**Parental Mentalization and Children’s Externalizing Problems: A Systematic Review and Meta-Analysis**

[Supplemental Online Material]

Maitane Nieto-Retuerto, Barbara Torres-Gomez, and Itziar Alonso-Arbiol

Department of Clinical and Health Psychology & Research Methods, University of the Basque Country UPV/EHU

Authors Note

Maitane Nieto-Retuerto  https://orcid.org/0000-0002-8474-9988

Barbara Torres-Gomez  <https://orcid.org/0000-0003-3430-1090>

Itziar Alonso-Arbiol  https://orcid.org/0000-0002-4638-085X

**S1. Supplementary Online Materials**

**Newcastle-Ottawa Scale Adapted for Cross-Sectional Studies**

**Selection** (Maximum 4 stars):

1. Representativeness of the sample:
	1. Truly representative of the average in the target population. \* (random sampling)
	2. Somewhat representative of the average in the target population. \* (non-random sampling)
	3. Selected group of users.
	4. No description of the sampling strategy.
2. Sample size:
	1. Justified and satisfactory. \*
	2. Not justified.
3. Ascertainment of the measurement tool:
	1. Validated measurement tool. \*\*
	2. Non-validated measurement tool, but the tool is available or described. \*
	3. No description of the measurement tool.

**Comparability** (Maximum 1 star)**:**

1. The potential confounders were investigated by subgroup analysis or multivariable analysis.
	1. The study investigates potential confounders. \*
	2. The study does not investigate potential confounders.

**Outcome** (Maximum 3 stars):

1. Assessment of the outcome:
	1. Cross-report. \*\*
	2. Self-report. \*\*
	3. No description.
2. Statistical test:
	1. The statistical test used to analyze the data is clearly described and appropriate. \*
	2. The statistical test is not appropriate, not described or incomplete.

This scale has been adapted from the Newcastle-Ottawa Scale (Wells et al., 2000) for cohort studies to perform a quality assessment of experimental studies for the systematic review and meta-analysis: “Parental Mentalization and Children’s Externalizing Problems: A Systematic Review and Meta-Analysis.”.

In the assessment of the outcome, our measures allow for self-report, therefore we have considered both self-report and cross-report.

We have not selected the *Non-response rate,* due to its non-applicability to the studies in our meta-analysis.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  **Author** | **Selection** (Maximum 4 stars): | **Comparability** (Max 1 star): | **Outcome** (Max 3 stars): | **Score**(range 0-8) |
|  |  | **Representativeness sample** | **Sample size** | **Ascertainment** | **Potential confounders** | **Assessment outcome** | **Statistical test** |
| **1** | Benbassat & Priel (2012) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **2** | Borelli et al. (2019) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **3** | Borelli et al. (2021) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **4** | Brown et al. (2023) | \* |  | \*\* | \* | \*\* | \* | 7 |
| **5** | Camisasca et al. (2018) |  |  | \*\* | \* | \*\* | \* | 6 |
| **6** | Carlone & Milan (2021) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **7** | Centifanti et al. (2016) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **8** | Charpentier-Mora et al. (2022) | \* |  | \*\* | \* | \*\* | \* | 7 |
| **9** | Colonnesi et al. (2019) |  |  | \*\* | \* | \*\* | \* | 6 |
| **10** | Colonnesi et al. (2020) |  |  | \*\* | \* | \*\* | \* | 6 |
| **11** | Condon et al. (2019) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **12** | Dieleman et al. (2020) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **13** | Dollberg et al. (2021) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **14** | Dollberg et al. (2023) |  |  | \*\* | \* | \*\* | \* | 6 |
| **15** | Ensink, Bégin, Normandin & Fonagy (2016) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **16** | Ensink, Bégin, Normandin & Fonagy (2017)  | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **17** | Ensink, Bégin, Normandin, Godbout & Fonagy (2017) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **18** | Ensink, Lereoux et al. (2017)  | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **19** | Feniger-Schaal & Koren-Karie (2021) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **20** | Ghanbari et al. (2023) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **21** | Gray et al. (2015) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **22** | Ha et al. (2011) | \* | \* | \* | \* | \*\* | \* | 7 |
| **23** | Hobby et al. (2023) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **24** | Hughes et al. (2017) | \* |  | \*\* | \* | \*\* | \* | 7 |
| **25** | Khoshroo & Seyed Mousavi (2022) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **26** | Kochanska & An (2023) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **27** | Konijn et al. (2020) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **28** | Koren Swisa et al. (2023) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **29** | Lunn et al. (2019) | \* |  | \* | \* | \*\* | \* | 6 |
| **30** | Meins et al. (2013) |  | \* | \*\* | \* | \*\* | \* | 7 |
| **31** | Menashe-Grinberg & Atzaba-Poria (2023) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **32** | Salo et al. (2021) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **33** | Salo et al. (2022) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **34** | Shai & Belsky (2017) | \* | \* | \* | \* | \*\* | \* | 7 |
| **35** | Shalev et al. (2023) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **36** | Smaling et al. (2016) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **37** | Suardi et al. (2020) | \* |  | \*\* | \* | \*\* | \* | 7 |
| **38** | Wade et al. (2021) | \* | \* | \*\* | \* | \*\* | \* | 8 |
| **39** | Walker et al. (2012) | \* |  | \*\* | \* | \*\* | \* | 7 |
| **40** | Yatziv et al. (2020) | \* |  | \*\* | \* | \*\* | \* | 7 |
| **41** | Zeegers et al. (2020) | \* |  | \*\* | \* | \*\* | \* | 7 |
| **42** | Zimmer-Gembeck et al. (2019) | \* | \* | \*\* | \* | \*\* | \* | 8 |

**S2. Supplementary Online Materials**

*Range Years of the Child for Each Study*

In order to perform moderation analyses, the range years of the child for each study was considered. To determine age groups, the stages considered have been early childhood (0-5 years), middle childhood (6-12), and adolescence (13-18).

|  |  |  |
| --- | --- | --- |
| **Author** | **Range of Children’s Age** | **Group for the moderation analysis** |
| Benbassat & Priel (2012) | 14-18 years | Adolescence |
| Borelli et al. (2019) | 8.03–17.67 years | Adolescence |
| Borelli et al. (2021) | 12- 17 years | Adolescence |
| Brown et al. (2023) | 13,34 months | Early childhood |
| Camisasca et al. (2018) | 12-22 months | Early childhood |
| Carlone & Milan (2021) | 3- 18 years | Age range too broad (no group) |
| Centifanti et al. (2016) | 58–64 months | Middle childhood |
| Charpentier-Mora et al. (2022) | 8-12 years | Middle childhood |
| Colonnesi et al. (2021) | 2-18 years | Age range too broad (no group) |
| Colonnesi et al. (2023) | 4,5 years | Early childhood |
| Condon et al. (2019) | 4-9 years | Age range too broad (no group) |
| Dieleman et al. (2020) | 11.74-18.80 years | Adolescence |
| Dollberg et al. (2021) | 3-5 years | Early childhood |
| Dollberg et al. (2023) | 3-6 years | Early childhood |
| Ensink, Bégin, Normandin & Fonagy (2016) | 7-12 years | Middle childhood |
| Ensink, Bégin, Normandin & Fonagy (2017) | 35–143 months | Age range too broad (no group) |
| Ensink, Bégin, Normandin, Godbout & Fonagy (2017) | 5-12 years | Middle childhood |
| Ensink, Lereoux et al. (2017) | 7-12 years | Middle childhood |
| Feniger-Schaal & Koren-Karie (2021) | 5- 12 years | Middle childhood |
| Ghanbari et al. (2023) | 2-12 years | Age range too broad (no group) |
| Gray et al. (2015) | 3-5 years | Early childhood |
| Ha et al. (2011) | 7 -11 years | Middle childhood |
| Hobby et al. (2023) | 51 months | Early childhood |
| Hughes et al. (2017) | 12 years | Adolescence |
| Khoshroo & Seyed Mousavi (2022)-o | 6-10 years | Middle childhood |
| Khoshroo & Seyed Mousavi (2022)-y | 3-6 years | Early childhood |
| Kochanska & An (2023) | 3 years | Early childhood |
| Konijn et al. (2020) | 1.8–17.3 years | Age range too broad (no group) |
| Koren Swisa et al. (2024) | 3-5 years | Early childhood |
| Lunn et al. (2019) | 9-16 years | Age range too broad (no group) |
| Meins et al. (2013) | 42-64 months | Early childhood |
| Menashe-Grinberg & Atzaba-Poria (2023) | 1–6.5 years  | Age range too broad (no group) |
| Salo et al. (2021) | 12 months | Early childhood |
| Salo et al. (2022) | 30 months | Early childhood |
| Shai & Belsky (2017) | 54 months  | Early childhood |
| Shalev et al. (2023) | 1.30-9.39 years | Age range too broad (no group |
| Smaling et al. (2016) | 20 months | Early childhood |
| Suardi et al (2020) | 12-42 months | Early childhood |
| Wade et al. (2021) | 18 months | Early childhood |
| Walker et al. (2012) | 36-54 months | Early childhood |
| Yatziv et al. (2020) | 66 months | Early childhood |
| Zeegers et al. (2020) | 3- 11 years | Age range too broad (no group) |
| Zimmer-Gembeck et al. (2019) | 29 to 83 months | Early childhood |

 *Note*. O = Older (middle) childhood; Y = Younger (early) childhood.

**S3. Supplementary Online Materials**

|  |  |
| --- | --- |
| Mentalization Questionnaires  | *k* |
|  Mind-Mindedness (MM; Meins & Fernyhough, 2015) | 14 |
| Parental Reflective Functioning Questionnaire (PRFQ; Luyten et al., 2017) | 11 |
| Parent Development Interview (PDI; Sleed et al., 2020) | 8 |
| Insightfulness (Oppenheim & Koren-Karie, 2002) | 2 |
| Five Minutes Speech Sample (FMSS, Gottschalk & Gleser, 1969) | 2 |
| Social Stories (Sharp et al., 2007)  | 2 |
| Reflective Function Questionnaire (RFQ; Fonagy & Ghinai, 2008) | 1 |
| Working Model of the Child Interview (WMCI; Zeanah & Benoit, 1995)Parental embodied mentalizing | 11 |
| Behavior Problem Questionnaires | *k* |
| The Child Behavior Checklist (CBCL; Achenbach, 1999) | 23 |
| Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) | 9 |
| Brief Infant-Toddler Social and Emotional Assessment (BITSEA*;* Briggs-Gowan & Carter, 2007) | 2 |
| The infant–toddler social and emotional assessment (ITSEA; Carter et al., 2003) | 2 |
| Behavior Assessment System for Children Second Edition (BASC-2; Kamphaus& Reynolds, 2004) | 1 |
| Behavioral and Emotional Screening System (BASC-3 BESS; Kamphaus & Reynolds, 2015) | 1 |
| Eyberg Child Behavior Inventory (ECBI; Eyberg & Pincus, 1999) | 1 |
| Social Competence and Behavior Evaluation Scale (SCBE-30; LaFreniere, & Dumas,1996) | 1 |
| Youth Self-Report (YSR; Achenbach & Rescorla, 2001)Teacher Report Form (Achenbach, 1991). | 11 |
| Moderators  | *k* |
| Parental mental health —e.g., depression, trauma, anxiety, stress— | 19 |
| Children mental health —e.g., sexual abuse— | 11 |
| Parental sensitivity | 6 |
| Parenting behaviors | 6 |
| Children theory of mind/mentalization | 4 |
| Attachment | 2 |
| Child temperament | 2 |
| Studies with interventions programs | 7 |

*Questionnaires and Moderators Analyzed in the Studies Included in the Meta-Analysis*

**References**

Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF proﬁles*.

Achenbach, T. M. (1999). The Child Behavior Checklist and related instruments. In M. E. Maruish (Ed.), The use of psychological testing for treatment planning and outcomes assessment (pp. 429–466). Lawrence Erlbaum Associates.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and proﬁles*.

Briggs-Gowan, M. J., & Carter, A. S. (2007). Brief Infant Toddler Social Emotional Assessment (BITSEA). <https://doi.org/10.1037/t14990-000>

Carter, A. S., Briggs-Gowan, M. J., Jones, S. M., & Little, T. D. (2003). The infant–Toddler Social and Emotional Assessment (ITSEA): Factor structure, reliability, and validity. *Journal of Abnormal Child Psychology*, *31*, 495-514. [https://doi.org/10.1023/A:1025449031360](https://doi.org/10.1023/A%3A1025449031360)

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

Fonagy, P., & Ghinai, R. A. (2008). *A self-report measure of mentalizing development and preliminary test of reliability and validity of the Reﬂective Function Questionnaire (RFQ)* [Unpublished manuscript]. University College of London.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>

Gottschalk, L. A., & Gleser, G. C. (1969). *The measurement of psychological states through the content analysis of verbal behavior*. University of California Press.

Kamphaus, R. W., & Reynolds, C. R. (2007). BASC–2 Behavioral and Emotional Screening System (BASC-2 BESS). <https://doi.org/10.1037/t29902-000>

Kamphaus, R. W., & Reynolds, C. R. (2015). *Behavior Assessment System for Children–Third Edition (BASC-3): Behavioral and emotional Screening System (BESS).* Pearson.

LaFreniere, P. J., & Dumas, J. E. (1996). Social Competence and Behavior Evaluation in children ages 3 to 6 years: The short form (SCBE-30). *Psychological Assessment*, *8*(4), 369-377. <https://doi.org/10.1037/1040-3590.8.4.369>

Luyten, P., Mayes, L. C., Nijssens, L., & Fonagy, P. (2017). The parental reflective functioning questionnaire: Development and preliminary validation. *PloS ONE,* *12*(5), e0176218. <https://doi.org/10.1371/journal.pone.0176218>

Meins, E., & Fernyhough, C. (2015). *Mind-mindedness coding manual, Version 2.2*. Unpublished manuscript. University of York, UK.

Oppenheim, D., & Koren-Karie, N. (2002). Mothers’ insightfulness regarding their children’s internal worlds: The capacity underlying secure child-mother relationships. *Infant Mental Health Journal, 23*(6), 593–605. <https://doi.org/10.1002/imhj.10035>

Sharp, C., Croudace, T. J., & Goodyer, I. M. (2007). Biased mentalizing in children aged seven to 11: Latent class confirmation of response styles to social scenarios and associations with psychopathology. *Social Development*, *16*(1), 181-202. <https://doi.org/10.1111/j.1467-9507.2007.00378.x>

Sleed, M., Slade, A., & Fonagy, P. (2020). Reflective functioning on the parent development interview: Validity and reliability in relation to socio-demographic factors. *Attachment & Human Development, 22*(3), 310-331. <https://doi.org/10.1080/14616734.2018.1555603>

Wells, G. A., Shea, B., O’Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2000). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses.

Zeanah, C. H., & Benoit, D. (1995). Clinical applications of a parent perception interview in infant mental health. *Child and Adolescent Psychiatric Clinics of North America*, *4*(3), 539-554. [https://doi.org/10.1016/S1056-4993(18)30418-8](https://doi.org/10.1016/S1056-4993%2818%2930418-8)