s10903-016-0499-7>
- Sangalang, C. C., & Vang, C. (2017b). Intergenerational Trauma in Refugee Families: A Systematic Review. *Journal of Immigrant and Minority Health*, 19(3), 745–754. <https://doi.org/10.1007/s10903-016-0499-7>
- Santos, Á. O. dos, da Silva, E. S., Couto, L. M., Reis, G. V. L., & Belo, V. S. (2023). The use of artificial intelligence for automating or semi-automating biomedical literature analyses: A scoping review. *Journal of Biomedical Informatics*, 142, 104389. <https://doi.org/10.1016/j.jbi.2023.104389>
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, 52(4), 1893–1907. <https://doi.org/10.1007/s11135-017-0574-8>
- Saxena, S., Paraje, G., Sharan, P., Karam, G., & Sadana, R. (2006). The 10/90 divide in mental health research: Trends over a 10-year period. *The British Journal of Psychiatry*, 188(1), 81–82. <https://doi.org/10.1192/bjp.bp.105.011221>

- Schaefer, E. S. (1959). A circumplex model for maternal behavior. *The Journal of Abnormal and Social Psychology*, 59(2), 226–235.
<https://doi.org/10.1037/h0041114>
- Scharpf, F., Kaltenbach, E., Nickerson, A., & Hecker, T. (2021). A systematic review of socio-ecological factors contributing to risk and protection of the mental health of refugee children and adolescents. *Clinical Psychology Review*, 83, 101930.
<https://doi.org/10.1016/j.cpr.2020.101930>
- Scharpf, F., Mkinga, G., Neuner, F., Machumu, M., & Hecker, T. (2021). Fuel to the fire: The escalating interplay of attachment and maltreatment in the transgenerational transmission of psychopathology in families living in refugee camps. *Development and Psychopathology*, 33(4), 1308–1321.
<https://doi.org/10.1017/S0954579420000516>
- Scharpf, F., Paulus, M., Christner, N., Beerbaum, L., Kammermeier, M., & Hecker, T. (2023). Intergenerational transmission of mental health risk in refugee families: The role of maternal psychopathology and emotional availability. *Development and Psychopathology*, 1–14. <https://doi.org/10.1017/S0954579423000846>
- Schoppe-Sullivan, S. J., Kotila, L., Jia, R., Lang, S. N., & Bower, D. J. (2013). Comparisons of levels and predictors of mothers' and fathers' engagement with their preschool aged children. *Early Child Development and Care*, 183(3–4), 498–514. <https://doi.org/10.1080/03004430.2012.711596>
- Scott, J. K. (2020). *Spanking attitudes and parents' biased information-seeking, distrust in research, and overestimations of scientific knowledge*.
<https://hdl.handle.net/2152/88564>
- Seagle, E. E., Dam, A. J., Shah, P. P., Webster, J. L., Barrett, D. H., Ortmann, L. W., Cohen, N. J., & Marano, N. N. (2020). Research ethics and refugee health: A

- review of reported considerations and applications in published refugee health literature, 2015-2018. *Conflict and Health*, 14(1), 39.
<https://doi.org/10.1186/s13031-020-00283-z>
- Sear, R. (2016). Beyond the nuclear family: An evolutionary perspective on parenting. *Current Opinion in Psychology*, 7, 98–103.
<https://doi.org/10.1016/j.copsyc.2015.08.013>
- Shawaf, N., & El Asmar, F. (2017). *'We're Not There Yet...' Voices of refugees from Syria in Lebanon*. Oxfam. <https://doi.org/10.21201/2017.9620>
- Shergold, P., Benson, K., & Piper, M. (2019). *Investing in refugees, investing in Australia the findings of a review into integration, employment and settlement outcomes for refugees and humanitarian entrants in Australia*. Dept. of the Prime Minister and Cabinet.
- Shreffler, J., & Huecker, M. R. (2023). Diagnostic Testing Accuracy: Sensitivity, Specificity, Predictive Values and Likelihood Ratios. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK557491/>
- Sim, A., Bowes, L., & Gardner, F. (2018). Modeling the effects of war exposure and daily stressors on maternal mental health, parenting, and child psychosocial adjustment: A cross-sectional study with Syrian refugees in Lebanon. *Global Mental Health*, 5, e40. <https://doi.org/10.1017/gmh.2018.33>
- Sim, A., Bowes, L., & Gardner, F. (2019). The Promotive Effects of Social Support for Parental Resilience in a Refugee Context: A Cross-Sectional Study with Syrian Mothers in Lebanon. *Prevention Science*, 20(5), 674–683.
<https://doi.org/10.1007/s11121-019-0983-0>
- Sim, A., Fazel, M., Bowes, L., & Gardner, F. (2018). Pathways linking war and displacement to parenting and child adjustment: A qualitative study with Syrian

refugees in Lebanon. *Social Science and Medicine*, 200, 19–26.

<https://doi.org/10.1016/j.socscimed.2018.01.009>

Sim, A. L., Bowes, L., Maignant, S., Magber, S., & Gardner, F. (2021). Acceptability and Preliminary Outcomes of a Parenting Intervention for Syrian Refugees. *Research on Social Work Practice*, 31(1), 14–25.

<https://doi.org/10.1177/1049731520953627>

Singh, A., Nemiro, A., Malik, A., Guimond, M.-F., Nduwimana, E., Likindikoki, S., Annan, J., & Tol, W. A. (2021). Cultural adaptation of a scalable psychological intervention for Burundian refugee adolescents in Tanzania: A qualitative study. *Conflict and Health*, 15(1), 73. <https://doi.org/10.1186/s13031-021-00391-4>

Skuban, E. M., Shaw, D. S., Gardner, F., Supplee, L. H., & Nichols, S. R. (2006). The correlates of dyadic synchrony in high-risk, low-income toddler boys. *Infant Behavior and Development*, 29(3), 423–434.

<https://doi.org/10.1016/j.infbeh.2006.02.004>

Smetana, J. G., & Ahmad, I. (2018). Heterogeneity in Perceptions of Parenting Among Arab Refugee Adolescents in Jordan. *Child Development*, 89(5), 1786–1802. <https://doi.org/10.1111/cdev.12844>

Sousa, C. A., Akesson, B., & Siddiqi, M. (2025). Parental Resilience in Contexts of Political Violence: A Systematic Scoping Review of 45 Years of Research. *Trauma, Violence, & Abuse*, 26(1), 41–57.

<https://doi.org/10.1177/15248380241270048>

Stern, C., Lizarondo, L., Carrier, J., Godfrey, C., Rieger, K., Salmond, S., Apóstolo, J., Kirkpatrick, P., & Loveday, H. (2020). Methodological guidance for the conduct of mixed methods systematic reviews. *JBIS Evidence Synthesis*, 18(10), 2108–2118. <https://doi.org/10.11124/JBISRIR-D-19-00169>

- Stewart, P. E. (2015). You Moved Up, Did You Forget Us?: The Influence of African American Intra-familial Social Mobility on Extended Family Relationships. *Journal of African American Studies*, 19(2), 214–232.
<https://doi.org/10.1007/s12111-015-9300-6>
- Sulaiman-Hill, C. M., & Thompson, S. C. (2011). Sampling challenges in a study examining refugee resettlement. *BMC International Health and Human Rights*, 11, 2. <https://doi.org/10.1186/1472-698X-11-2>
- Summerscales, R. L., Argamon, S., Bai, S., Hupert, J., & Schwartz, A. (2011). Automatic Summarization of Results from Clinical Trials. *2011 IEEE International Conference on Bioinformatics and Biomedicine*, 372–377.
<https://doi.org/10.1109/BIBM.2011.72>
- Syam, H., Venables, E., Sousse, B., Severy, N., Saavedra, L., & Kazour, F. (2019). “With every passing day I feel like a candle, melting little by little.” experiences of long-term displacement amongst Syrian refugees in Shatila, Lebanon. *Conflict and Health*, 13, 45. <https://doi.org/10.1186/s13031-019-0228-7>
- Syriani, E., David, I., & Kumar, G. (2023). *Assessing the Ability of ChatGPT to Screen Articles for Systematic Reviews* (arXiv:2307.06464). arXiv.
<https://doi.org/10.48550/arXiv.2307.06464>
- Talebi, N. (2024). Talking Past Each Other: Politics of Knowledge Production in Transnational Power Relations. *Nordic Journal of Migration Research*, 14(4).
<https://doi.org/10.33134/njmr.776>
- Tatton, A. (2023). Using the Theory of Family Boundary Ambiguity and Ambiguous Loss to Understand the Experiences of Foster Carers’ Own Children. *The British Journal of Social Work*, 53(1), 198–215. <https://doi.org/10.1093/bjsw/bcac139>

- Terhürne, P., Schwartz, B., Baur, T., Schiller, D., Eberhardt, S. T., André, E., & Lutz, W. (2022). Validation and application of the Non-Verbal Behavior Analyzer: An automated tool to assess non-verbal emotional expressions in psychotherapy. *Frontiers in Psychiatry, 13*. <https://doi.org/10.3389/fpsyt.2022.1026015>
- Thabet, A. A., Ibraheem, A. N., Shivram, R., Winter, E. A., & Vostanis, P. (2009). Parenting support and PTSD in children of a war zone. *The International Journal of Social Psychiatry, 55*(3), 226–237. <https://doi.org/10.1177/0020764008096100>
- Thompson, R. A. (2021). Eyes to see and ears to hear: Sensitivity in research on attachment and culture. *Attachment & Human Development, 23*(2), 222–230. <https://doi.org/10.1080/14616734.2020.1828548>
- Timshel, I., Montgomery, E., & Dalgaard, N. T. (2017). A systematic review of risk and protective factors associated with family related violence in refugee families. *Child Abuse & Neglect, 70*, 315–330. <https://doi.org/10.1016/j.chiabu.2017.06.023>
- Tobin, S. A., Momani, F., & Al Yakoub, T. (2022). ‘The war has divided us more than ever’: Syrian refugee family networks and social capital for mobility through protracted displacement in Jordan. *Journal of Ethnic and Migration Studies, 0*(0), 1–18. <https://doi.org/10.1080/1369183X.2022.2090157>
- Tol, W. A., Barbui, C., Galappatti, A., Silove, D., Betancourt, T. S., Souza, R., Golaz, A., & Ommeren, M. van. (2011). Mental health and psychosocial support in humanitarian settings: Linking practice and research. *The Lancet, 378*(9802), 1581–1591. [https://doi.org/10.1016/S0140-6736\(11\)61094-5](https://doi.org/10.1016/S0140-6736(11)61094-5)
- Tol, W. A., Le, P. D., Harrison, S. L., Galappatti, A., Annan, J., Baingana, F. K., Betancourt, T. S., Bizouerne, C., Eaton, J., Engels, M., Hijazi, Z., Horn, R. R., Jordans, M. J. D., Kohrt, B. A., Koyiet, P., Panter-Brick, C., Pluess, M., Rahman, A., Silove, D., ... Van Ommeren, M. (2023). Mental health and psychosocial

- support in humanitarian settings: Research priorities for 2021–30. *The Lancet Global Health*, S2214109X23001286. [https://doi.org/10.1016/S2214-109X\(23\)00128-6](https://doi.org/10.1016/S2214-109X(23)00128-6)
- Tol, W. A., Song, S., & Jordans, M. J. D. (2013). Annual Research Review: Resilience and mental health in children and adolescents living in areas of armed conflict – a systematic review of findings in low- and middle-income countries. *Journal of Child Psychology and Psychiatry*, 54(4), 445–460. <https://doi.org/10.1111/jcpp.12053>
- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5), 859–883. <https://doi.org/10.1037/0033-2909.133.5.859>
- Twum-Danso Imoh, A., Rabello de Castro, L., & Naftali, O. (2022). Studies of childhoods in the Global South: Towards an epistemic turn in transnational childhood research? *Third World Thematics: A TWQ Journal*, 7(1–3), 1–16. <https://doi.org/10.1080/23802014.2022.2161619>
- Ungar, M. (2009). Overprotective Parenting: Helping Parents Provide Children the Right Amount of Risk and Responsibility. *The American Journal of Family Therapy*, 37(3), 258–271. <https://doi.org/10.1080/01926180802534247>
- Ungar, M. (2010). Families as Navigators and Negotiators: Facilitating Culturally and Contextually Specific Expressions of Resilience. *Family Process*, 49(3), 421–435. <https://doi.org/10.1111/j.1545-5300.2010.01331.x>
- Ungar, M., & Hadfield, K. (2019). The differential impact of environment and resilience on youth outcomes. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, 51(2), 135–146. <https://doi.org/10.1037/cbs0000128>

- Ungar, M., Theron, L., & Hölzge, J. (2023). Multisystemic approaches to researching young people's resilience: Discovering culturally and contextually sensitive accounts of thriving under adversity. *Development and Psychopathology*, 35(5), 2199–2213. <https://doi.org/10.1017/S0954579423000469>
- UNHCR. (1951). *The 1951 Refugee Convention*. <https://www.unhcr.org/3b66c2aa10>
- UNHCR. (2004, June 10). *Protracted Refugee Situations*.
<https://www.unhcr.org/excom/standcom/40c982172/protracted-refugee-situations.html>
- UNHCR. (2007). *Sources, methods and data considerations*.
<https://www.unhcr.org/sites/default/files/legacy-pdf/4981c2f02.pdf>
- UNHCR. (2012). *Dadaab—World's biggest refugee camp 20 years old*. UNHCR.
<https://www.unhcr.org/news/stories/dadaab-worlds-biggest-refugee-camp-20-years-old>
- UNHCR. (2018). *Shelter and infrastructure*. Global Focus.
<https://reporting.unhcr.org/shelter-and-infrastructure-99>
- UNHCR. (2019). *Handbook on Procedures and Criteria for Determining Refugee Status and Guidelines on International Protection Under the 1951 Convention and the 1967 Protocol Relating to the Status of Refugees*.
<https://www.refworld.org/type,HANDBOOK,UNHCR,,5cb474b27,0.html>
- UNHCR. (2020). *Protracted Refugee Situations Explained*.
<https://www.unrefugees.org/news/protracted-refugee-situations-explained/>
- UNHCR. (2021a). *Rwanda Country Refugee Response Plan (CRP)*.
<https://reporting.unhcr.org/sites/default/files/2021%20Rwanda%20Country%20Refugee%20Response%20Plan.pdf>

UNHCR. (2021b). *Situation Syria Regional Refugee Response* [Dataset].

<https://data.unhcr.org/en/situations/syria/location/36>

UNHCR. (2021c). *UNHCR Global Trends Report 2021*. United Nations High

Commissioner for Refugees. <https://data.unhcr.org/en/documents/details/93791>

UNHCR. (2021d, November 10). *Refugee Data Finder*. UNHCR.

<https://www.unhcr.org/refugee-statistics/>

UNHCR. (2022). *Refugee host countries by income level*. UNHCR Refugee Statistics.

<https://www.unhcr.org/refugee-statistics/insights/explainers/refugee-host-countries-income-level.html>

UNHCR. (2023a). *Somalia Refugee Crisis Explained*.

<https://www.unrefugees.org/news/somalia-refugee-crisis-explained/>

UNHCR. (2023b). *Uganda*. Global Focus.

<https://reporting.unhcr.org/operational/operations/uganda>

UNHCR. (2023c). *UNHCR - Refugee Statistics*. UNHCR. <https://www.unhcr.org/refugee-statistics/>

UNHCR. (2023d). *Unlocking Potential: The Right to Education and Opportunity*.

UNHCR. (2024a). *Data finder* [Population Statistics; CSV]. UNHCR Refugee Statistics

Database. [https://www.unhcr.org/refugee-statistics/download?data_finder%5BdataGroup%5D=displacement&data_finder%5Bdataset%5D=population&data_finder%5BdisplayType%5D=totals&data_finder%5BpopulationType%5D%5B%5D=REF&data_finder%5BpopulationType%5D%5B%5D=ASY&data_finder%5BpopulationType%5D%5B%5D=IDP&data_finder%5BpopulationType%5D%5B%5D=OIP&data_finder=on&data_finder%5Byear__filterType%5D=range&data_finder%5Byear__rangeFrom%5D=2022&data_finder%5Byear__rangeTo%5D=2024&data_finder%](https://www.unhcr.org/refugee-statistics/download?data_finder%5BdataGroup%5D=displacement&data_finder%5Bdataset%5D=population&data_finder%5BdisplayType%5D=totals&data_finder%5BpopulationType%5D%5B%5D=REF&data_finder%5BpopulationType%5D%5B%5D=ASY&data_finder%5BpopulationType%5D%5B%5D=IDP&data_finder%5BpopulationType%5D%5B%5D=OIP&data_finder=on&data_finder%5Byear__filterType%5D=range&data_finder%5Byear__rangeFrom%5D=2022&data_finder%5Byear__rangeTo%5D=2024&data_finder%5B)

5Bcoo__displayType%5D=all&data_finder%5Bcoa__displayType%5D=custom&
 data_finder%5Bcoa__country%5D%5B%5D=96&data_finder%5Bcoa__country
 %5D%5B%5D=96&data_finder%5Byear__%5D=&data_finder%5Bcoo__%5D=
 &data_finder%5Bcoa__%5D=&data_finder%5Badvanced__%5D=&data_finder
 %5Bsubmit%5D=

UNHCR. (2024b). *Global Appeal*. Global Focus. <https://reporting.unhcr.org/global-appeal-2024>

UNHCR. (2024c). *Global trends*. <https://www.unhcr.org/global-trends-report-2023>

UNHCR. (2024d). *Lebanon*. Global Focus.
<https://reporting.unhcr.org/operational/operations/lebanon>

UNHCR. (2025). *Global appeal*.

UNICEF. (2023, June 13). *Number of displaced children reaches new high of 43.3 million*. <https://www.unicef.org/press-releases/number-displaced-children-reaches-new-high-433-million>

Vaage, A. B., Thomsen, P. H., Rousseau, C., Wentzel-Larsen, T., Ta, T. V., & Hauff, E. (2011). Paternal predictors of the mental health of children of Vietnamese refugees. *Child and Adolescent Psychiatry and Mental Health*, 5.
<https://doi.org/10.1186/1753-2000-5-2>

Van de Mortel, T. (2008). Faking it: Social desirability response bias in self-report research. *Thea van de Mortel*, 25.

van Ee, E., Kleber, R. J., & Jongmans, M. J. (2016). Relational Patterns Between Caregivers With PTSD and Their Nonexposed Children: A Review. *Trauma, Violence, & Abuse*, 17(2), 186–203. <https://doi.org/10.1177/1524838015584355>

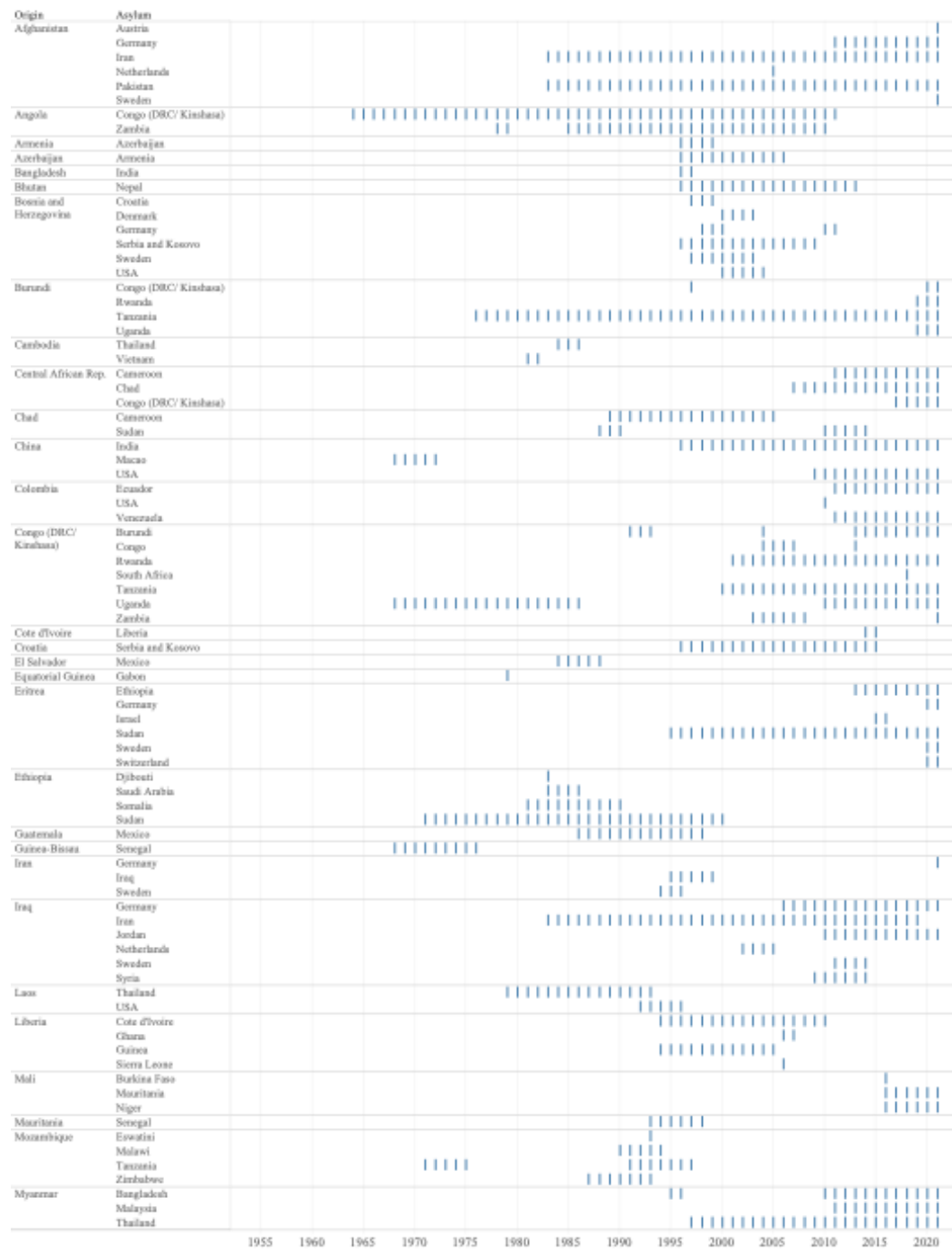
- van Ee, E., Sleijpen, M., Kleber, R. J., & Jongmans, M. J. (2013). Father-Involvement in a Refugee Sample: Relations between Posttraumatic Stress and Caregiving. *Family Process*, 52(4), 723–735. <https://doi.org/10.1111/famp.12045>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. <https://doi.org/10.48550/ARXIV.1706.03762>
- Vaughn, B., Deinard, A., & Egeland, B. (1980). Measuring temperament in pediatric practice. *The Journal of Pediatrics*, 96(3 Pt 1), 510–514. [https://doi.org/10.1016/s0022-3476\(80\)80712-8](https://doi.org/10.1016/s0022-3476(80)80712-8)
- Wallace, B. C., Kuiper, J., & Sharma, A. (2016). *Extracting PICO Sentences from Clinical Trial Reports using Supervised Distant Supervision*.
- Wals, A. E. J. (2010). Mirroring, Gestaltswitching and transformative social learning: Stepping stones for developing sustainability competence. *International Journal of Sustainability in Higher Education*, 11(4), 380–390. <https://doi.org/10.1108/14676371011077595>
- Walsh, F. (2006). *Strengthening family resilience, 2nd ed* (pp. xvi, 384). The Guilford Press.
- Walsh, F. (2012). Chapter 8 Successful Aging and Family Resilience. *Annual Review of Gerontology and Geriatrics*, 32(1), 151–172. <https://doi.org/10.1891/0198-8794.32.153>
- Walsh, F. (2021). Family Resilience: A Dynamic Systemic Framework. In M. Ungar (Ed.), *Multisystemic Resilience: Adaptation and Transformation in Contexts of Change* (p. 0). Oxford University Press. <https://doi.org/10.1093/oso/9780190095888.003.0015>

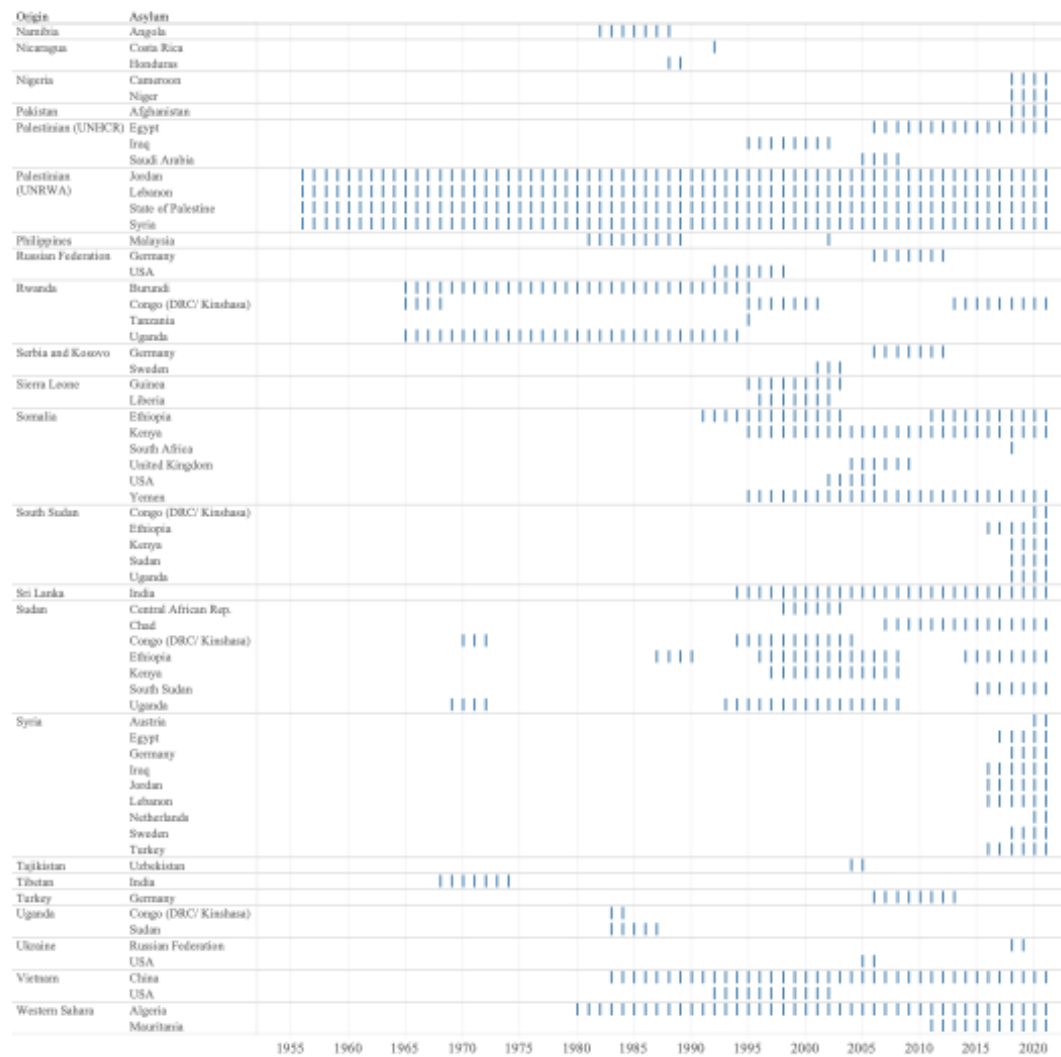
- Wang, S., Scells, H., Koopman, B., & Zuccon, G. (2023). *Can ChatGPT Write a Good Boolean Query for Systematic Review Literature Search?* (arXiv:2302.03495). arXiv. <https://doi.org/10.48550/arXiv.2302.03495>
- Wang, Z., Nayfeh, T., Tetzlaff, J., O'Brien, P., & Murad, M. H. (2020). Error rates of human reviewers during abstract screening in systematic reviews. *PLoS ONE*, 15(1), e0227742. <https://doi.org/10.1371/journal.pone.0227742>
- WHO. (2014). *Mental health*. <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response>
- WHO. (2020). *Nurturing care for children living in humanitarian settings*. <https://iris.who.int/bitstream/handle/10665/337762/9789240016132-eng.pdf>
- Wiepkema, F. (2013). *Quality of father-child interaction in refugee and asylum seeker families: The relation among posttraumatic stress and parental caregiving* [Master Thesis]. <https://studenttheses.uu.nl/handle/20.500.12932/12408>
- Willen, S. S. (2019). *Fighting for dignity: Migrant lives at Israel's margins* (1st edition). PENN University of Pennsylvania Press.
- Wilson, S., & Durbin, C. E. (2010). Effects of paternal depression on fathers' parenting behaviors: A meta-analytic review. *Clinical Psychology Review*, 30(2), 167–180. <https://doi.org/10.1016/j.cpr.2009.10.007>
- Wissö, T., & Bäck-Wiklund, M. (2021a). Fathering Practices in Sweden During the COVID-19: Experiences of Syrian Refugee Fathers. *Frontiers in Sociology*, 6, 721881. <https://doi.org/10.3389/fsoc.2021.721881>
- Wissö, T., & Bäck-Wiklund, M. (2021b). Fathering Practices in Sweden During the COVID-19: Experiences of Syrian Refugee Fathers. *Frontiers in Sociology*, 6. <https://www.frontiersin.org/articles/10.3389/fsoc.2021.721881>

- Yaffe, Y. (2023). Systematic review of the differences between mothers and fathers in parenting styles and practices. *Current Psychology*, 42(19), 16011–16024.
<https://doi.org/10.1007/s12144-020-01014-6>
- Yalim, A. C., & Critelli, F. (2023). Gender roles among Syrian refugees in resettlement contexts: Revisiting empowerment of refugee women and needs of refugee men. *Women's Studies International Forum*, 96, 102670.
<https://doi.org/10.1016/j.wsif.2022.102670>
- Yang, B. (2023). Chapter 11—Vibration and control of multiple-degree-of-freedom systems. In B. Yang (Ed.), *Stress, Strain, and Structural Dynamics (Second Edition)* (pp. 529–625). Academic Press. <https://doi.org/10.1016/B978-0-12-818563-6.00001-8>
- Yashika, Y., & Prakash, D. (2025). Identity Threat: A Literature Review and Future Research Agenda Using Theories Context Methodology & Antecedents Decisions Outcomes (TCM & ADO) Framework. *Basic and Applied Social Psychology*, 47(2), 57–70. <https://doi.org/10.1080/01973533.2024.2418857>
- Yoosefi Lebni, J., Enayat, H., Irandoost, S. F., & Dehghan, A. A. (2022). Exploring the Challenges of Afghan Refugee Women Facing COVID-19: A Qualitative Study in Iran. *Frontiers in Public Health*, 10, 838965.
<https://doi.org/10.3389/fpubh.2022.838965>
- Young, A. (1995). *The Harmony of Illusions: Inventing Post-Traumatic Stress Disorder*. Princeton University Press. <https://www.jstor.org/stable/j.ctt7swhj>

Supplements

Supplementary Table 1: List of protracted refugee situations

Protracted Refugee Situations: Place and Time



Note. The blue bars indicate the years in which a protracted refugee situation was ongoing in a given country. The cut-off point for a protracted refugee situation was 24,000 refugees or more from a given country of origin remaining for five years or longer in a given asylum country. The data was extracted from the United Nations High Commissioner for Refugees (UNHCR) on the 28th of Feb of July 2023 using the Data Finder tool (last updated on the 16th of June 2022). For Palestinian refugees under the United Nations Relief and Works Agency (UNRWA) mandate, the data was extracted on the same platform, but the data source was UNRWA. The data's licence is under Creative Commons, which allows for sharing (copying and redistributing the material in any medium or format Adapt) and remixing, transforming, and building upon the material for any purpose, even commercially.

Supplementary Table 2: Quality appraisal raw results

	Total score	Theoretical foundations	Research aim	Setting & population	Study design	Sampling approach	Data collection rationale	Tool format/content	Collection procedure	Recruitment data	Analytic justification	Analysis appropriateness	Stakeholder consideration	Strengths/limitations
Adfam (2009)	28	3	3	3	2	2	3	2	2	1	0	3	3	1
Afifi et al. (2016)	33	3	3	3	3	3	2	3	3	3	3	2	1	1
Akesson & Sousa (2020)	27	3	3	3	2	3	0	3	2	3	0	3	0	2
Akhter & Kuskabe (2014)	21	3	2	3	3	3	0	0	2	2	0	3	0	0
Al-Natour et al. (2022)	27	2	2	2	3	2	3	3	2	2	2	3	0	1
Alnuaimi (2018)	29	2	3	3	3	3	2	3	3	2	1	3	0	1
Atallah (2017)	37	3	3	3	3	3	2	3	3	3	3	3	2	3
Babuc (2022)	29	3	3	3	2	2	3	3	3	3	0	3	0	1

Bella my et al. (201 7)	24	1	3	3	3	2	3	2	3	1	0	3	0	0
Bunn (202 2)	38	3	3	3	3	3	3	3	3	2	3	3	3	3
Dokt orlari (201 9)	18	0	2	3	2	2	0	2	3	2	0	1	0	1
Fell meth et al. (201 8)	29	3	3	3	3	2	1	3	3	3	2	3	0	0
Heck er et al. (202 2)	31	3	3	2	3	2	2	3	3	3	1	3	1	2
Inga bire & Richt ers (202 0)	24	3	3	3	3	2	0	3	2	2	0	3	0	0
Jabb ar & Zaza (201 9)	30	3	3	3	3	3	1	3	2	3	1	3	1	1
Lem bcke et al. (202 0)	32	3	2	2	2	1	3	3	3	3	3	3	1	3
Lewa ndo Hund t et al. (200 4)	18	3	1	3	2	2	1	1	2	1	0	1	0	1
Lewe k & Nabe r	18	3	3	1	2	1	1	2	1	0	1	1	0	2

(2017)															
Meyer et al. (2013)	30	3	3	3	1	2	3	2	3	2	2	3	0	3	
Meyer et al. (2017)	37	3	3	3	3	3	2	3	3	3	3	3	2	3	
Miller et al. (2020)	35	3	3	2	3	3	3	3	3	3	0	3	3	3	
NRC (2015)	19	0	3	3	3	2	1	0	2	3	0	1	0	1	
Puffer et al. (2017)	39	3	3	3	3	3	3	3	3	3	3	3	3	3	
Rhee et al. (2012)	34	3	3	3	3	3	3	3	3	3	3	3	0	1	
Scharpf et al. (2021)	35	3	3	2	2	2	3	3	3	3	3	3	2	3	
Sim et al., (2019)	33	3	3	3	3	2	3	2	3	1	3	3	1	3	
Sim, Bowes, et al. (2018b)	32	3	3	3	2	1	3	3	3	1	3	3	1	3	
Sim, Fazel, et al.	29	3	3	3	2	2	2	2	3	1	2	3	1	2	

(2018)														
Syam et al. (2019)	27	2	3	3	2	2	0	3	3	3	0	3	1	2
Yoos ef Lebn i et al. (2022)	30	3	3	3	3	3	0	3	3	2	0	3	2	2

Note. This table details the quality of the studies. Overall, almost all the peer-reviewed literature demonstrated a strong degree of theoretical or conceptual grounding, while a few grey literatures work lacked such a basis. Research aims are generally well-defined, but there is room for improvement with greater specificity, particularly in studies that included mixed methods or explored complex interactions between mental health and parenting in the environment. Detailed descriptions of research settings and target populations were often present, but sometimes, studies that were otherwise rigorous needed more specificity. While most studies employed designs appropriate for their aims, some inconsistencies exist, highlighting the need for a more rigorous selection of methodologies aligned with specific research questions. Sampling strategies vary in rigour; some studies demonstrate careful consideration of sampling, while others could benefit from more robust approaches to ensure alignment with the respective goals of the study. Low scores were common when it came to the rationale behind data collection tools, which was often explained poorly or not at all; in particular, many studies did not explain how the interview type related to their research question, the population, or the overall methodological approach, and why a combination of interviewing approaches was used or not. Often, the format and content of data collection tools (e.g., type of questions asked) seem suited to address research aims, yet a few studies showed content that was disjointed from the specific research questions. Data collection procedures are usually described, with one exception, with some other studies benefitting from providing more comprehensive accounts to enhance replicability. Recruitment data is typically provided; however, some studies had incomplete data that did not allow for a full picture of the recruitment process and outcome. Analytic methods were, in many cases, not justified at all, while, in other cases, deeper explanations linking the chosen method to the study's specific research questions would have been beneficial. In terms of the appropriateness of analytic methods, most studies demonstrated a strong fit, but others could be improved by selecting methods more closely aligned with their research aims. Very few studies showed evidence of the integration of refugee perspectives in research design, indicating a critical gap in collaborative and participatory

approaches. A majority of studies did not offer a thorough discussion of strengths and limitations, with only a handful demonstrating critical depth.

Supplementary Table 3: Sociodemographic characteristics of the qualitative study

Father's name	Mother's name	Father's age	Mother's age	Father's education	Mother's education	Father's work status	Area of residence
Faris	Fatma	33	38	Primary	Primary	Freelance	Amman
Fawaz	Fayrooz	40	37	Primary	Unschoolled	No job	Amman
Hosam	Hend	33	43	Secondary	Primary	Metal worker	Amman
Majed	Marwa	24	35	Primary	Primary	Furniture painter	Amman
Mohammed	Marah	25	29	Primary	Secondary	Freelance	Amman
Rami	Ruba	34	36	Primary	Unschoolled	No job	Amman
Rashdan	Rawan	36	40	Primary	Secondary	Furniture painter	Amman
Saeed	Sahar	27	37	Primary	Primary	Builder	Amman
Zaid	Zaytona	35	42	Primary	Primary	Builder	Amman
Dawod	Dalal	24	36	Secondary	Secondary	No job	Zaatari
Ibrahim	Ibtisam	27	33	Secondary	Secondary	Freelance	Zaatari
Jabir	Joud	24	34	University	Secondary	Teacher	Zaatari
Kareem	Kisma	43	45	Secondary	Primary	School guard	Zaatari
Luay	Layan	28	36	Secondary	Secondary	School guard	Zaatari
Sami	Sarah	40	47	Secondary	Primary	Builder	Zaatari

Participant names shown in this table are pseudonyms.

Supplementary Table 4: Demographics of sample agreeing

Group	N	Child age (avg)	Child gender	Child in School	Mother works	Father lives in house	Urban vs camp	High density (2.51+ persons per room)
Agreed to all recording	244	6.28	48% female	65% in school	11.5%	91% Father present	73% Amman	41.9%
Agreed to recording at the initial phase but missed some/all follow-up	26	6.63	58% female	81% in school	8%	88.5% Father present	81% Amman	33.3%
Refused some/all recording at the initial phase but agreed to be recorded at follow-up	16	6.28	40% female	75% in school	25%	87.5% Father present	44% Amman	53.3%

Refused all	32	6.40	56% female	69% in school	3%	75% Father present	84% Amma n	51.7%
-------------	----	------	---------------	------------------	----	--------------------------	------------------	-------

The “refused all” group shows the highest housing density, lowest father presence (75%), and slowest maternal employment (3%). We considered the upper half of answers as “high density” scores.

Supplementary Table 5: OBSESS and self-reported measures with low reliability

a) Parenting self-reports’ correlations with OBSES at the initial data collection stage

	1	2	3	4
1. Sensitivity	—	.41**	.03	−.06
2. Synchrony		—	.05	−.06
3. Mother–Child Closeness			—	−.20**
4. Mother–Child Conflict				—

b) Parenting self-reports’ correlations with OBSES at follow-up

	1	2	3	4	5	6	7
1. Sensitivity	—	.44**	−.15*	0.06	.14*	−.10	0.01
2. Synchrony		—	−.10	0.07	0.09	−.08	−.06
3. Mother–Child Conflict			—	−.26**	−.30**	0.08	.36**

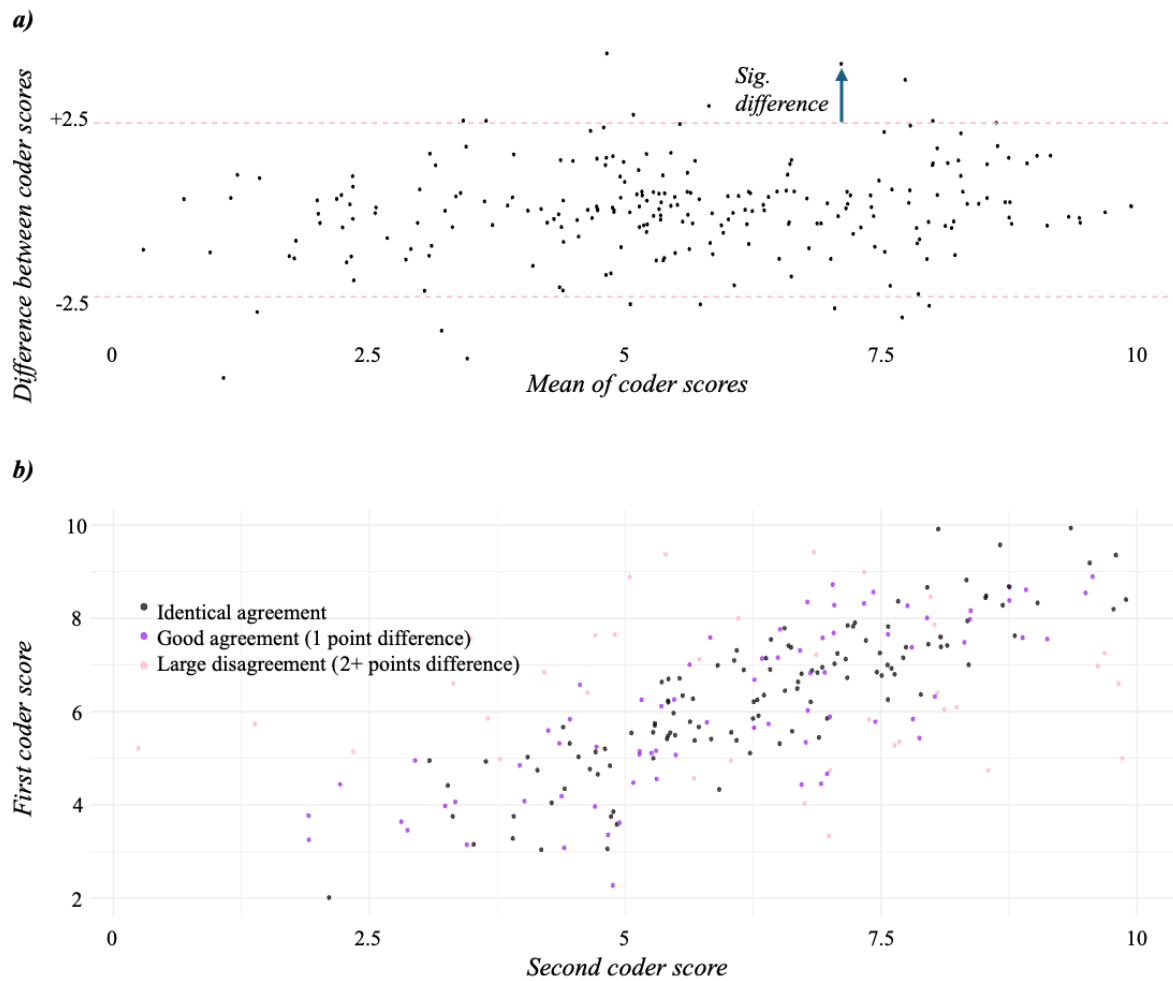
4. Mother–Child Closeness	—	.35**	0.02	–.21**
5. Positive Parenting		—	0.07	–.15*
6. Inconsistent Discipline			—	0.06
7. Poor Supervision				—

Supplementary Table 6: Detail of the iterative process to engineer the prompt across the different stages

<i>Title and abstract</i>	We tested GPT-4's ability to screen citations, where we faced challenges with the model generating all the responses for multiple citations. Segregating the citations from the prompts allowed GPT-4's to respond to all citations. We checked how GPT-4 was assessing decisions, which revealed that the model often misjudged studies when keywords were missing in citations but present in full abstracts. This suggested that GPT-4 did not reliably refer to the abstract when testing, but rather to the title. We refined the prompts to include the full abstracts and monitored the correctness of GPT-4's decision when keywords were not mentioned in the citations. GPT-4 was unable to produce a final decision when faced with complex queries and large textual data volumes. To mitigate this, we moved from including all the inclusion/exclusion criteria at once to presenting each criterion separately and reduced the number of titles/abstracts in each message.
---------------------------	--

<i>Full text</i>	Initially, we slightly modified the prompt from the title/abstract phase, mainly altering decision-making language to better suit the full-text stage requirements, where we replaced ‘Yes/maybe’ with ‘Include’ and ‘No’ with ‘Exclude.’ For consistency, we continued restricting GPT-4’s output to the final decision. Next, although we retained the core instructions from the title/abstract prompt, due to the token limitations of GPT-4, we segmented the study into smaller parts. To ensure consistency, we ensured that each text segment sent to GPT-4 for screening included the relevant inclusion/exclusion criteria within the same prompt.
<i>Extraction</i>	Prompts first included generic instructions like creating a table for the participants in the study. This initial set of prompts underwent team review, leading to a refinement process aimed at providing more detailed and explicit instructions. For example, the table prompt was rewritten to explicitly specify the total number of participants, the average age for the entire sample, and a breakdown by female/male/non-binary categories. A total of five studies were used to pilot these specific prompts.

Supplementary figure 1: Normal Range (a) and Extent of Differences in Coding (b)



Supplementary Figure 1a. shows that there are no systematic differences in coding score between two coders on the same video, whereas figure 1b. shows that most disagreement in coding score between two coders on the same video are less than 2-points difference.