Results for Postnatal RCT

A sample of 61 postpartum women was recruited, with 29 randomly assigned to the intervention group and 32 to the control group.

Two observations (PMH047AP and 0405AP) had missing R-DAS pre-test scores. Three observations (PMH001PA, PMH424AP, and 1041AP) had missing post-test scores for EPDS, GAD-7, and R-DAS. One observation (2115AP) had missing post-test scores for GAD-7 and R-DAS.

To address this an imputation strategy helped ensure that the missing data did not undermine the validity of the analysis and preserved statistical integrity.

The data indicates that overall, both groups displayed comparable distributions across most demographic characteristics, with slight variations in specific categories.

With respect to nationality, the majority of participants in both groups were Maltese nationals, with a slightly higher proportion in the intervention group (89.7% vs. 81.3%). The Fisher's Exact Test ($p = 0.478$) indicated no significant difference between groups in terms of nationality distribution.

In terms of relationship status, most participants were married or in a registered partnership (65.6% in the control group vs. 65.5% in the intervention group), while similar percentages were observed for cohabiting participants (21.9% vs. 27.6%). Other relationship categories had only a few participants, contributing to the lack of significant differences ($\chi2(4) = 3.361$, $p = 0.499$).

Regarding education level, a majority in both groups had completed tertiary education (62.5% in control vs. 62.1% in intervention), while smaller proportions had post-secondary (28.1% vs. 34.5%) or secondary education (9.4% vs. 3.4%). The slight variations were not statistically significant ($\chi2(2) = 1.013, p = 0.603$). Labour status was also largely similar across groups, with most participants being employed or self-employed (84.4% in control vs. 93.1% in intervention). A small number were unemployed (6.3% vs. 3.4%) or taking care of the house and/or family (9.4% vs. 3.4%), with no statistically significant differences ($\chi2(2) = 1.189$, $p = 0.552$).

The impact of COVID-19 showed some variability, though it did not reach statistical significance ($\chi2(2) = 4.568$, $p = 0.102$). A higher percentage of participants in the control group reported saving more (46.9% vs. 41.4%) or struggling to make ends meet (31.3% vs. 13.8%), whereas the intervention group had a larger proportion reporting that they saved less (44.8% vs. 21.9%). Regarding psychotropic medication, a higher percentage of participants in the control group were taking such medication (21.9% vs. 6.9%). However, the Fisher's Exact Test ($p = 0.151$) indicated that this difference was not statistically significant, suggesting that medication usage was comparable between groups.

Overall, while minor variations were observed, both groups showed comparable distributions across most demographic and lifestyle characteristics. The Chi-Square Test was used to compare categorical variables with multiple levels, while Fisher's Exact Test was applied for binary variables where expected counts were low (i.e., Nationality and Psychotropic Medication). These results suggest that randomization was effective in balancing baseline characteristics between the control and intervention groups.

Analysis 1: Differences between pre-post-measures for the control and intervention groups separately

This section investigates the differences in scores for the Edinburgh Postnatal Depression Scale (EPDS), Generalized Anxiety Disorder-7 (GAD-7), and the Revised Dyadic Adjustment Scale (R-DAS) between pre- and post-intervention assessments, analyzed separately for the control and intervention groups. These measures were selected to capture key aspects of psychological and relational well-being during the study period. Whereas participants needed to score below the cut-off point either on the EPDS or the GAD7 or on both to be eligible for the study, the R-DAS was used as a criterion of relational wellbeing between the parents, which in turn would reflect on the quality of their parenting relationship. Participants were eligible for the study irrespective of their score on the RDAS.

To assess changes within each group over time, the paired t-test was employed. This statistical test evaluates whether the mean difference between pre- and post-measurements within a group is significantly different from zero, assuming the data are approximately normally distributed. Using the paired t-test, this analysis aims to:

- Determine whether significant changes occurred within each group (control and intervention) for EPDS, GAD-7, and R-DAS.
- Highlight the magnitude and direction of changes in mental health and relationship quality scores for each measure.

By analysing the pre- and post-scores separately for the control and intervention groups, this study provides insight into the effectiveness of the intervention and any observed patterns of change within each group.

- **Control Group**

The changes in scores for the **EPDS**, **GAD-7**, and **R-DAS** were analysed separately for the control group using paired t-tests. The table below summarizes the mean pre- and post-intervention scores, the change in scores, and the 95% confidence intervals (CIs) for the change in scores:

	Pre	Post	Change of score	95% CI of Change of Scores
EPDS	13.87 (4.556)	9.28 (5.419)	-4.59 (5.260)	[-6.49, -2.70]
GAD-7	10.19 (4.060)	6.56 (5.382)	-3.63 (6.814)	[-6.08, -1.17]
R-DAS	49.47 (9.899)	49.56 (9.837)	0.094 (6.850)	[-2.38,2.56]

Table 1: Pre- and post-intervention mean scores, change in scores, and 95% Confidence Intervals for the control

This table highlights the mean changes in scores for the control group across the three measures, along with the associated variability and confidence intervals. Below are the results of the paired t-tests conducted to assess differences in pre- and post-measurements for the control group across the EPDS, GAD-7, and R-DAS.

- **EPDS**

t(31) = 4.940, p-value < 0.001, Cohen's d = 0.873, 95% CI for Cohen's d [0.460, 1.277]

The t-test result was t(31) = 4.940, p-value < 0.001, indicating a statistically significant reduction in depressive symptoms post-intervention. The Cohen's d value of 0.873 suggests a large effect size, meaning the control had a substantial impact on reducing depressive symptoms. The 95% CI for Cohen's d is [0.460, 1.277], further supporting that the effect is large and meaningful.

- **GAD-7**

t(31) = 3.009, p-value = 0.003, Cohen's d = 0.532, 95% CI for Cohen's d [0.157, 0.899]

The t-test result was t(31) = 3.009, p-value = 0.003, indicating a statistically significant reduction in anxiety symptoms in the control group post-intervention. The Cohen's d value of 0.532 suggests a moderate effect size, meaning the reduction in anxiety symptoms was of moderate magnitude. The 95% CI for Cohen's d is [0.157, 0.899], further supporting that the effect size is moderate and significant.

- **R-DAS**

t(31) = -0.077, p-value = 0.939, Cohen's d = -0.014, 95% CI for Cohen's d [-0.360, 0.333]

The t-test result was t(31) = -0.077, p-value = 0.939, indicating that there was no statistically significant change in R-DAS scores in the control group post-intervention. The Cohen's d value of -0.014 suggests an extremely small effect size, meaning that the intervention had virtually no impact. The 95% CI for Cohen's d is [-0.360, 0.333], which includes zero, further confirming that there was no meaningful effect.

- **Intervention Group**

The changes in scores for the **EPDS, GAD-7**, and **R-DAS** were analysed separately for the intervention group using paired t-tests. The table below summarizes the mean pre- and post-intervention scores, the change in scores, and the 95% confidence intervals (CIs) for the change in scores:

	Pre	Post	Change of score	95% CI of Change of Scores
EPDS	12.97 (3.469)	8.90 (3.697)	-4.07 (4.391)	[-5.74, -2.40]
GAD-7	11.69 (3.901)	7.86 (4.926)	-3.83 (4.907)	[-5.69, -1.96]
R-DAS	48.14 (10.589)	48.97 (10.940)	0.83 (8.320)	[-2.34, 3.99]

Table 2: Pre- and post-intervention mean scores, change in scores, and 95% Confidence Intervals for the intervention

This table highlights the mean changes in scores for the intervention group across the three measures, along with the associated variability and confidence intervals. Below are the results of the paired t-tests conducted to assess differences in pre- and post-measurements for the intervention group across the EPDS, GAD-7, and R-DAS.

- **EPDS**

t(28) = 4.990, p-value < 0.001, Cohen's d = 0.927, 95% CI for Cohen's d [0.484, 1.358]

The t-test result was t(28) = 4.990, p-value < 0.001, indicating a statistically significant reduction in EPDS scores in the intervention group post-intervention. The Cohen's d value of 0.927 suggests a large effect size, meaning the intervention had a substantial impact on reducing depressive symptoms. The 95% CI for Cohen's d is [0.484, 1.358], which confirms that the effect size is large, as the confidence interval does not include zero and falls within a range indicative of a meaningful reduction in symptoms.

- **GAD-7**

t(28) = 4.201, p-value < 0.001, Cohen's d = 0.780, 95% CI for Cohen's d [0.358, 1.192]

The t-test result was t(28) = 4.201, p-value < 0.001, indicating a statistically significant reduction in GAD-7 scores in the intervention group post-intervention. The Cohen's d value of 0.780 suggests a large effect size, meaning the intervention had a substantial impact on reducing anxiety symptoms. The 95% CI for Cohen's d is [0.358, 1.192], which further supports the large effect size, as the confidence interval does not include zero and reflects a meaningful reduction in anxiety symptoms.

- **R-DAS**

t(28) = -0.536, p-value = 0.596, Cohen's d = -0.099, 95% CI for Cohen's d [-0.463, 0.266]

The t-test result was t(28) = -0.536, p-value = 0.596, indicating that there was no statistically significant change in R-DAS scores in the intervention group post-intervention. The Cohen's d value of -0.099 suggests an extremely small effect size, meaning that the intervention had no impact on the scores. The 95% CI for Cohen's d is [-0.463, 0.266], which includes zero, confirming that the effect is negligible and that there was no meaningful change in R-DAS scores.

Conclusion

In conclusion, this analysis demonstrated that both the intervention and control groups experienced significant reductions in depressive symptoms (EPDS) and anxiety symptoms (GAD-7) from pre- to post-intervention.

- EPDS: Large effect sizes were observed in both groups (d = 0.873 in control, d = 0.927 in intervention), indicating a substantial reduction in depressive symptoms regardless of group allocation.
- GAD-7: The effect sizes were moderate in the control group (d = 0.532) and large in the intervention group (d = 0.780), suggesting meaningful reductions in anxiety symptoms in both conditions, with a somewhat stronger effect in the intervention group.
- R-DAS: No significant changes in relationship distress were observed in either group, with negligible effect sizes, indicating that the intervention did not impact relationship quality.

Of note, 21.9% of the control group was on psychotropic medication compared to only 6.9% in the intervention group. While Fisher's Exact Test indicated no significant difference in psychotropic medication use between groups ($p = 0.151$), the control group had a higher proportion of participants taking medication (21.9%) compared to the intervention group (6.9%). Given the small sample size, it is possible that medication use may have contributed to symptom reductions, particularly in the control group. However, the intervention group, which had fewer medicated participants, also showed significant reductions in EPDS and GAD-7 scores, suggesting that the intervention itself played a key role in improving symptoms. It is important to note that relationship distress (as measured by R-DAS) was not a requirement for study participation. Consequently, only 13 out of 31 participants in the intervention group and 8 out of 30 in the control group were classified as distressed on the R-DAS at baseline. In contrast, higher proportions of participants met clinical thresholds for anxiety (GAD-7: 20/31 in IG, 23/30 in CG) and depression (EPDS: 22/31 in IG, 19/30 in CG). Given that many participants already reported healthy relationship dynamics, their scores may have been close to a ceiling effect, leaving little room for measurable improvement. This may explain why no significant changes were observed in R-DAS scores, despite reductions in anxiety and depressive symptoms.

In conclusion, the analysis revealed that both the intervention and control groups showed significant reductions in EPDS and GAD-7 scores, indicating improvements in depressive and anxiety symptoms, respectively. Both groups exhibited large effect sizes for EPDS and moderate effect sizes for GAD-7, indicating meaningful improvements in these areas. Of note is the fact that 21.9% were on psychotropic medication in the Control Group as opposed to 6.9% in the Intervention Group. No significant changes were observed in R-DAS scores for either group, with extremely small effect sizes, indicating that the intervention did not impact relationship distress. This result may be attributed to the fact that most mothers in both groups did not report relationship distress at baseline—16 out of 29 in the Intervention Group and 22 out of 31 in the Control Group. As a result, their potential for improvement in R-DAS scores was limited, likely due to a ceiling effect, where participants with already healthy relationships had little room for measurable change.

Analysis 2: Differences between the intervention and control groups, controlling for time.

This analysis examines whether there are significant differences in EPDS, GAD-7, and R-DAS scores between the intervention and control groups, while accounting for the effect of time (pre- and post-intervention). A repeated-measures ANOVA is employed to control for time and assess group differences in change scores over time.

The primary objectives of the analysis are as follows:

- To determine whether the intervention group shows greater improvement in EPDS, GAD-7, and R-DAS scores compared to the control group.
- To evaluate the interaction effect of group (intervention vs. control) and time (pre vs. post), assessing whether the change in scores over time differs significantly between the groups.

Additionally, the analysis includes the calculation of effect size (Partial Eta Squared) to quantify the magnitude of the group differences while controlling for time. Confidence intervals for the mean change scores are provided to assess the precision of the observed differences. This comprehensive approach enables the evaluation of the intervention's effectiveness while accounting for natural variations over time and baseline differences between groups.

The table presents the pre- and post-intervention scores for the EPDS, GAD-7, and R-DAS scales in both the control and intervention groups, along with the change in scores and the corresponding 95% confidence intervals (CIs) for these changes.

Control group				
	Pre	**Post**	**Change of score**	**95% CI of Change of Scores**
EPDS	13.87 (4.556)	9.28 (5.419)	-4.59 (5.260)	[-6.49, -2.70]
GAD-7	10.19 (4.060)	6.56 (5.382)	-3.63 (6.814)	[-6.08, -1.17]
R-DAS	49.47 (9.899)	49.56 (9.837)	0.094 (6.850)	[-2.38,2.56]
Intervention group				
	Pre	**Post**	**Change of score**	**95% CI of Change of Scores**
EPDS	12.97 (3.469)	8.90 (3.697)	-4.07 (4.391)	[-5.74, -2.40]
GAD-7	11.69 (3.901)	7.86 (4.926)	-3.83 (4.907)	[-5.69, -1.96]
R-DAS	48.14 (10.589)	48.97 (10.940)	0.83 (8.320)	[-2.34, 3.99]

Table 3: Pre- and post- intervention scores for EPDS, GAD-7, and R-DAS in control and intervention groups

- EPDS

In the control group:

- Pre-Intervention: The mean EPDS score was 13.87, with a standard deviation of 4.556, indicating moderate depressive symptoms with moderate variability among participants.
- Post-Intervention: The mean EPDS score dropped to 9.28, with a standard deviation of 5.419, suggesting a reduction in depressive symptoms, though some variability remains in the outcomes.
- The mean reduction was -4.59, with a standard deviation of 5.260, reflecting considerable variability in how much depressive symptoms improved across participants.

In the intervention group:

- Pre-Intervention: The mean EPDS score was 12.97, with a standard deviation of 3.469, indicating moderate depressive symptoms with significant variability among participants.
- Post-Intervention: The mean EPDS score dropped to 8.90, with a standard deviation of 3.697, suggesting a reduction in depressive symptoms, although variability in the outcomes still exists.
- The mean reduction was -4.07, with a standard deviation of 4.391, showing a considerable range in how much depressive symptoms improved.

Comparison Between Groups

The interaction term "**Time * Treatment Group**" examines whether the change in **EPDS** scores over time (pre- to post-intervention) differs between the control and intervention groups:

F(1, 59) = 0.177, p = 0.676, Partial Eta Squared = 0.003, 95% CI for Partial Eta Squared [0,0.081]

- **F(1, 59) = 0.177, p = 0.676**, indicating that the change in EPDS scores over time does not differ significantly between the control and intervention groups.
- The **Partial Eta Squared** value of **0.003** suggests a very small effect size, with minimal practical significance between the groups.
- The **95% Confidence Interval (CI)** for Partial Eta Squared is **[0, 0.081],** confirming that the effect size is very small, as the entire confidence interval is close to zero.

Conclusion: The analysis revealed that both the control and intervention groups showed significant reductions in EPDS scores from pre- to post-intervention, suggesting an improvement

in depressive symptoms. However, the interaction term for time and treatment group did not indicate a significant difference in the change of EPDS scores between the two groups ($F(1, 59) = 0.177$, $p = 0.676$). The effect size was very small, as indicated by the Partial Eta Squared value of 0.003, which suggests minimal practical significance. These findings suggest that, while both groups experienced improvements in depressive symptoms, the intervention did not yield a significantly greater change in EPDS scores compared to the control group over time.

- GAD-7

In the **control** group:

- Pre-Intervention: The mean GAD-7 score was 10.19, with a standard deviation of 4.060, indicating moderate levels of anxiety among participants.
- Post-Intervention: The mean GAD-7 score decreased to 6.56, with a standard deviation of 5.382, suggesting a reduction in anxiety symptoms over time.
- The mean reduction in GAD-7 score was -3.63, with a standard deviation of 6.814, indicating considerable variability in how participants responded.

In the **intervention** group:

- Pre-Intervention: The mean GAD-7 score was 11.69, with a standard deviation of 3.901, indicating moderate levels of anxiety with some variation among participants.
- Post-Intervention: The mean GAD-7 score decreased to 7.86, with a standard deviation of 4.926, showing a notable reduction in anxiety symptoms following the intervention.
- The mean reduction in GAD-7 score was -3.83, with a standard deviation of 4.907, indicating moderate variability in individual responses.

Comparison Between Groups

The interaction term "**Time * Treatment Group**" examines whether the change in **GAD-7** scores over time (pre- to post-intervention) differs between the control and intervention groups:

F(1, 59) = 0.017, p = 0.895, Partial Eta Squared = 0.0003, 95% CI for Partial Eta Squared [0,0.045]

- **F(1, 59) = 0.017**, **p = 0.895**, indicating that the change in GAD-7 scores over time does not differ significantly between the control and intervention groups.
- The **Partial Eta Squared** value of **0.0003** suggests an extremely small effect size, implying that the treatment group (intervention vs. control) has a negligible influence on the change in GAD-7 scores over time.
- The **95% Confidence Interval (CI)** for Partial Eta Squared is **[0, 0.045]**, which further confirms that the effect size is very small, as the entire confidence interval is close to zero.

Conclusion: The analysis of the GAD-7 scores revealed that there were no significant differences in the change of anxiety symptoms over time between the control and intervention groups. The

interaction term for time and treatment group showed no significant effect ($F(1, 59) = 0.017$, $p = 0.895$), with an extremely small effect size (Partial Eta Squared = 0.0003). This suggests that, although both groups experienced reductions in GAD-7 scores, the intervention did not result in a significantly greater change in anxiety symptoms compared to the control group.

- R-DAS

In the **control** group:

- Pre-Intervention: The mean R-DAS score was 49.47, with a standard deviation of 9.899, suggesting a moderate level of relationship quality with some variability among participants.
- Post-Intervention: The mean R-DAS score increased slightly to 49.56, with a standard deviation of 9.837, indicating negligible change in relationship satisfaction.
- The mean change in score was 0.094, with a standard deviation of 6.850, reflecting considerable variability in individual responses.

In the **intervention** group:

- Pre-Intervention: The mean R-DAS score was 48.14, with a standard deviation of 10.589, indicating a similar baseline level of relationship satisfaction compared to the control group, with notable variability among participants.
- Post-Intervention: The mean R-DAS score slightly increased to 48.97, with a standard deviation of 10.940, suggesting minimal improvement in relationship quality.
- The mean change was 0.83, with a standard deviation of 8.320, showing that responses varied widely across individuals.

Comparison Between Groups

The interaction term "**Time * Treatment Group**" examines whether the change in **R-DAS** scores over time (pre- to post-intervention) differs between the control and intervention groups:

F(1, 59) = 0.142, p = 0.707, Partial Eta Squared = 0.002, 95% CI for Partial Eta Squared [0,0.078]

- **F(1, 59) = 0.142, p = 0.707**, indicating that the change in R-DAS scores over time does not differ significantly between the control and intervention groups.
- The **Partial Eta Squared** value of **0.002** suggests an extremely small effect size, implying that the treatment group (intervention vs. control) has a negligible influence on the change in R-DAS scores over time.
- The **95% Confidence Interval (CI)** for Partial Eta Squared is **[0, 0.078]**, which further confirms that the effect size is very small, as the entire confidence interval is close to zero.

Conclusion: The analysis of R-DAS scores indicated that there were no significant differences in the change of distress related to relationships over time between the control and intervention groups. The interaction term for time and treatment group showed no significant effect (F(1, 59) = 0.142, p = 0.707), with an extremely small effect size (Partial Eta Squared = 0.002). These findings suggest that the intervention did not lead to a significantly greater change in R-DAS compared to the control group over time.

While Fisher's Exact Test did not show a statistically significant difference in psychotropic medication use between groups (p = 0.151), the control group had more than three times as many participants on medication (21.9% vs. 6.9%). Given the small sample size, it is possible that a true difference was not detected due to limited statistical power. If medication contributed to symptom improvements, it may have influenced the reductions in EPDS and GAD-7 scores in the control group. Future studies with larger samples could further explore the role of psychotropic medication in influencing intervention outcomes.

A **Sensitivity analysis** was carried out for the study because some participants on both the Intervention and the Control Group did not adhere to the conditions of the group they were in. This was carried out by the statistician, who was removed from the results that were obtained.

In the sensitivity analyses the following cases were excluded:

Intervention Group (3 out of 29):

- **0387AP:** Attended only one therapy session and then stopped; no phone calls recorded.
- **PMH001AP:** Attended only four sessions and then stopped; no phone calls recorded.
- **1041AP:** Attended only two sessions and then stopped; no phone calls recorded.

Control Group (6 out of 32):

- **PMH136:** No phone calls recorded.
- **0197AP:** Only one phone call recorded.
- **PMH424AP:** Attended one phone call; therapy started eight weeks later.
- **1626AP:** No phone calls recorded; sought counselling elsewhere.
- **PMH152AP:** Two of the five calls were carried out face-to-face.
- **PMH2056:** Attended only one phone call and then started therapy elsewhere.

Reallocation of Cases:

Two cases, one from each group, switched their allocation group:

- **2130AP:** No therapy sessions but completed four phone calls; reallocated to the control group.
- **PMH153AP:** Completed five phone calls, was prescribed medication, and attended therapy in Mtarfa through her GP; reallocated to the intervention group.

Analysis 1: Pre- vs. Post-Measure Comparison: Control vs. Intervention Groups

This section investigates the differences in scores for the Edinburgh Postnatal Depression Scale (EPDS), Generalized Anxiety Disorder-7 (GAD-7), and Revised Dyadic Adjustment Scale (R-DAS) between pre- and post-intervention assessments, analysed separately for the control and intervention groups. These measures were selected to capture critical aspects of psychological and relational well-being during the study period.

To assess changes within each group over time, the paired t-test was employed. This statistical test evaluates whether the mean difference between pre- and post-measurements within a group is significantly different from zero, assuming the data is approximately normally distributed. By using the paired t-test, this analysis aims to:

- Determine whether significant changes occurred within each group (control and intervention) for EPDS, GAD-7, and R-DAS.
- Highlight the magnitude and direction of changes in mental health and relationship quality scores for each measure.

By analysing the pre- and post-scores separately for the control and intervention groups, this study provides insight into the effectiveness of the intervention and any observed patterns of change within each group.

- **Control Group**

The changes in scores for the **EPDS**, **GAD-7**, and **R-DAS** were analysed separately for the control group using paired t-tests. The table below summarizes the mean pre- and post-intervention scores, the change in scores, and the 95% confidence intervals (CIs) for the change in scores:

	Pre	Post	Change of score	95% CI of Change of Scores
EPDS	13.31 (4.946)	8.46 (5.736)	-4.85 (5.387)	[-7.02, -2.67]
GAD-7	10.58 (4.091)	6.00 (5.499)	-4.58 (7.044)	[-7.42, -1.73]
R-DAS	50.58 (9.441)	50.46 (9.709)	-0.12 (6.855)	[-2.88, 2.65]

Table 4: Pre- and Post-Intervention mean scores, Change in Scores, and 95% Confidence Intervals for the control

This table highlights the mean changes in scores for the control group across the three measures, along with the associated variability and confidence intervals. Below are the results of the paired t-tests conducted to assess differences in pre- and post-measurements for the control group across the EPDS, GAD-7, and R-DAS.

- **EPDS**

t(25) = 4.587, p-value < 0.001, Cohen's d = 0.900, 95% CI for Cohen's d [0.436, 1.351]

The paired t-test revealed a significant reduction in EPDS scores (-4.85) for the control group from pre- to post-measurement, t(25) = -3.313, p = 0.003. The effect size, measured by Cohen's d, was

0.900, indicating a large effect size. The 95% confidence interval for Cohen's d ranged from 0.436 to 1.351, demonstrating a substantial decrease in depressive symptoms over time.

- **GAD-7**

t(25) = 3.313, p-value = 0.003, Cohen's d = 0.872, 95% CI for Cohen's d [0.423, 1.069]

The paired t-test revealed a significant reduction in GAD-7 scores for the control group from pre- to post-measurement, t(25) = -3.313, p = 0.003. The effect size, measured by Cohen's d, was 0.872, indicating a large effect size. The 95% confidence interval for Cohen's d ranged from 0.423 to 1.069, suggesting a substantial improvement in anxiety symptoms over time.

- **R-DAS**

t(25) = 0.086, p-value = 0.932, Cohen's d = 0.017, 95% CI for Cohen's d [-0.368, 0.401]

The paired t-test indicated no significant change in R-DAS scores for the control group from pre- to post-measurement, t(25) = 0.086, p = 0.932. The effect size, measured by Cohen's d, was 0.017, suggesting a negligible effect. The 95% confidence interval for Cohen's d ranged from -0.368 to 0.401, further confirming the lack of meaningful change in relationship quality over the study period.

- **Intervention Group**

The changes in scores for the **EPDS, GAD-7**, and **R-DAS** were analysed separately for the intervention group using paired t-tests. The table below summarizes the mean pre- and post-intervention scores, the change in scores, and the 95% confidence intervals (CIs) for the change in scores:

	Pre	Post	Change of score	95% CI of Change of Scores
EPDS	13.42 (3.690)	8.88 (3.756)	-4.54 (4.641)	[-6.41, -2.66]
GAD-7	12.12 (3.734)	8.19 (5.052)	-3.92 (6.183)	[-6.01, -1.84]
R-DAS	47.31 (9.983)	48.85 (11.277)	1.54 (8.125)	[-1.74, 4.82]

Table 5: Pre- and Post-Intervention mean scores, Change in Scores, and 95% Confidence Intervals for the intervention

This table highlights the mean changes in scores for the intervention group across the three measures, along with the associated variability and confidence intervals. Below are the results of the paired t-tests conducted to assess differences in pre- and post-measurements for the intervention group across the EPDS, GAD-7, and R-DAS.

- **EPDS**

t(25) = 4.986, p-value < 0.001, Cohen's d = 0.978, 95% CI for Cohen's d [0.502, 1.441]

The p-value < 0.001 indicates that the result is statistically significant, meaning that there was a significant reduction in depressive symptoms following the intervention. Cohen's d = 0.978 suggests a large effect size, which indicates a substantial impact of the intervention on reducing depression. The 95% confidence interval for Cohen's d ([0.502, 1.441]) confirms that the true effect size is likely to fall within this range, further supporting the conclusion that the intervention had a large and meaningful effect on depressive symptoms.

- **GAD-7**

t(25) = 3.876, p-value < 0.001, Cohen's d = 0.760, 95% CI for Cohen's d [0.316, 1.192]

The significant p-value ($p < 0.001$) indicates a statistically significant reduction in anxiety symptoms post-intervention. Cohen's d = 0.760 suggests a medium-to-large effect size, indicating that the intervention had a moderate to substantial impact on reducing anxiety symptoms. The 95% confidence interval for Cohen's d ([0.316, 1.192]) supports this by showing that the true effect size is likely to fall within this range, which still suggests a meaningful effect.

- **R-DAS**

t(25) = -0.965, p-value = 0.344, Cohen's d = -0.189, 95% CI for Cohen's d [-0.575, 0.200]

The p-value = 0.344 indicates that the result is not statistically significant, meaning that there is no strong evidence to suggest a change in relationship satisfaction after the intervention. Cohen's d = -0.189 represents a very small effect size, suggesting that the intervention had little to no impact on relationship satisfaction. The 95% confidence interval for Cohen's d ([-0.575, 0.200]) includes zero, further indicating that the intervention likely had a negligible effect on R-DAS scores.

Conclusion

Both control and intervention groups showed significant reductions in depressive and anxiety symptoms, suggesting improvements in mental health over time. However, relationship satisfaction (measured by R-DAS) did not change significantly in either group, indicating that the intervention had no impact on relational quality. However, most mothers (14 out of 27 in the IG and 19 out of 26 in the CG) did not report relationship distress at baseline, which implies that their potential for improvement in RDAS would be limited. This could explain why the mean change in R-DAS was small and not statistically significant.

Analysis 2: Differences between the intervention and control groups, controlling for time.

In this analysis, we examine whether there are significant differences in EPDS, GAD-7 and R-DAS scores between the intervention group and control group, while accounting for the effect of time (pre- and post-intervention). A repeated-measures ANOVA is used to control for time and assess group differences in change scores over time.

The primary focus is to determine:

- Whether the intervention group shows greater improvement in EPDS, GAD-7 and R-DAS scores compared to the control group.
- The interaction effect of group (intervention vs. control) and time (pre vs. post), indicating whether the change in scores over time differs significantly between the groups.

The analysis also includes the calculation of effect size (Partial Eta Squared) to quantify the magnitude of the group difference while controlling for time. Confidence intervals for the mean change scores are provided to assess the precision of the observed differences. This approach allows us to evaluate the effectiveness of the intervention while accounting for natural variations over time and the baseline differences between groups.

This table displays the pre- and post-intervention scores for EPDS, GAD-7, and R-DAS in both the control and intervention groups, along with the change in scores, the corresponding 95% confidence intervals for the changes, and the mean and standard deviation in parentheses.

Control group				
	Pre	**Post**	**Change of score**	**95% CI of Change of Scores**
EPDS	13.31 (4.946)	8.46 (5.736)	-4.85 (5.387)	[-7.02, -2.67]
GAD-7	10.58 (4.091)	6.00 (5.499)	-4.58 (7.044)	[-7.42, -1.73]
R-DAS	50.58 (9.441)	50.46 (9.709)	-0.12 (6.855)	[-2.88, 2.65]
Intervention group				
	Pre	**Post**	**Change of score**	**95% CI of Change of Scores**
EPDS	13.42 (3.690)	8.88 (3.756)	-4.54 (4.641)	[-6.41, -2.66]
GAD-7	12.12 (3.734)	8.19 (5.052)	-3.92 (6.183)	[-6.01, -1.84]

R-DAS	47.31 (9.983)	48.85 (11.277)	1.54 (8.125)	[-1.74, 4.82]

Table 1: Pre- and Post- Intervention Scores for EPDS, GAD-7, and R-DAS in Control and Intervention Groups

- EPDS

In the **control** group:

- The EPDS score decreased by 4.85 points, indicating a reduction in depressive symptoms.
- The 95% CI [-7.02, -2.67] does not include 0, suggesting that the reduction is statistically significant.
- The variability (SD = 5.387) shows that not all individuals experienced the same level of improvement, but the overall trend points to a meaningful improvement in depressive symptoms.

In the **intervention** group:

- The EPDS score decreased by 4.54 points, again reflecting an improvement in depressive symptoms.
- The 95% CI [-6.41, -2.66] excludes 0, confirming this reduction is statistically significant as well.
- The standard deviation of the change (SD = 4.641) suggests slightly less variability compared to the control group.

Comparison Between Groups

The interaction term "**Time * Treatment Group**" examines whether the change in **EPDS** scores over time (pre- to post-intervention) differs between the control and intervention groups:

F(1, 50) = 0.049, p = 0.826, Partial Eta Squared = 0.001, 95% CI for Partial Eta Squared [0,0.071]

- **F(1, 50) = 0.049:** This F-value is quite small, indicating that there is little difference between how the EPDS scores changed over time between the control and intervention groups.
- **p = 0.826:** The p-value is greater than 0.05, meaning the interaction effect is not statistically significant. In other words, there is no evidence that the two groups experienced different changes in EPDS scores over time.
- **Partial Eta Squared = 0.001 with 95% CI [0, 0.071]:** The effect size is exceptionally small, indicating that only **0.1% of the variance** in EPDS scores is explained by group differences. The confidence interval includes zero, which reinforces the absence of any meaningful effect. The upper limit (0.071) represents a very small effect, showing the results align with negligible group differences.

From the statistical test results:

- The **interaction effect between time and group** was **not significant** (**p = 0.826**), indicating that the reduction in EPDS scores was **similar across the control and intervention groups**.
- This implies that while depressive symptoms improved over time, the intervention did not lead to a significantly greater reduction in symptoms compared to the control condition.

- GAD-7

In the **control** group:

- The GAD-7 score decreased by 4.58 points, indicating a reduction in anxiety symptoms.
- The 95% CI [-7.42, -1.73] does not include 0, suggesting that the reduction is statistically significant.
- The relatively large standard deviation (SD = 7.044) of the change indicates substantial variability in the extent to which individuals experienced improvements in anxiety symptoms.

In the **intervention** group:

- The GAD-7 score decreased by 3.92 points, also reflecting an improvement in anxiety symptoms.
- The 95% CI [-6.01, -1.84] excludes 0, confirming that the reduction is statistically significant.
- The standard deviation of the change (SD = 6.183) indicates variability in individual responses, although it is somewhat smaller than in the control group.

Comparison Between Groups

The interaction term "**Time * Treatment Group**" examines whether the change in **G**AD-7 scores over time (pre- to post-intervention) differs between the control and intervention groups:

F(1, 50) = 0.146, p = 0.704, Partial Eta Squared = 0.003, 95% CI for Partial Eta Squared [0,0.091]

- **F(1, 50) = 0.146:** The F-statistic is very small, indicating a minimal difference between groups for GAD-7 scores.
- **p = 0.704:** The interaction between time (pre vs. post) and group (control vs. intervention) is not statistically significant. This means there is no evidence that the intervention group showed a greater reduction in GAD-7 scores compared to the control group. Both groups experienced similar improvements in anxiety symptoms over time, regardless of the intervention.

- **Partial Eta Squared = 0.003 with 95% CI [0, 0.091]:** The effect size is extremely small, suggesting that only **0.3% of the variance** in GAD-7 scores is explained by group differences. The confidence interval includes zero, reinforcing the lack of a meaningful effect. The upper limit (0.091) still represents a small effect, showing the result is consistent with minimal to no group difference.

From the statistical test results:

- Both the control and intervention groups experienced statistically significant reductions in GAD-7 scores, indicating improved anxiety symptoms.
- However, the intervention did not lead to significantly greater reductions in anxiety symptoms compared to the control group, as indicated by the non-significant interaction effect ($p = 0.704$) and the negligible effect size.

- R-DAS

In the **control** group:

- The mean R-DAS score decreased slightly by 0.12 points from pre- to post-assessment.
- The 95% confidence interval [-2.88, 2.65] includes 0, suggesting that the change is not statistically significant.
- The small magnitude of change and its associated variability (SD = 6.855) indicate that the participants in the control group largely maintained their dyadic adjustment levels over time, with no meaningful improvement or decline.

In the **intervention** group:

- The mean R-DAS score increased by 1.54 points, indicating a slight improvement in dyadic adjustment for the intervention group.
- However, the 95% confidence interval [-1.74, 4.82] includes 0, meaning this change is not statistically significant.
- The variability (SD = 8.125) of the change in scores shows that participants' experiences varied, with some improving more than others, and some possibly worsening.

Comparison Between Groups

Interaction between Time and Treatment Group (time * Treatment Group)

The interaction term "**Time * Treatment Group**" examines whether the change in **EPDS** scores over time (pre- to post-intervention) differs between the control and intervention groups:

F(1, 50) = 0.629, p = 0.431, Partial Eta Squared = 0.012, 95% CI for Partial Eta Squared [0,0.127]

- **F(1, 50) = 0.629**: The F-statistic is quite small, indicating that the difference between groups (control vs. intervention) is not substantial.
- **p = 0.431**: The p-value is greater than 0.05, meaning the difference in change scores between the control and intervention groups is not statistically significant.
- **Partial Eta Squared = 0.012 with 95% CI [0, 0.127]**: This effect size suggests a **small effect**. Thus, only about **1.2% of the variance** in the change scores can be explained by group membership (intervention vs. control). The confidence interval includes zero, which supports the conclusion that the effect is not significant. The upper limit (0.127) suggests that, even under the most optimistic scenario, the effect would still be relatively small.

From the statistical test results:

- There is no statistically significant difference between the intervention and control groups for R-DAS scores. While the observed effect size is small (Partial Eta Squared = 0.012), the confidence interval reinforces that the effect could plausibly be zero or at most small in magnitude.
- The intervention did not produce a statistically significant improvement in R-DAS scores compared to the control group, and the effect size was small. This suggests limited impact of the intervention on relationship satisfaction as measured by R-DAS.

Conclusion

The results suggest that while both the control and intervention groups showed improvements in depressive symptoms (EPDS) and anxiety symptoms (GAD-7), the intervention did not lead to a significantly greater reduction in these symptoms compared to the control group. Similarly, there was no significant improvement in relationship satisfaction (R-DAS) for either group, with very

small effect sizes observed across all measures. The sensitivity analyses did not provide results that were different from the ones obtained in the ITT.