

## **FIERCE dataset guide**

**Authors:** Julia Michalek, Mays Al-Hamad, Rana Dajani, Amal El Kharouf, Qusai Khraisha, Isabelle Mareschal, Lina Qtaishat, Sophie von Stumm, & Kristin Hadfield

The data collection for all timepoints took place in participants' homes, in Arabic, and the surveys were completed using KoBoToolbox offline survey tool. Two female fieldworkers visited participants' homes at each timepoint and simultaneously collected data from the mother and child. Two male fieldworkers collected data from the fathers over the phone at two timepoints. The fieldworkers were all native Arabic speakers and were trained on all measures before the data collection process. Before the start of the data collection process at each timepoint, all measures were also piloted with families who did not take part in the FIERCE project. All information was collected verbally by the fieldworkers and data collection in each household lasted on average 1 hour. Visual scales were used to aid the participants' understanding of the Likert scales.

Wherever possible, we used locally developed measures in Arabic (e.g., Father Involvement Questionnaire), or measures developed elsewhere but previously validated in Arabic (e.g., Dyadic Adjustment Scale). If this was not possible, we used measures originally developed and validated in English (e.g., Reader Self Concept Scale) and used a team translation framework, where questions were translated and back translated by members of the research group. This was done to ensure that all items were understandable and valid in Arabic and were relevant in the local context. The full surveys used at all three timepoints are available in English and Arabic on the OSF page in the Planning and Measures section. The information on which measures were completed at which timepoint are available on the OSF page and at the end of this document.

The data from all surveys and the syntax used to compute the final variables are available on the OSF page (Knowledge Translation/FIERCE data). All the measures included in the study are presented in Table 1 (pages 25-26).

### **Abbreviations:**

**T1 = Timepoint 1**

**T2 = Timepoint 2**

**T3 = Timepoint 3**

**C = Child**

**M = Mother**

**F = Father**

## Table of Contents

<b>GENERAL INFORMATION.....</b>	<b>4</b>
<b>Demographic information .....</b>	<b>4</b>
<b>Study relevant information .....</b>	<b>4</b>
<b>Exclusions .....</b>	<b>4</b>
<b>WLR intervention fidelity .....</b>	<b>4</b>
<b>Reliability measures.....</b>	<b>5</b>
<b>MEASURES.....</b>	<b>6</b>
<b>1. Reading .....</b>	<b>6</b>
Holistic Assessment of Learning and Development Outcomes - Literacy .....	6
Parents Digital Literacy Questionnaire .....	6
Young Reader Motivation Questionnaire .....	7
Reader Self-Concept Scale – attitude towards reading subscale .....	7
Adult Reading Motivation Scale.....	8
<b>2. Mental health and wellbeing .....</b>	<b>8</b>
K6 Psychological Distress Scale.....	8
Human Insecurity and Distress Scale – Insecurity subscale .....	8
Depression, Anxiety, and Stress Scale – Anxiety subscale .....	9
Centre for Epidemiological Studies – Depression Scale .....	9
PTSD Checklist for DSM-5.....	9
Warwick-Edinburgh Mental Wellbeing Scale .....	10
Insomnia Severity Index .....	10
Holistic Assessment of Learning and Development Outcomes – SEL (Social and Emotional Learning)	11
<b>2.1 Mothers reporting on child mental health.....</b>	<b>11</b>
Paediatric Symptoms Checklist .....	11
<b>3. Family dynamics .....</b>	<b>12</b>
Arabic Revised Dyadic Adjustment Scale .....	12
Pianta Child-Parent Relationship Scale .....	12
Father Involvement Questionnaire.....	13
General Functioning Subscale .....	14
Alabama Parenting Questionnaire – short form.....	14
Parental Self-Agency Measure – positive items .....	15
Parental Burnout Assessment .....	15
Arabic Experience in Close Relationships Scale - Revised 16.....	16
<b>4. Other .....</b>	<b>16</b>
Household Wealth Index.....	16
COVID-19 Household Environment Scale – Part 2 .....	16
Traumatic Events Checklist .....	17
Collective Efficacy Scale.....	17
Outgroup Attitudes Scale.....	18
Short Form Spirituality-S-4 .....	18
<b>5. Cognitive tasks .....</b>	<b>18</b>
Go/NoGo .....	19
Dot probe .....	19
Simon task.....	19
Trust task.....	20
Sharing tasks .....	20
Raven’s Progressive Matrices.....	21
Delayed gratification.....	21

Number recall (Digit Span Task).....	21
Continuous Performance Task (CPT, Space Adventure) .....	21
<b>6. Observational tasks.....</b>	<b>22</b>
Mother-child observational tasks .....	22
Puzzle .....	22
<i>Table 1. All variables collected at the three timepoints .....</i>	<i>25</i>
<b>REFERENCES.....</b>	<b>26</b>

## **GENERAL INFORMATION**

### **Demographic information**

Age of all participants, gender of participating children, years living in Jordan, education information, languages spoken, employment information (mothers, fathers).

### **Study relevant information**

The participating mothers and fathers were asked if they were able to read, to estimate the parents' literacy levels at each timepoint (mothers at T1-T3, father at T1-T2). The mothers reported on the numbers of people in the household at T1. Information on the participating mothers' weight and height, and the participating children' mid-upper arm circumference (MUAC) was collected at T3. All measurements were taken by the fieldworkers using a scale and a tape measure. The data from the mothers was then used to calculate their body mass index (BMI, calculated as  $\text{weight}/\text{height}^2$ ). At T2 and T3 the fieldworkers reported on whether the mother was present in the room during the child data collection (0 = *No*, 1 = *Yes, but it didn't affect it*, 2 = *Yes, and it affected the testing*).

### **Exclusions**

The total scores for all variables were computed using listwise deletion (i.e., if any value was missing from a given measure, the total score for that measure was a system missing value).

Additionally, to ensure that any issues which occurred during data collection were taken into account during the analyses, the fieldworkers completed 'Data collection field challenges' documents for the 3 testing timepoints. These documents were then used to code additional variables for participant exclusion: T1/T2/T3\_surveyIssues. These variables include information on the types of challenges present during data collection for each family at each timepoint and are coded as follows: 0 (*no issues*), 1 (*lack of privacy*), 2 (*issues with the mother's survey*), 3 (*issues with the child's survey*), and 4 (*issues with both surveys*). Those participants whose surveys included field challenges were excluded from the analyses reported below (descriptive statistics and internal reliability).

### **WLR intervention fidelity**

The fidelity of the intervention was assessed at three timepoints. Firstly, from the beginning of the reading sessions, families were followed up weekly for 12 weeks on average to record the children's attendance and the ambassadors' sessions information. This data was then used to calculate the percentage of sessions the children attended (*PercentSessionsAttended*:  $\text{number of sessions attended} / \text{number of sessions available}$ ) and the percentage of the weeks when the children attended at least one session (*PercentWeeksAttended*:  $\text{number of weeks when the children attended a session} / \text{number of weeks when a session was available}$ ).

Secondly, 12 months after the start of the programme, a subset of families were followed up once to record the number of sessions the children attended in the past three months (*NoSessionsAfter12months*). Finally, another follow up was conducted with a subset of families who took part in the T3 data collection, to record the number of sessions the children attended between the previous follow up and the time of the T3 data collection (*NoSessionsAtT3*).

## Reliability measures

The internal reliability of all survey measures was calculated using Cronbach's Alpha.

To determine consistency among raters for the components of HALDO (literacy and socio-emotional learning), the interrater reliability analysis was performed using the interclass correlation coefficients (ICC). In order to do this, the HALDO responses from a subset of participating children ( $n = 74$  at T1,  $n = 92$  at T2, and  $n = 27$  at T3) were independently coded by two fieldworkers, and ICC model 1 with two raters was conducted using these variables. The variables below are these interrater reliability analyses (ICC).

### Variables:

- Expressive language:
  - LIT11\_T1\_rel
  - LIT11\_T2\_rel
  - LIT11\_T3\_rel
- Letter identification
  - T1\_LETTER\_IDENTIFICATION\_rel
  - T2\_LETTER\_IDENTIFICATION\_rel
  - T3\_LETTER\_IDENTIFICATION\_rel
- Reading comprehension
  - LIT5\_T1\_rel
  - LIT5\_T2\_rel
  - LIT5\_T3\_rel
- Self-concept
  - T2\_SEL\_selfConcept\_rel
  - T3\_SEL\_selfConcept\_rel
- Ambitions
  - T2\_SEL\_ambitions\_rel
  - T3\_SEL\_ambitions\_rel
- Empathy
  - T2\_SEL\_empathy\_rel
  - T3\_SEL\_empathy\_rel

## MEASURES

### 1. Reading

#### Holistic Assessment of Learning and Development Outcomes - Literacy

HALDO; D'Sa, Krupar, & Westrope, 2019

This is an adaptive measure consisting of three components: letter identification, expressive language, and reading comprehension. Children begin by completing common letter identification (LIT1), which is followed by expressive language (LIT1.1) if they cannot identify any common letters, and by uncommon letter identification (LIT2) and reading comprehension (LIT3-5) if they can identify at least one common letter. Children's responses are marked as 0 (*incorrect*), 1 (*correct*), or 999 (*refused/skipped*). The total scores range from 0 to 15, with higher scores indicating higher degree of literacy. Children completed this measure at all three timepoints. The interrater reliability (ICC) is reported for each HALDO LIT component.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
x			x		

Variables:

- Expressive language:
  - T1\_LIT1.1:  $N = 103$ ;  $M = 4.74$ ;  $SD = 3.33$ ; range = 0-10; ICC = .998
  - T2\_LIT1.1:  $N = 92$ ;  $M = 5.63$ ;  $SD = 3.27$ ; range = 0-10; ICC = .998
  - T3\_LIT1.1:  $N = 12$ ;  $M = 4.67$ ;  $SD = 3.17$ ; range = 0-10; ICC = 1.000
- Letter identification:
  - T1\_LETTER\_IDENTIFICATION:  $N = 201$ ;  $M = 5.27$ ;  $SD = 3.35$ ; range = 1-10; ICC = .999
  - T2\_LETTER\_IDENTIFICATION:  $N = 200$ ;  $M = 5.26$ ;  $SD = 3.34$ ; range = 1-10; ICC = .996
  - T3\_LETTER\_IDENTIFICATION:  $N = 90$ ;  $M = 7.11$ ;  $SD = 3.07$ ; range = 1-10; ICC = .999
- Reading comprehension:
  - T1\_LIT5:  $N = 19$ ;  $M = 3.95$ ;  $SD = 1.13$ ; range = 1-5; ICC = 1.000
  - T2\_LIT5:  $N = 20$ ;  $M = 4.35$ ;  $SD = 1.18$ ; range = 0-5; ICC = 1.000
  - T3\_LIT5:  $N = 35$ ;  $M = 4.37$ ;  $SD = 1.22$ ; range = 0-5; ICC = .978
- HALDO total score:
  - T1\_HALDO\_TOTAL:  $N = 304$ ;  $M = 5.26$ ;  $SD = 3.78$ ; range = 0-15
  - T2\_HALDO\_TOTAL:  $N = 289$ ;  $M = 5.61$ ;  $SD = 3.76$ ; range = 0-15
  - T3\_HALDO\_TOTAL:  $N = 102$ ;  $M = 8.32$ ;  $SD = 4.82$ ; range = 0-15

#### Parents Digital Literacy Questionnaire

PPCATR; Ozturk & Ohi, 2018

This mother-reported measure consists of 7 items on a 5-point Likert scale ranging from 0 (*never*) to 4 (*always*). Mothers' perception of the study child's attitude towards reading score

is calculated by summing all items and can range from 0 to 28. This data was collected from the participating mothers at all three timepoints.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>x</b>		

Variables:

- T1\_M\_PPCATR:  $N = 399$ ;  $M = 18.24$ ;  $SD = 5.73$ ; range = 0-28;  $\alpha = .715$
- T2\_M\_PPCATR:  $N = 292$ ;  $M = 19.57$ ;  $SD = 5.63$ ; range = 0-28;  $\alpha = .796$
- T3\_M\_PPCATR:  $N = 97$ ;  $M = 20.43$ ;  $SD = 5.20$ ; range = 4-28;  $\alpha = .798$

### Young Reader Motivation Questionnaire

#### YRMQ, Coddington & Guthrie, 2009

The subscales used were 4 items measuring the efficacy for reading and 3 items measuring the reading orientation. YRMQ is a 4-point Likert scale ranging from 1 (*no, never*) to 4 (*yes, always*). The total scores are calculated by summing all items and can range from 7 to 28, with higher scores indicating a more positive attitude towards reading. This child-reported data was collected at all three timepoints.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>x</b>		

Variables:

- T1\_YRMQ:  $N = 303$ ;  $M = 21.63$ ;  $SD = 4.93$ ; range = 7-28;  $\alpha = .793$
- T2\_YRMQ:  $N = 290$ ;  $M = 20.16$ ;  $SD = 5.15$ ; range = 7-28;  $\alpha = .782$
- T3\_YRMQ:  $N = 101$ ;  $M = 19.74$ ;  $SD = 4.67$ ; range = 7-28;  $\alpha = .760$

### Reader Self-Concept Scale – attitude towards reading subscale

#### RSCS, Chapman & Tunmer, 1995

8-item *yes/no* measure assessing child's attitudes towards reading. The total score is calculated by summing all items and can range from 0 to 8, with higher scores indicating that children have a more positive attitude towards reading. This data was collected from the participating children at all timepoints.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>x</b>		

Variables:

- T1\_RSCS:  $N = 300$ ;  $M = 7.07$ ;  $SD = 1.60$ ; range = 0-8;  $\alpha = .787$
- T2\_RSCS:  $N = 189$ ;  $M = 6.87$ ;  $SD = 1.60$ ; range = 0-8;  $\alpha = .724$
- T3\_RSCS:  $N = 102$ ;  $M = 6.70$ ;  $SD = 1.84$ ; range = 0-8;  $\alpha = .791$

Adult Reading Motivation Scale  
ARMS; Schutte & Malouff, 2007

6 items measuring mothers' and fathers' attitudes towards reading. ARMS is a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The total score is calculated by summing all items and can range from 6 to 30, with higher scores indicating more positive attitudes towards reading. This scale was completed by participating mothers at T1 and T2, and fathers at T1.

T	Person who completed the measure:			Reporting on:		
	Mother	Father	Child	Mother	Father	Child
1,2	x			x		
1		x			x	

Variables:

- T1\_M\_ARM\_total:  $N = 249$ ;  $M = 22.86$ ;  $SD = 3.73$ ; range = 10-30;  $\alpha = .712$
- T2\_M\_ARM\_total:  $N = 250$ ;  $M = 23.32$ ;  $SD = 3.43$ ; range = 12-30;  $\alpha = .684$
- T1\_F\_ARM\_total:  $N = 136$ ;  $M = 21$ ;  $SD = 4.78$ ; range = 7-30;  $\alpha = .666$

## 2. Mental health and wellbeing

K6 Psychological Distress Scale  
K6; Kessler et al., 2003

6-item, 5-point Likert scale measuring distress, with answers ranging from 1 (*all of the time*) to 5 (*none of the time*). The total score is calculated by summing all items and can range from 6 to 30, with lower scores indicating higher levels of psychological distress. This scale was completed by the participating mothers at T1 and T3.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
x			x		

Variables:

- T1\_M\_K6\_total:  $N = 272$ ;  $M = 15.23$ ;  $SD = 5.63$ ; range = 6-30;  $\alpha = .810$
- T3\_M\_K6\_total:  $N = 99$ ,  $M = 16.13$ ,  $SD = 4.50$ , range = 6-27,  $\alpha = .748$

Human Insecurity and Distress Scale – Insecurity subscale  
HIDS-I, Ziadni et al., 2011

10 items on a 4-point Likert scale ranging from 1 (*never*) to 4 (*always*). The total insecurity score is calculated by summing all items and can range from 10 to 40, with higher scores indicating more fear and worry. This data was collected from the participating mothers at T1.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
x			x		



Variables:

- T1\_M\_HIS\_total:  $N = 275$ ;  $M = 34.38$ ;  $SD = 4.95$ ; range = 17-40;  $\alpha = .782$

#### Depression, Anxiety, and Stress Scale – Anxiety subscale

DASS-21-A, Henry & Crawford, 2005

7 items related to anxiety on a 4-point Likert scale ranging from 0 (*did not apply to me – never*) to 3 (*applied to me very much, or most of the time – almost always*). The participants were asked to what extent each statement applies to their feelings over the past week. The total score is derived by summing all scores and can range from 0 to 21, with higher score indicating more anxiety symptoms. This anxiety data was collected from the participating mothers and fathers at T2.

T	Person who completed the measure:			Reporting on:		
	Mother	Father	Child	Mother	Father	Child
2	x			x		
2		x			x	

Variables:

- T2\_M\_DASS21:  $N = 289$ ;  $M = 6.60$ ;  $SD = 4.60$ ; range = 0-20;  $\alpha = .785$
- T2\_F\_DASS:  $N = 104$ ;  $M = 2.51$ ;  $SD = 3.29$ ; range = 0-16;  $\alpha = .798$

#### Centre for Epidemiological Studies – Depression Scale

CES-D Radloff, 1977

10 items on a 4-point Likert scale ranging from 0 (*rarely, or none of the time*) to 3 (*most, or almost all the time*) related to depressive symptoms. The participants were asked about the way they have felt or behaved over the past week. Two positive items are reverse scored (item 5 and item 8). The scores are then summed and can range from 0 to 30, with higher scores indicating that a participant has more depressive symptoms. This data was collected from the participating mothers and fathers at T2.

T	Person who completed the measure:			Reporting on:		
	Mother	Father	Child	Mother	Father	Child
2	x			x		
2		x			x	

Variables:

- T2\_M\_CESD:  $N = 284$ ;  $M = 13.95$ ;  $SD = 6.27$ ; range = 0-30;  $\alpha = .791$
- T2\_F\_CESD:  $N = 102$ ;  $M = 9.18$ ;  $SD = 6.60$ ; range = 0-24;  $\alpha = .852$

#### PTSD Checklist for DSM-5

PCL-5, Blevins et al., 2015

20 items, 5-point Likert scale with responses ranging from 0 (*not at all*) to (*extremely*) to a list of problems that people might have in response to a very stressful experience. The participants were asked to keep their worst event in mind and indicate how much they have been bothered

by each problem in the past month. The total symptom severity score is calculated by summing all scores and can range from 0 to 80, with higher scores indicating more symptoms of post-traumatic stress disorder. This data was collected from a subset ( $n = 129$ ) of participating mothers at T2.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>x</b>		

Variables:

- T2\_M\_PCL5:  $N = 124$ ;  $M = 34.96$ ;  $SD = 16.24$ ; range = 0-68;  $\alpha = .920$

### Warwick-Edinburgh Mental Wellbeing Scale

WEMWBS; Tennant et al., 2007

This 14-item, 5-point Likert scale measures mothers' and fathers' feelings of wellbeing and optimism in the past two weeks, with answers ranging from 1 (*none of the time*) to 5 (*all of the time*). The total score is derived by summing all items, and ranges from 14 to 70, with higher scores suggestive of higher wellbeing. This data was collected from a subset ( $n = 130$ ) of the participating mothers at T2, from all fathers at T2, and from all mothers at T3.

T	Person who completed the measure:			Reporting on:		
	Mother	Father	Child	Mother	Father	Child
<b>2,3</b>	<b>x</b>			<b>x</b>		
<b>2</b>		<b>x</b>			<b>x</b>	

Variables:

- T2\_M\_WEMWBS:  $N = 126$ ;  $M = 50.08$ ;  $SD = 9.07$ ; range = 28-69;  $\alpha = .824$
- T3\_M\_WEMWBS:  $N = 99$ ,  $M = 52.35$ ,  $SD = 8.21$ , range = 25-68,  $\alpha = .817$
- T2\_F\_WEMWBS:  $N = 94$ ;  $M = 57.13$ ;  $SD = 9.75$ ; range = 23-70;  $\alpha = .871$

### Insomnia Severity Index

ISI; Morin et al., 2011

A subset of 5 items on a 5-point Likert scale. 3 items measure current sleep problems, with answers ranging from 0 (*none*) to 4 (*very severe*). 1 item measures participants' satisfaction with their current sleep pattern from 0 (*very satisfied*) to 4 (*very dissatisfied*), and 1 item measures the level of worry about sleep from 0 (*not worried at all*) to 4 (*very much worried*). All items are summed to create the total score, which can range from 0 to 20, with higher scores indicating more sleep problems. This data was collected from the participating mothers at T3.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>x</b>		

Variables:

- T3\_ISI:  $N = 99$ ;  $M = 9.17$ ;  $SD = 4.40$ ; range = 0-20;  $\alpha = .713$

## Holistic Assessment of Learning and Development Outcomes – SEL (Social and Emotional Learning)

HALDO; D'Sa, Krupar, & Westrope, 2019

16-item adaptive subscale measuring three components: self-concept, ambitions, and empathy. The children's responses are marked as 0 (*incorrect*), 1 (*correct*), or 999 (*refused/skipped*). The total score ranges from 0 to 16, with higher scores indicating higher levels of socio-emotional learning (subscale scores range: 0-6 for self-concept, 0-5 for ambitions, and 0-5 for empathy). The participating children completed this measure at T2 and T3. The interrater reliability (ICC) is reported for each HALDO SEL component below.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
x			x		

Variables:

- T2\_SEL\_total:  $N = 271$ ;  $M = 8.51$ ;  $SD = 3.41$ ; range = 0-16
- T2\_SEL\_selfConcept:  $N = 282$ ;  $M = 4.28$ ;  $SD = 1.37$ ; range = 0-6; ICC = .982
- T2\_SEL\_ambitions:  $N = 282$ ;  $M = 2.03$ ;  $SD = 1.58$ ; range = 0-5; ICC = .965
- T2\_SEL\_empathy:  $N = 284$ ;  $M = 2.19$ ;  $SD = 1.43$ ; range = 0-5; ICC = .976
- T3\_SEL\_total:  $N = 93$ ;  $M = 9.35$ ;  $SD = 3.14$ ; range = 2-15
- T3\_SEL\_selfConcept:  $N = 98$ ;  $M = 4.70$ ;  $SD = 1.27$ ; range = 1-6; ICC = .982
- T3\_SEL\_ambitions:  $N = 98$ ;  $M = 1.84$ ;  $SD = 1.26$ ; range = 0-5; ICC = .980
- T3\_SEL\_empathy:  $N = 98$ ;  $M = 2.73$ ;  $SD = 1.41$ ; range = 0-5; ICC = .952

### **2.1 Mothers reporting on child mental health**

#### Paediatric Symptoms Checklist

PSC-17, Jellinek et al., 1988

The PSC consists of three subscales measuring separate constructs: child internalising, externalising, and attention problems, on a 3-point Likert scale ranging from 0 (*never*) to 2 (*often*). The total scores for the three subscales are calculated by summing all scores within each subscale, with higher scores indicating the presence of more symptomatology. This data was collected from the participating mothers at T2, about the study child.

The internalising subscale (PSC-Int) includes 5 items and contains statements related to the child's feelings of anxiety and depression. PSC-Int scores can range from 0 to 10.

The externalising subscale (PSC-Ext) includes 7 items related to the child's conduct and behavioural problems. PSC-Ext scores can range from 0 to 14.

The attention problems subscale (PSC-Att) includes 5 items which relate to the child's inattention and hyperactivity. PSC-Att scores can range from 0 to 10.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
x			x		

Variables:

- T2\_PSCExt:  $N = 291$ ;  $M = 5.09$ ;  $SD = 3.06$ ; range = 0-14;  $\alpha = .742$
- T2\_PSCInt:  $N = 294$ ;  $M = 3.21$ ;  $SD = 2.02$ ; range = 0-9;  $\alpha = .594$
- T2\_PSCAtt:  $N = 293$ ;  $M = 5.19$ ;  $SD = 2.38$ ; range = 0-10;  $\alpha = .646$

### 3. Family dynamics

#### Arabic Revised Dyadic Adjustment Scale

A-RDAS; Elanbari, 2015 [originally developed by Spanier et al., 1976]

A subset of 8 questions pertaining to the quality of the mother-father relationship. 4 items measure how often participants engage in positive activities with their partner on a 6-point Likert scale ranging from 0 (*never*) to 5 (*more often*). 4 items measure negative aspects of a relationship on a 6-point Likert scale ranging from 0 (*all of the time*) to 5 (*never*). The total DAS score is calculated by summing all items and can range between 0 and 40, with higher scores indicating better-quality spousal relationship. This data was collected from the participating mothers at T1 and T2.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
x			mother-father relationship		

Variables:

- T1\_M\_DAS\_total:  $N = 261$ ;  $M = 22.33$ ;  $SD = 6.59$ ; range = 1-34;  $\alpha = .773$
- T2\_M\_DAS\_total:  $N = 257$ ;  $M = 26.11$ ;  $SD = 8.47$ ; range = 1-39;  $\alpha = .827$

#### Pianta Child-Parent Relationship Scale

Pianta, 1992; Driscoll & Pianta, 2011

Short form containing 15 items on a 5-point Likert scale ranging from 1 (*definitely does not apply*) to 5 (*definitely applies*). 7 positive items are summed to create a measure of mother-child closeness (items 1, 3, 5-7, 9, 15) where scores can range between 7 and 35, with higher scores indicating more closeness in mother-child relationships. 8 negative items are summed to create a measure of mother-child conflict (items 2, 4, 8, 10-14) where scores can range between 8 and 40, with higher scores indicating more conflict in mother-child relationships. This data was collected from the participating mothers are all three timepoints.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
x			mother-child relationship		

Variables:

- T1\_M\_PiantaCloseness:  $N = 290$ ;  $M = 32.90$ ;  $SD = 2.65$ ; range = 17-35;  $\alpha = .527$
- T2\_M\_PiantaCloseness:  $N = 286$ ;  $M = 32.57$ ;  $SD = 2.60$ ; range = 22-35;  $\alpha = .628$
- T3\_M\_PiantaCloseness:  $N = 98$ ;  $M = 31.63$ ;  $SD = 3.67$ ; range = 15-35;  $\alpha = .768$
- T1\_M\_PiantaConflict:  $N = 279$ ;  $M = 23.39$ ;  $SD = 6.57$ ; range = 8-39;  $\alpha = .667$
- T2\_M\_PiantaConflict:  $N = 285$ ;  $M = 22.85$ ;  $SD = 6.41$ ; range = 9-36;  $\alpha = .717$

- T3\_M\_PiantaConflict:  $N = 99$ ;  $M = 23.88$ ;  $SD = 5.88$ ; range = 8-36;  $\alpha = .701$

### Father Involvement Questionnaire

FIQ; Hein et al., 2020

20 items on a 4-point Likert scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*) measuring the perceived engagement of the father in three contexts. 9 items (items 1-9) assess the fathers' fathering involvement with the child. 2 negatively worded questions (items 8 and 9) are reverse-scored and all items in this subscale are summed for the total father-child subscale score, which can range between 9 and 36. 6 items (items 10-15) assess the fathers' fathering involvement with the mother and the total scores on this father-mother subscale can range between 6 and 24. 5 items (items 16-20) assess the fathers' fathering involvement with the community, with the possible total scores ranging between 5 and 20. All 20 items are also summed for the total score of the father's involvement. Higher scores indicate higher father engagement.

The FIQ was completed by the participating mothers and fathers to measure the perception of the fathers' involvement from the perspective of both parents. The mothers at T1 and T2 completed the measurement with reverse scoring, i.e., 1 = *strongly agree*, 4 = *strongly disagree*. Corrected variables are saved as T1\_M\_FIQ1-20Rev and T2M\_FIQ1-20\_Rev.

Additionally, spousal discordance was calculated by subtracting the mother's scores from the father's scores for each subscale and each timepoint. This gives an indication of how differently the mothers and fathers view the level of the fathers' involvement with the child, mother, and community. As we were interested in how differently the mothers and fathers view the fathers' involvement (regardless of the direction of this difference), all spousal discordance scores are absolute values.

This data was collected from the participating mothers and fathers at T1 and T2.

T	Person who completed the measure:			Reporting on:		
	Mother	Father	Child	Mother	Father	Child
1,2	x			father's involvement with others		
1,2	x			father's involvement with others		

### Variables:

- T1\_M\_FIQTot:  $N = 186$ ;  $M = 58.93$ ;  $SD = 9.95$ ; range = 24-78;  $\alpha = .893$
- T1\_M\_FIQChild:  $N = 205$ ;  $M = 26.46$ ;  $SD = 5.06$ ; range = 11-36;  $\alpha = .824$
- T1\_M\_FIQMother:  $N = 211$ ;  $M = 19.40$ ;  $SD = 3.76$ ; range = 6-24;  $\alpha = .894$
- T1\_M\_FIQCommunity:  $N = 207$ ;  $M = 12.84$ ;  $SD = 3.29$ ; range = 5-20;  $\alpha = .752$
- T1\_F\_FIQ\_total:  $N = 155$ ;  $M = 65.84$ ;  $SD = 6.50$ ; range = 47-80;  $\alpha = .713$
- T1\_F\_FIQchild:  $N = 156$ ;  $M = 30.23$ ;  $SD = 3.81$ ; range = 19-36;  $\alpha = .621$
- T1\_F\_FIQmother:  $N = 158$ ;  $M = 22.44$ ;  $SD = 1.87$ ; range = 14-24;  $\alpha = .671$
- T1\_F\_FIQcommunity:  $N = 157$ ;  $M = 13.19$ ;  $SD = 3.66$ ; range = 5-20;  $\alpha = .685$
- T1\_FIQMother\_Diff:  $N = 147$ ;  $M = 3.99$ ;  $SD = 3.16$ ; range = 0-18
- T1\_FIQChild\_Diff:  $N = 139$ ;  $M = 5.32$ ;  $SD = 4.32$ ; range = 0-22
- T1\_FIQCommunity\_Diff:  $N = 139$ ;  $M = 3.37$ ;  $SD = 2.54$ ; range = 0-13
- T2\_M\_FIQTot:  $N = 241$ ;  $M = 58.88$ ;  $SD = 8.44$ ; range = 27-76;  $\alpha = .883$

- T2\_M\_FIQChild:  $N = 252$ ;  $M = 25.92$ ;  $SD = 4.99$ ; range = 9-36;  $\alpha = .856$
- T2\_M\_FIQMother:  $N = 248$ ;  $M = 19.52$ ;  $SD = 3.20$ ; range = 9-24;  $\alpha = .849$
- T2\_M\_FIQCommunity:  $N = 251$ ;  $M = 13.37$ ;  $SD = 2.38$ ; range = 5-19;  $\alpha = .704$
- T2\_F\_FIQTot:  $N = 103$ ;  $M = 67.17$ ;  $SD = 5.41$ ; range = 51-77;  $\alpha = .644$
- T2\_F\_FIQChild:  $N = 103$ ;  $M = 30.32$ ;  $SD = 3.50$ ; range = 18-36;  $\alpha = .659$
- T2\_F\_FIQMother:  $N = 104$ ;  $M = 23.12$ ;  $SD = 1.53$ ; range = 12-24;  $\alpha = .659$
- T2\_F\_FIQComm:  $N = 104$ ;  $M = 13.62$ ;  $SD = 3.43$ ; range = 5-19;  $\alpha = .683$
- T2\_FIQMother\_Diff:  $N = 100$ ;  $M = 3.64$ ;  $SD = 2.73$ ; range = 0-11
- T2\_FIQChild\_Diff:  $N = 100$ ;  $M = 4.44$ ;  $SD = 3.44$ ; range = 0-16
- T2\_FIQComm\_Diff:  $N = 99$ ;  $M = 2.95$ ;  $SD = 2.36$ ; range = 0-10
- T2\_FIQTot\_Diff:  $N = 96$ ;  $M = 7.95$ ;  $SD = 5.88$ ; range = 0-25

### General Functioning Subscale

GF6+; Boterhoven de Haan et al., 2015

This subscale of the General Functioning scale contains 6 items on a 4-point Likert scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). The total score is calculated by averaging all items and can range between 1 and 4, with higher scores indicating better family functioning. This data was collected from the participating fathers at T1 and T2, and the participating mothers at T3.

T	Person who completed the measure:			Reporting on:		
	Mother	Father	Child	Mother	Father	Child
3	x			family functioning		
1,2	x			family functioning		

Variables:

- T1\_F\_GF6:  $N = 156$ ;  $M = 3.67$ ;  $SD = .37$ ; range = 2-4;  $\alpha = .721$
- T2\_F\_GF6:  $N = 101$ ;  $M = 3.75$ ;  $SD = .37$ ; range = 2-4;  $\alpha = .734$
- T3\_M\_GF6:  $N = 99$ ;  $M = 1.77$ ;  $SD = .48$ ; range = 1-3.17;  $\alpha = .860$

### Alabama Parenting Questionnaire – short form

APQ; Elgar et al., 2007

9 item measure of parenting styles in three contexts: positive parenting, inconsistent discipline, and poor supervision. The APQ uses a 5-point Likert scale ranging from 1 (*never*) to 5 (*always*). Items 1, 6, and 7 are summed to calculate the positive parenting score (APQPositivePar), items 2, 4, and 9 are summed for the inconsistent discipline total score (APQDisc), and items 3, 5, and 8 are summed for the poor supervision score (APQSuper). The total scores from all three subscales can range between 3 and 15. This data was collected from the participating mothers at T2.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
x			x		

Variables:

- T2M\_APQPositivePar:  $N = 290$ ;  $M = 13.91$ ;  $SD = 1.52$ ; range = 7-15;  $\alpha = .629$
- T2M\_APQDisc:  $N = 289$ ;  $M = 10.27$ ;  $SD = 2.43$ ; range = 3-15;  $\alpha = .418$
- T2M\_APQSuper:  $N = 289$ ;  $M = 4.34$ ;  $SD = 1.99$ ; range = 3-12;  $\alpha = .580$

### Parental Self-Agency Measure – positive items

PSA; Dumka et al., 1996

5 positive items measuring maternal self-efficacy on a 7-point Likert scale from 1 (*rarely*) to 7 (*always*). The total score is calculated by summing all items and can range between 5 and 35, with higher scores indicating that the mothers felt they had more self-agency as a parent. This data was collected from the participating mothers at T2.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>x</b>		

Variables:

- T2M\_PSATot:  $N = 291$ ;  $M = 30.70$ ;  $SD = 4.13$ ; range = 11-35;  $\alpha = .780$

### Parental Burnout Assessment

PBA; Roskam et al., 2018

23 items measuring how often the participants have negative feelings or experiences in relation to their children, on a 7-point Likert scale ranging from 0 (*never*) to 6 (*every day*). All items are summed for the total PBA score which can range from 0 to 138, with higher scores indicating higher levels of parental burnout. This data was collected from the participating fathers at T2.

Given that mother and child data was collected in the same room and at the same time by two fieldworkers, although we wanted to implement the PBA to mothers, many of the questions are highly negative about parenting and we did not think that it would be acceptable to ask some of these questions in the study child's presence. Consequently, as a research team, we chose 4 selected items from the PBA which would not be an issue if the child overheard, piloted these with mothers before T3, had discussions among the fieldworkers and other research team members, and ultimately administered these 4 items to the participating T3 mothers. This selected PBA measure can range from 0 to 28, and higher scores similarly indicate higher levels of burnout.

T	Person who completed the measure:			Reporting on:		
	Mother	Father	Child	Mother	Father	Child
<b>3</b>	<b>x</b>			<b>x</b>		
<b>2</b>		<b>x</b>			<b>x</b>	

Variables:

- T2\_F\_PBA:  $N = 93$ ;  $M = 18.03$ ;  $SD = 18.74$ ; range = 0-104;  $\alpha = .881$
- T3\_M\_PBA:  $N = 99$ ;  $M = 14.71$ ;  $SD = 5.94$ ; range = 0-24;  $\alpha = .754$

## Arabic Experience in Close Relationships Scale - Revised 16

ECRS-R-16, Kazarian & Taher, 2012

The ECRS includes 16 items on a 7-point Likert scale ranging from 1 (*strongly agree*) to 7 (*strongly disagree*). Seven positive items (items 1-7) are reverse scored. The total score is then calculated by summing the items and can range from 16 to 112, with higher scores indicating higher anxious and avoidant attachment. This measure was collected from the mothers at T3.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
x			x		

Variables:

- T3\_ECRS:  $N = 90$ ;  $M = 77.81$ ;  $SD = 15.76$ ; range = 29-112;  $\alpha = .846$

## 4. Other

### Household Wealth Index

Panter-Brick et al., 2009, 2018

12 *yes/no* items measuring relative poverty. In this measure, participants report on their households having the following items: TV, satellite, smartphone, car, refrigerator, computer, oven with gas, bedframe (not only a mattress), washing machine, heater, fan, and water heater. This is a checklist measure of locally relevant household wealth and is calculated by summing all items. The total scores can range from 0 to 12, with lower scores indicating that the households were more impoverished.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
x			household		

Variables:

- T1M\_HouseholdWealth:  $N = 297$ ;  $M = 7.96$ ;  $SD = 1.85$ ; range = 3-12

### COVID-19 Household Environment Scale – Part 2

CHES; Behar-Zusman et al., 2020

This is a 29-item, 5-point Likert scale measure of the impacts of the COVID-19 pandemic on family functioning, with answers ranging from 1 (*much less than before*) to 5 (*much more than before*). It consists of two subsections. The negative items (1-15) measure changes in household conflict as compared to pre-COVID. The negative items are summed for the total conflict score, which can range between 15 and 75, with higher scores indicating more household conflict than before the COVID-10 pandemic. The family conflict measure was collected from the participating fathers at T2.

The positive items (17-30) measure the change in family cohesion ('togetherness') as compared to pre-COVID (1 item on physical intimacy was removed from the original subscale in this population). The total togetherness score is calculated by summing the positive items and can range between 14 and 70, with higher scores indicating better family cohesion than



before the COVID-19 pandemic. The data from the cohesion subscale was collected from the fathers at T1 and T2.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>change in family functioning</b>		

Variables:

- T1\_F\_CHES\_cohesion:  $N = 122$ ;  $M = 54.49$ ;  $SD = 6.96$ ; range = 31-69;  $\alpha = .802$
- T2\_F\_CHES\_cohesion:  $N = 75$ ;  $M = 52.52$ ;  $SD = 6.32$ ; range = 34-64;  $\alpha = .754$
- T2\_F\_CHES\_conflict:  $N = 52$ ;  $M = 44.50$ ;  $SD = 4.25$ ; range = 33-56;  $\alpha = .749$

### Traumatic Events Checklist

TEC, Panter-Brick et al., 2009

20 *yes/no* checklist questions assessing exposure to traumatic events related to war. The total TEC score is calculated by summing all items. It can range from 0 to 20, with higher scores indicating exposure to a larger number of traumatic events. We implemented this measure in T2, but then it became clear that answering the questions was unduly distressing to mothers, so we replaced it with the less distressing PTSD and wellbeing questionnaires.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>x</b>		

Variables:

- T2M\_TEC:  $N = 133$ ;  $M = 7.52$ ;  $SD = 4.69$ ; range = 0-18

### Collective Efficacy Scale

Sampson et al., 1997

10 items on a 5-point Likert scale measure social cohesion amongst neighbours and include two sections. The *informal social control* section includes items 1-5 with answers ranging from 1 (*very likely*) to 5 (*very unlikely*), which assess the likelihood of the neighbours intervening when there is trouble. The *social cohesion and trust* section include items 6-10 with answers ranging from 1 (*strongly agree*) to 5 (*strongly disagree*), which assess how likely the neighbours are to support each other. Two negative items (questions 9 and 10) are reverse coded, and all items are then averaged to calculate the total score, which can range from 1 to 5, with higher scores indicating higher social cohesion. This data was collected from the mothers at T3.

Some of the participating mothers ( $n = 32$ ) were asked an incorrect question in the Collective Efficacy Scale (ColLES) and were therefore excluded from the ‘complete’ measure (T3\_ColLES\_complete). To incorporate the answers from all participating mothers, an adjusted variable with 9 items (excluding item 5) was also calculated (T3\_ColLES\_complete).

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>neighbourhood social cohesion</b>		

Variables:

- T3\_CollES\_complete:  $N = 67$ ;  $M = 2.74$ ;  $SD = .67$ ; range = 1.5-4.4;  $\alpha = .774$
- T3\_CollES\_incomplete:  $N = 98$ ;  $M = 2.68$ ;  $SD = .68$ ; range = 1.44-4.56;  $\alpha = .778$

### Outgroup Attitudes Scale

Ozkan et al., 2021, using items adapted from Bagci et al., 2018

2 items measuring perceived discrimination of Syrian refugees in Jordan on a 7-point Likert scale ranging from 1 (*not at all*) to 7 (*all the time*). The two items are summed, and the total score can range from 2 to 14, with higher scores indicating more perceived discrimination. This data was collected from the mothers at T3.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>x</b>		

Variables:

- T3\_OAS\_1:  $N = 99$ ;  $M = 7.16$ ;  $SD = 3.87$ ; range = 2-14;  $\alpha = .738$

### Short Form Spirituality-S-4

ISS, Kira et al., 2021

Selected questions used in this population were 4 items on a 4-point Likert scale ranging from 1 (*not true about me*) to 4 (*mostly true about me*). The total ISS score is calculated by summing all questions and can range from 4 to 16, with higher scores indicating a higher level of spirituality. This data about the mothers' religious belief was collected from the participating mothers at T3.

Person who completed the measure:			Reporting on:		
Mother	Father	Child	Mother	Father	Child
<b>x</b>			<b>x</b>		

Variables:

- T3\_SFS:  $N = 99$ ;  $M = 15.57$ ;  $SD = .88$ ; range = 11-16;  $\alpha = .507$

## **5. Cognitive tasks**

The data from the following cognitive tasks has not yet been coded and is not yet part of the database. Once the coding of these tasks is finalised, the data will be included in the database and the guide will be updated.

Unless specified otherwise, all cognitive tasks were administered using Matlab (Mathworks) and Psychtoolbox-3 (Brainard, 1997).

### Go/NoGo

Adapted from Hare et al. (2005)

This task measures affective cognitive inhibition and was completed independently by mothers and children at T1. Over 120 trials, the participants were briefly presented with images of different actors expressing an emotion (anger, sadness, neutral expression). On a given block, the participants were presented with target ('no-go') emotional stimuli amongst neutral stimuli and vice versa (i.e., target neutral stimuli amongst emotional stimuli). The participants' goal was to press 'spacebar' on the target trials and withhold their response on the non-target trials, whilst their performance and reaction times were measured, resulting in measures of cognitive inhibition to angry and sad emotional expressions.

Person who completed the measure:		
Mother	Father	Child
x		x

### Dot probe

Adapted from Mcleod et al. (1986)

This task measures affective attention bias and was completed independently by the children at T1 and the mothers at T2. Over 100 trials, the participants were briefly presented with an image of an emotional expression (anger, sadness) on either the left or the right side of the screen, paired with an image of a neutral expression on the opposite side of the screen. The stimuli were then followed by an image of a coin (probe) appearing behind the emotional face or the neutral face. The participants were asked to indicate the location of the coin by pressing a corresponding key on the keyboard as quickly and as accurately as possible. During the task we measured the reaction time and accuracy to assess their attention bias to angry and sad emotional expressions.

Person who completed the measure:		
Mother	Father	Child
x		x

### Simon task

Adapted from Davidson et al. (2006)

This task measures cognitive inhibition and was completed by the children at T2. Over 120 trials, the children were presented with a single dot (either pink or yellow) on the left or right side of the screen and were instructed to indicate the position of the dot by pressing the left or right arrow key on the keyboard as quickly and accurately as possible, depending on the rules explained by the fieldworker, which differed for each block. The rules depended on the colour of the dot, e.g., children were asked to respond 'same side' if the dot was pink, and 'opposite side' if the dot was yellow. We measured the children's accuracy and reaction times to assess different aspects of their cognitive inhibition.

Person who completed the measure:		
Mother	Father	Child
		x

### Trust task

Adapted from Neil et al. (2022)

This task measures the perception of trustworthiness and was completed by the children at T2. The children were presented with different-identities computer generated faces (with neutral expressions) that varied in the level of trustworthiness (computationally modelled dimension of trustworthiness). The children were requested to indicate if each of the faces was trustworthy or untrustworthy, and their responses were then plotted to extract their trustworthiness bias (the level of trustworthiness at which the child was equally likely to respond 'trustworthy' or 'not trustworthy', i.e., the face trust level leading to 0.5 proportion of trustworthy responses) and sensitivity (the ability to detect changes in the level of trustworthiness of the stimulus).

Person who completed the measure:		
Mother	Father	Child
		x

### Sharing tasks

Adapted from Fehr et al. (2008)

These tasks measure the willingness to share with others and were completed by the children at T2 and T3. The data from these tasks was input using KoBoToolbox.

Sharing task 1. In this task, the children are presented with a number of sweets and different possibilities to share them with other children that they know ('in-group') or do not know ('out-group'), and who are not present during the task. The children were asked to make a decision about how many sweets they want to receive and how many they want to share with the other child. Depending on the block, they can select from the following options: 1) both children (the study child and 'the other child') receive one sweet; 2) the study child receives two sweets and the other child receives none; 3) the study child receives one sweet and the other child receives none; or 4) the study child receives no sweets and the other child receives two. This task was completed by the children at T2.

Sharing task 2 (Envelope task). In this task, the children are told they are going to receive 4 sweets and that they can share them with another child who we would not have the time to test. They can distribute their sweets however they want (i.e., they have the freedom to decide how many sweets to keep for themselves and how many to share with another child) and place them in an envelope so that the fieldworkers cannot see how many sweets each child decided to share. The results of this task are the number of sweets the children decided to share (0-4), with higher score indicating higher tendency to share. This task was completed by the children at T3.

Person who completed the measure:		
Mother	Father	Child
		x

### Raven's Progressive Matrices

#### Raven & Raven (2004)

This task measures reasoning and was completed by children at T3. This non-verbal task consists of viewing geometric designs with part of the design missing. The children are asked to select the missing pattern from a set of design options, with the task becoming progressively difficult. The performance is measured by the number of correct trials (and/or when the patterns stop if the child is no longer able to continue).

Person who completed the measure:		
Mother	Father	Child
		<b>x</b>

### Delayed gratification

#### Adapted from Duckworth & Seligman (2005)

This task is a measure of delayed gratification and was completed by the children at T3 with the use of KoBoToolbox. The children were informed that they will receive sweets as part of the task and were asked to select one of the options: they can receive one sweet right now, or they can wait until the end of the data collection that day and receive two sweets. By measuring children's responses (*wait/no wait*) we assess their ability to delay gratification and wait for a better reward.

Person who completed the measure:		
Mother	Father	Child
		<b>x</b>

### Number recall (Digit Span Task)

#### Adapted from Davis & Pratt (1995)

This task is a measure of working memory and was completed by children at T3 with the use of KoBoToolbox. The children were asked to repeat a sequence of digits with increasing difficulty levels (increasing number of digits in a sequence, starting with 2 and increasing to 9 digits). This task measures children's ability to keep certain amount of information in their minds and use it according to the instructions.

Person who completed the measure:		
Mother	Father	Child
		<b>x</b>

### Continuous Performance Task (CPT, Space Adventure)

#### Adapted from Klee & Garflinkel (1983)

This task measures sustained attention and was completed by the children at T3. At each trial the children were presented with an image of a different-coloured spaceship. The children were instructed to look out for the red spaceship (target), and their goal was to press 'spacebar' on target trials and withhold their response on non-target trials (other colours). Using their accuracy score and reaction times in this task, we measured children's ability to sustain their attention for a prolonged period of time.

Person who completed the measure:		
Mother	Father	Child
		<b>x</b>

## 6. Observational tasks

The data from the following observational tasks has not yet been coded and is not yet part of the database. Once the coding of these tasks is finalised, the data will be included in the database and the guide will be updated.

### Mother-child observational tasks

The participating mothers and/or children were recorded during the observational tasks and the videos will be coded based on a manual adapted from the Coding of Attachment Related Parenting (CARP; Matias, 2006) and developed based on the data gathered for the videos. All observational tasks are coded in terms of three main domains: sensitive responding, positive mutuality, and disrupted responsiveness.

#### Shared book reading

This task involves the mother and the child reading book together and measures the mother-child relationship quality. Participants could do this task for up to 10 minutes. This data was collected at T1 and T2.

Person who completed the measure:		
Mother	Father	Child
<b>x</b>		<b>x</b>

#### Free play

This task involves the participating mother and child playing together freely with two sets of toys given to them by the fieldworker. Participants could do this task for up to 5 minutes. It is used to measure mother-child relationship quality. This data was collected at T1 and T2.

Person who completed the measure:		
Mother	Father	Child
<b>x</b>		<b>x</b>

### Puzzle

This task measures child task perseverance (grit). Children could do this task for up to 5 minutes. It involves the child being asked to build 7 pre-specified objects using Duplo pieces, whilst being video recorded. Children were asked to complete as many of the objects as possible. We will code various behaviours related to how the child engages with and solves the puzzle. This data was collected at T3.

Person who completed the measure:		
Mother	Father	Child
		<b>x</b>

**This research was funded by an Education and Learning in Crises award from the British Academy (ERICC\190091).**

[illegible]



Interfaith Spirituality Scale	Religiosity	4							x <sup>†</sup>	
<i>Cognitive tasks</i>										
Go/NoGo	Affective cognitive inhibition		x		x					
Dot probe	Affective attention				x	x				
Simon	Cognitive control (inhibition & flexibility)							x		
Trust	Perception of trustworthiness							x		
Sharing	Willingness to share with others							x		x
Raven's progressive matrices	Cognitive and spatial skills									x
Delayed gratification	Delayed gratification									x
Number recall	Working memory									x
Continuous Performance Task	Sustained attention									x
<i>Observational tasks</i>										
Shared book reading (10 min)	Mother-child relationship quality		x		x	x		x		
Free play (5 min)	Mother-child relationship quality		x		x	x		x		
Puzzle (5 min)	Task perseverance (grit)									x

**Table 1. All variables collected at the three timepoints**

HALDO = Holistic Assessment of Learning and Development Outcomes, PPCATR = Parents' perceived child attitudes towards reading, DASS = Depression Anxiety Stress Scales, CES-D = Centre for Epidemiologic Studies - Depression Scale, PTSD Checklist for DSM-5 = PCL-5, WEMWBS-S = The Warwick-Edinburgh Mental Wellbeing Scale, short form. The COVID-19 Household Environment Scale has a togetherness subscale and a conflict subscale; they are indicated above with a + and a – symbol, respectively.

\* We had been using the Trauma Events Checklist with the mothers in T2, but some of the mothers found this questionnaire quite distressing to complete. We tried a few things to make it less distressing but it continued to be difficult for them to complete and so about halfway through T2, we stopped using this and instead added the PCL-5 and the WEMWBS-S for the mothers.

<sup>†</sup> Select items only.

## **REFERENCES**

- Bagci, S. C., Piyale, Z. E., Bircek, N. I., & Ebcim, E. (2018). Think beyond contact: Reformulating imagined intergroup contact theory by adding friendship potential. *Group Processes & Intergroup Relations*, 21(7), 1034–1052.
- Behar-Zusman, V., Chavez, J. V., & Gattamorta, K. (2020). Developing a Measure of the Impact of COVID-19 Social Distancing on Household Conflict and Cohesion. *Family process*, 59(3), 1045–1059.
- Blevins, C., Weathers, F., Davis, M., Witte, T., & Domino, J. (2015). The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and Initial Psychometric Evaluation. *Journal of Traumatic Stress*, 28, 489–498.
- Boterhoven de Haan, K. L., Hafekost, J., Lawrence, D., Sawyer, M. G., & Zubrick, S. R. (2015). Reliability and validity of a short version of the general functioning subscale of the McMaster Family Assessment Device. *Family process*, 54(1), 116–123.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.
- Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology*, 87(1), 154-167.
- Coddington, C. S., & Guthrie, J. T. (2009). Teacher and student perceptions of boys' and girls' reading motivation. *Reading Psychology*, 30(3), 225-249.
- D'Sa, N., Krupar, A., & Westrope, C. (2019). Feasible measurement of learning in emergencies: lessons from Uganda. *Forced Migration Review*, 60, 74-75.
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44(11), 2037–2078.
- Davis, H.L. and Pratt, C. (1995), The development of children's theory of mind: The working memory explanation. *Australian Journal of Psychology*, 47, 25-31.
- Driscoll, K., & Pianta, R. C. (2011). Mothers' and fathers' perceptions of conflict and closeness in parent-child relationships during early childhood. *Journal of Early Childhood and Infant Psychology*, (7), 1-24.
- Duckworth, A. L., Tsukayama, E., & Kirby, T. A. (2013). Is it really self-control? Examining the predictive power of the delay of gratification task. *Personality and Social Psychology Bulletin*, 39, 843-855.

- Dumka, L. E., Stoerzinger, H. D., Jackson, K. M., & Roosa, M. W. (1996). Examination of the cross-cultural and cross-language equivalence of the parenting self-agency measure. *Family Relations*, 45(2), 216-222.
- Elanbari G. (2015). *Translation and validation of the Arabic-Revised Dyadic Adjustment Scale (A-RDAS)*. [Doctoral thesis, St Mary's University Texas].
- Elgar, F. J., Waschbusch, D. A., Dadds, M. R., & Sigvaldason, N. (2007). Development and Validation of a Short Form of the Alabama Parenting Questionnaire. *Journal of Child and Family Studies*, 16, 243-259.
- Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, 454, 1079–1083.
- Hare T.A., Tottenham N., Davidson M.C., Glover G.H., Casey B.J. (2005) Contributions of amygdala and striatal activity in emotion regulation. *Biological Psychiatry* 57(6), 624–632
- Hein S, Bick J, Issa G, Aoude L, Maalouf C, Awar A, et al. (2020) Maternal perceptions of father involvement among refugee and disadvantaged families in Beirut, Lebanon. *PLoS ONE* 15(3): e0229670.
- Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 44(2), 227–239.
- Jellinek, M. S., Murphy, J. M., Robinson, J., Feins, A., Lamb, S., & Fenton, T. (1988). Pediatric Symptom Checklist: Screening school-age children for psychosocial dysfunction. *The Journal of Pediatrics*, 112(2), 201–209.
- Kazarian, S. S., & Taher, D. (2012). Adult attachment styles in Lebanese community sample: A brief experience in close relationships scale. *The Arab Journal of Psychiatry*, 23(1), 74–79.
- Kessler, R. C., Barker, P. R, Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M.J., Normand, S.L., Manderscheid, R. W., Walters, E. E., Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Arch Gen Psychiatry*, 60(2), 184-9.
- Kira, I., Özcan, N., Shuwiekh, H. , Kucharska, J. , Al-Huwailah, A. & Bujold-Bugeaud, M. (2021) Mental Health Dynamics of Interfaith Spirituality in Believers and Non-Believers: The Two Circuit Pathways Model of Coping with Adversities: Interfaith Spirituality and Will-to Exist, Live and Survive. *Psychology*, 12, 992-1024.
- Klee, S. H., & Garfinkel, B. D. (1983). The computerized continuous performance task: a new measure of inattention. *Journal of abnormal child psychology*, 11(4), 487–495.
- Macleod, C., Mathews, A., & Tata, P. (1986). Attentional Bias in Emotional Disorders. *Journal of Abnormal Psychology*, 95(1), 15–20.

- Matias, S. (2006). *Direct Observation of Parent-Child Interaction based on Attachment Theory*. [Doctoral thesis, King's College London].
- Morin C.M., Belleville G., Bélanger L. & Ivers H. (2011). The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. *SLEEP*, 34(5), 601-608.
- Neil, L., Viding, E., Armbruster-Genc, D., Lisi, M., Mareschal, I., Rankin, G., Sharp, M., Phillips, H., Rapley, J., Martin, P. and McCrory, E. (2022), Trust and childhood maltreatment: evidence of bias in appraisal of unfamiliar faces. *Journal of Child Psychology and Psychiatry*, 63, 655-662.
- Özkan, Z., Ergün, N., & Çakal, H. (2021). Positive versus negative contact and refugees' intentions to migrate: The mediating role of perceived discrimination, life satisfaction and identification with the host society among Syrian refugees in Turkey. *J Community Appl Soc Psychol*, 31, 438– 451.
- Ozturk, G., & Ohi, S. (2018). Understanding young children's attitudes towards reading in relation to their digital literacy activities at home. *Journal of Early Childhood Research*, 16(4), 393–406.
- 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.
- Panter- Brick, C., Hadfield, K., Dajani, R., Eggerman, M., Ager, A., & Ungar, M. (2018). Resilience in context: A brief and culturally grounded measure for Syrian refugee and Jordanian host- community adolescents. *Child Development*, 89(5), 1803-1820.
- Pianta, R. C. (1992), *Child-parent relationship scale* (Unpublished measure, University of Virginia)
- Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, 1(3), 385–401.
- Raven, J., & Raven, J. (2003). Raven Progressive Matrices. In R. S. McCallum (Ed.), *Handbook of nonverbal assessment*. Kluwer Academic/Plenum Publishers.
- Roskam, I., Brianda, M. E., & Mikolajczak, M. (2018). A Step Forward in the Conceptualization and Measurement of Parental Burnout: The Parental Burnout Assessment (PBA). *Frontiers in psychology*, 9, 758.
- Sampson, R.J., Raudenbush, S.W., Earls, F. (1997) Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy, *Science* 277(5328), 918-924.
- Schutte, N.S., & Malouff, J.M. (2007) Dimensions of Reading Motivation: Development of an Adult Reading Motivation Scale, *Reading Psychology*, 28:5, 469-489.
- Spanier, G.B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family*. 38, 15–28.

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., *et al.* (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes*, 5, 63.

Ziadni, M., Hammoudeh, W., Rmeileh, N. M. E. A., Hogan, D., Shannon, H., & Giacaman, R. (2011). Sources of Human Insecurity in Post-War Situations: The Case of Gaza. *Journal of Human Security*, 7(3).