

London, 31st January, 2017

Babywatching Study – Full Report

Page | 1

Authors: Francesca Lionetti, Shannon Snelling, Michael Pluess

Institution: Queen Mary University of London
Department of Biological and Experimental Psychology
School of Biological and Chemical Sciences
Email: m.pluess@qmul.ac.uk

Babywatching is a school-based programme. A parent, and their baby, visit a class of children, once a week, for half an hour, so the children can watch the interactions between mother and baby, and their developing relationship, guided by a trained Group Leader, who is most often a teacher supported by a Mentor.

Babywatching aims at reducing children's anxiety and aggression by promoting understanding and empathy. The current study tested B.A.S.E Babywatching intervention outcomes and the moderating role of children's environmental sensitivity (Pluess, 2015) on these outcomes.

Methods

Participants

A sample of 161 children, their school teachers and parents, were involved in the study. The study was conducted in collaboration with a primary and a junior school located in London. The intervention condition consisted of 88 children assigned to the intervention condition, and 73 to the control condition. At baseline children involved in the study were 4 to 10 years old ($M = 6.32$ years, $SD = 2.04$, 50% of the sample was 4 to 5 years old and 50% 7 to 10 years old; 53.8% of the sample were male).

Measures and procedures

Contact for Further Information: Dr Michael Pluess, Department of Biological and Experimental Psychology, Queen Mary University of London, London E1 4NS, United Kingdom, Phone: +44 (0)207-882-8004, Email: m.pluess@qmul.ac.uk

The intervention efficacy was evaluated using teacher and parent reported questionnaires. The Strengths and Difficulties Questionnaire (SDQ, teacher and parent report) was used to assess prosocial competences, conduct problems, emotional problems, hyperactivity and problems in the relationship with peers; the Rating of Children's and Adolescent's Behaviour (TRSBS questionnaire, teacher and parent report) and the Griffith Empathy Measure (GEM, parent report) was included in order to assess empathy. Also, parents (e.g., mother, father, or other relevant primary caregivers) were asked to complete the Highly Sensitive Child scale (HSC) in order to assess the child's general sensitivity to environmental influences. Except for the HSC scale, a temperamental and relatively stable trait and assessed only at the baseline (PRE), all questionnaires were administered according to the following time-points: at baseline before the intervention (PRE), 5 months into the intervention (MID), and at 10 month once the intervention was completed (POST).

Plan of analyses

In the result section below, we first report descriptive data, followed by results pertaining to the evaluation of intervention effects using latent growth curve models. Analyses were conducted with R with a maximum likelihood estimation of models with missing data (Muthén & Muthén, 1998–2010). This method uses all the data that are available in order to estimate the model without imputing the data and is supported in lavaan, the R package used for estimating longitudinal changes.

The estimation and prediction of the longitudinal development across the repeatedly assessed outcome variables were analysed with latent growth curve models (Curran & Muthén, 1999; Meredith & Tisak, 1990; Muthén & Curran, 1997). Given that variables presented a non-normal distribution of data, the MLR estimator was used. The factor loading for the three measures on the latent intercepts factor were fixed to 1.0 to represent the initial starting point of the trajectories. To define the linear metric of time, the factor loadings for the slope were set to 0, 1, and 2. The latent growth factors were regressed on the variable programme (2 = control group; 1 = experimental group) in order to test if the B.A.S.E. Babywatching intervention affects the change over time of the target variables (MacKinnon, 2008).

Contact for Further Information: Dr Michael Pluess, Department of Biological and Experimental Psychology, Queen Mary University of London, London E1 4NS, United Kingdom, Phone: +44 (0)207-882-8004, Email: m.pluess@qmul.ac.uk

If the intervention condition significantly predicted change over time, a latent growth multi-group analysis was run to compare the slope of the intervention condition with the slope of the control condition.

Results

Descriptive statistics and between group comparison

Baseline – observed values

Correlation values between age at assessment, in years, and study variables at the baseline, are reported in Table 1 ($N=156$). Overall, correlations between parent and teacher reports were low to medium, ranging from $r = .25$ for prosocial competence at the SDQ to $r = .40$ for empathy at the TRSBS. The associations between age at assessment (in years) and study variables were low to moderate in effect size (highest correlation was $r = .28$ for the conduct problem scale).

Pertaining to gender, females were higher than males on prosocial competences (SDQ, $t(144.22) = 3.51$, $p < .001$) and empathy (TRSBS, $t(140) = 4.1$, $p < .001$) and lower in conduct problems (SDQ, $t(138.88) = -2.26$, $p = .03$) and hyperactivity (SDQ, $t(129.91) = -3.69$, $p < .001$), based on teacher-report questionnaires. When considering parent-report questionnaires, only one significant gender difference emerged for empathy, with higher competence reported for females (GEM, $t(135.5) = 2.15$, $p = .03$).

Overall, children in the intervention group were significantly more at risk than children belonging to the control group. Specifically, they were rated higher by their teachers on conduct problems (SDQ, $t(144.49) = 3.362$, $p < .001$) and emotional problems (SDQ, $t(152.4) = 3.485$, $p < .001$). Furthermore, children belonging to the intervention group were rated by teachers as being less prosocial (SDQ, $t(153) = -2.636$, $p = .009$) and having less empathy (TRSBS, $t(151.82) = -1.935$, $p = .05$). Based on parents' reports, children belonging to the control group had higher scores regarding problem behaviours, and lower scores along scales evaluating positive social competences and empathy, but the differences were not significant. Descriptives for the baseline data are reported in Table 2.

Table 1. Correlation among study variables at the baseline. Pearson's *r* is a measure of association, values of .10, .30, and .50 represent small, moderate, and large effect size respectively

	Age	PS_P	CP_P	EP_P	HA_P	PP_P	EMP1_P	EMP2_P	PS_T	CP_T	EP_T	HA_T	PP_T	EMP1_T	HSC_P
Age	--	.086	.133	.143	.240*	.070	.164*	.064	.193*	-.283**	.064	-.164*	.016	.251*	-.023
PS_P		--	-.513**	-.269**	-.377**	-.337*	.579**	.498**	.252*	-.189*	-.116	-.313**	-.207*	.323**	.001
CP_P			--	.368**	.547**	.391**	-.392**	-.399**	-.206	.282**	.114	.331**	.264**	-.216*	-.216*
EP_P				--	.307**	.379**	-.244*	-.152*	-.027	.033	.357**	.053	.140	-.072	-.300*
HA_P					--	.332**	-.204	-.285**	-.240*	.238*	.145	.496**	.332**	-.240*	-.074
PP_P						--	-.405**	-.341**	-.347**	.264*	.222*	.353**	.441**	-.307**	-.180
EMP1_P							--	.632**	.316**	-.275**	-.139	-.270**	-.313**	.403**	.072
EMP2_P								--	.236*	.238*	-.027	-.308**	-.227**	.293**	-.045
PS_T									--	-.662**	-.301**	-.691**	-.341**	.763**	.055
CP_T										--	.252*	.709**	.283**	-.606**	-.115
EP_T											--	.256**	.585**	-.250**	-.180
HA_T												--	.329**	-.651**	-.075
PP_T													--	-.353**	-.046
EMP1_T														--	.148

Note. PS = Prosocial Competences; CP = Conduct Problems; EP = Emotional Problems; HA = Hyperactivity; EMP1 = empathy assessed with the TRSBS questionnaire; EMP2 = empathy assessed using the GEM Scale. P = parent report; T = teacher report; **p* < .05; ***p* < .001.

Table 2. Mean values (M) and standard deviations (SD) for all measures in the intervention and in the control group at the baseline according to parents and teacher's report.

	<i>Parent report</i> M(SD)		<i>Teacher report</i> M(SD)	
	Intervention	Control	Intervention	Control
PS	7.72 (1.90)	8.11 (1.73)	5.71 (2.77)	6.88 (2.79)
CP	1.66 (1.71)	1.50 (1.60)	1.73 (2.25)	0.77 (1.32)
EP	1.99 (2.26)	.156 (1.71)	2.66 (2.30)	1.38 (2.34)
HA	3.48 (2.55)	3.83 (2.50)	3.62 (3.27)	2.90 (2.72)
EMP1	19.46 (3.14)	20.28 (2.69)	17.21 (3.59)	18.32 (3.59)
EMP2	31.91 (18.24)	32.51 (17.35)	--	--

Note. PS = Prosocial Competences; CP = Conduct Problems; EP = Emotional Problems; HA = Hyperactivity; EMP1 = empathy assessed with the TRSBS questionnaire; EMP2 = empathy assessed using the GEM Scale.

Post-intervention – observed values

At the post-intervention assessment phase, correlation between parents and teachers' report have been computed only for 111 subjects given that for parents there were 45 missing questionnaires. Results, reported in Table 3, identified a stable trend for the association between parent and teachers' reports suggesting overall a moderate association between teacher and parent's report regarding children's behavioural features. At the post-assessment intervention, no statistically significant differences between the control and intervention groups were identified according to parents' report. Similarly, no group differences emerged according to teacher reports, suggesting that while the two groups were significantly different at the pre-intervention time point of assessment, they reached comparable competencies along the target variables at the post-intervention assessment.

Table 3. Correlation among study variables at the post-intervention assessment phase. Pearson's r is a measure of association, values of .10, .30, and .50 represent small, moderate, and large effect size

	PS_P	CP_P	EP_P	HA_P	PP_P	EMP1_P	EMP2_P	PS_T	CP_T	EP_T	HA_T	PP_T	EMP1_T
PS_P	--	-.527**	-.194*	-.339**	-.314**	.521**	.541**	.391**	-.145	-.154	-.275**	-.183**	.281**
CP_P		--	.503**	.439**	.353**	-.357**	-.345**	-.384**	.252**	.131	.304**	.243**	-.342**
EP_P			--	.182*	.226**	-.163	-.122	-.275**	.037	.459**	.164	.121	-.109
HA_P				--	.565**	-.409**	-.106	-.331**	.427**	.081	.565**	.334**	-.417**
PP_P					--	-.620**	-.148	-.338**	.619**	.198	.564**	.670**	-.520**
EMP1_P						--	.412**	.174	-.439**	.018	-.377**	-.450**	.332**
EMP2_P							--	.267**	-.008	-.089	-.129	-.081	.264**
PS_T								--	-.596**	-.211*	-.551**	-.397**	.741**
CP_T									--	.149	.716**	.734**	-.596**
EP_T										--	.346**	.474**	-.213**
HA_T											--	.651**	-.592**
PP_T												--	-.508**

Notes. PS = Prosocial Competences; CP = Conduct Problems; EP = Emotional Problems; HA = Hyperactivity; EMP1 = empathy assessed with the T RSBS questionnaire; EMP2 = empathy assessed using the GEM Scale. P = parent report; T = teacher report; * $p < .05$; ** $p < .001$.

Table 4. Mean values (M) and standard deviations (SD) for all measures in the intervention and in the control group at the post-intervention assessment according to parents and teacher's report.

	<i>Parent report</i> M(SD)		<i>Teacher report</i> M(SD)	
	Intervention	Control	Intervention	Control
PS	8.24 (1.78)	8.64 (1.41)	7.85 (2.34)	7.83 (2.42)
CP	1.51 (1.69)	1.02 (1.26)	1.03 (2.88)	0.69 (1.14)
EP	1.30 (1.79)	1.59 (1.89)	1.05 (1.53)	1.29 (1.97)
HA	3.01 (2.89)	3.22 (2.56)	2.22 (3.59)	2.39 (2.74)
EMP1	19.26 (4.55)	20.27 (3.71)	19.84 (4.20)	19.69 (3.65)
EMP2	34.18 (19.07)	36.82 (17.95)	--	--

Note. PS = Prosocial Competences; CP = Conduct Problems; EP = Emotional Problems; HA = Hyperactivity; EMP1 = empathy assessed with the TRSBS questionnaire; EMP2 = empathy assessed using the GEM Scale.

Teacher report – Main effects

Prosocial competences (SDQ – see Figure 1).

The intervention significantly predicted the intercept at the baseline, suggesting that the experimental and the control groups were significantly different before the intervention ($B = 1.109 (.445)$, $p = .013$). Specifically, children belonging to the intervention group had lower prosocial skills than their peers. Interesting to observe, the intervention significantly predicted the linear change of prosocial competences over time ($B = -.707 (.283)$, $p = .013$) and allowed children belonging to the intervention group to reach levels of social competences comparable to that of their peers in the control condition. At the post-intervention assessment, the two groups were comparable as shown in Figure 1 ($B = -.305 (.333)$, $p = .363$). Multi-group analyses suggested that both the intervention and control group slopes were significant. In both groups, an increase in social competences was observable but the effect size was double in the intervention group compared to the control group (intervention group, $B = 1.170 (0.193)$, $p < .001$; control group, $B = .510 (.190)$, $p = .007$).

Conduct problems (SDQ – see Figure 2).

The intervention significantly predicted the linear change of conduct problems over time ($B = .523 (.167)$, $p = .002$), but again there was no between group difference at the follow-up assessment ($B =$

.102 (.200), $p = .611$). Given the different intercept at the baseline between the two experimental conditions ($B = -.943 (.273)$, $p < .001$), data suggest that the intervention reduced conduct problems for children in the intervention condition to levels comparable to those children in the control condition. Multi-group analyses revealed that in the control group the slope was non-significant ($B = -.033 (0.092)$, $p = .719$) whereas it was in the intervention group it was ($B = -.560 (.143)$, $p < .001$).

Emotional problems (SDQ – see Figure 3).

The intervention significantly predicted the linear change of emotional problems over time ($B = .825 (.217)$, $p < .001$) but at the follow-up assessment the two groups did not differ ($B = .367 (.257)$, $p = .152$). Given the different intercepts at the baseline between the two experimental conditions ($B = -1.282 (.358)$, $p < .001$), data suggest that the intervention reduced emotional problems in the intervention condition to levels comparable to children belonging to the control condition. Multi-group analyses revealed that in the control group the slope was non-significant ($B = -.097 (0.143)$, $p = .497$) whereas it was in the intervention group it was ($B = -.828 (.143)$, $p < .001$).

Hyperactivity (SDQ). No relevant intervention effect was identified over time ($B = .413 (.312)$, $p = .185$) Problems with peers (SDQ). Similarly, no significant effect of the intervention for the problem with peers variable was identified ($B = .167 (.187)$, $p = .372$).

Empathy (TRSBS). Observed values suggested a quadratic slope but with three time points it wasn't possible to test a quadratic model. The effect of the intervention on TRSBS could not be reliably estimated. According to t-tests, no between groups difference at the baseline and at the follow up assessment was identified.

Teacher report –Interaction effects

Interaction with children's environmental sensitivity. Overall, HSC did not significantly interact with the intervention condition in predicting changes over time.

Interaction with children's age (primary vs. junior school). Multi-group analyses suggest that the effect was evident for children aged 4 to 5 but not for children aged 7 to 10.

Parent report

Page | 9

At the mid and post-intervention time points, there were 45 parents who no longer completed questionnaires. Results suggest no significant intervention effects according to parents but the high number of missing data at the mid and at the post-assessment phase requires caution regarding the interpretation of these findings. Furthermore, the tested models converged but the matrix was often not positive defined.

Discussion

The current study aimed at evaluating the effect of a school-based intervention programme, B.A.S.E. Babywatching, in promoting positive social-emotional and behavioural adjustment in children aged 4 to 10 years. For evaluating intervention efficacy, 161 children, their teachers and parents, coming from a Primary and a Junior School located in London, took part in the study and were evaluated three times across 10 months: before, during, and after the intervention.

According to our findings, children allocated to the intervention condition had significantly more problems before the intervention compared to children in the control condition but after the intervention, there were no longer any significant differences between groups. Hence, the results of the study suggest that the intervention promoted positive prosocial competences and reduced emotional and conduct problems over time in children belonging to the intervention group. These results emerged only with teacher reported data. The finding of significant differences between the two groups at baseline but not at follow up assessment, could be interpreted either as reflecting group differences due to the non-blind experimental condition, or to the non-random assignment of children to intervention and control conditions (i.e., children assigned to the intervention condition may have been those considered more at risk in terms of behavioural and emotional problems). The observation that parent-reported data proposes a similar but non-significant trend suggests that findings are not just merely the reflection of teachers' bias when completing questionnaires.

Contact for Further Information: Dr Michael Pluess, Department of Biological and Experimental Psychology, Queen Mary University of London, London E1 4NS, United Kingdom, Phone: +44 (0)207-882-8004, Email: m.pluess@qmul.ac.uk

Overall, no interaction with environmental sensitivity emerged but multi-group analyses suggested that the positive effect of the intervention was restricted to younger children attending the primary school and aged 4 to 5 years.

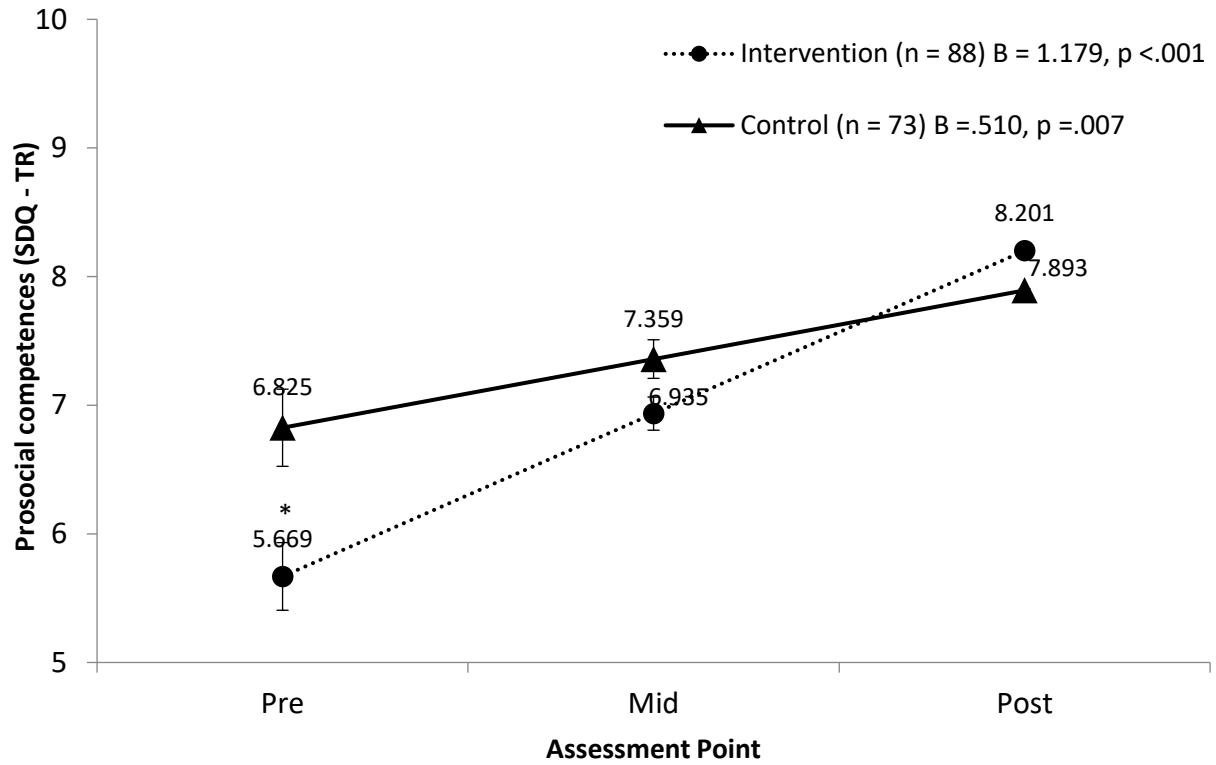
Page | 10

To conclude, results of this study suggest that Babywatching could be a promising programme to promote positive change in prosocial competence, conduct and emotional problems over time. To more reliably support these results, further studies adopting a randomized control trial study design, and using blind coders for the evaluation of the intervention, are required. Furthermore, future studies should test intervention effects in children of different ages in order to investigate whether treatment effects are strongest in younger versus older children.

FIGURES

Figure 1. Intervention effects on prosocial competences. M and SE based on predicted values.

Page | 11



* $p = .013$

Figure 2. Intervention effects on conduct problems. M and SE based on predicted values.

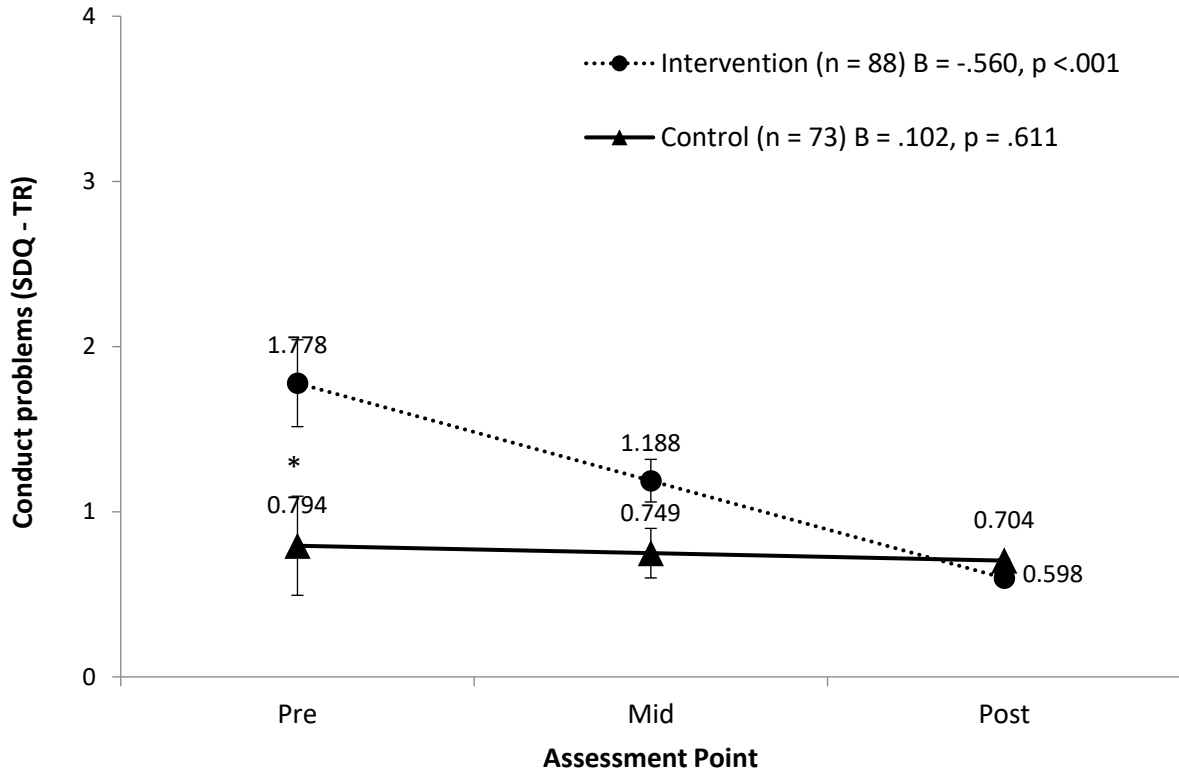
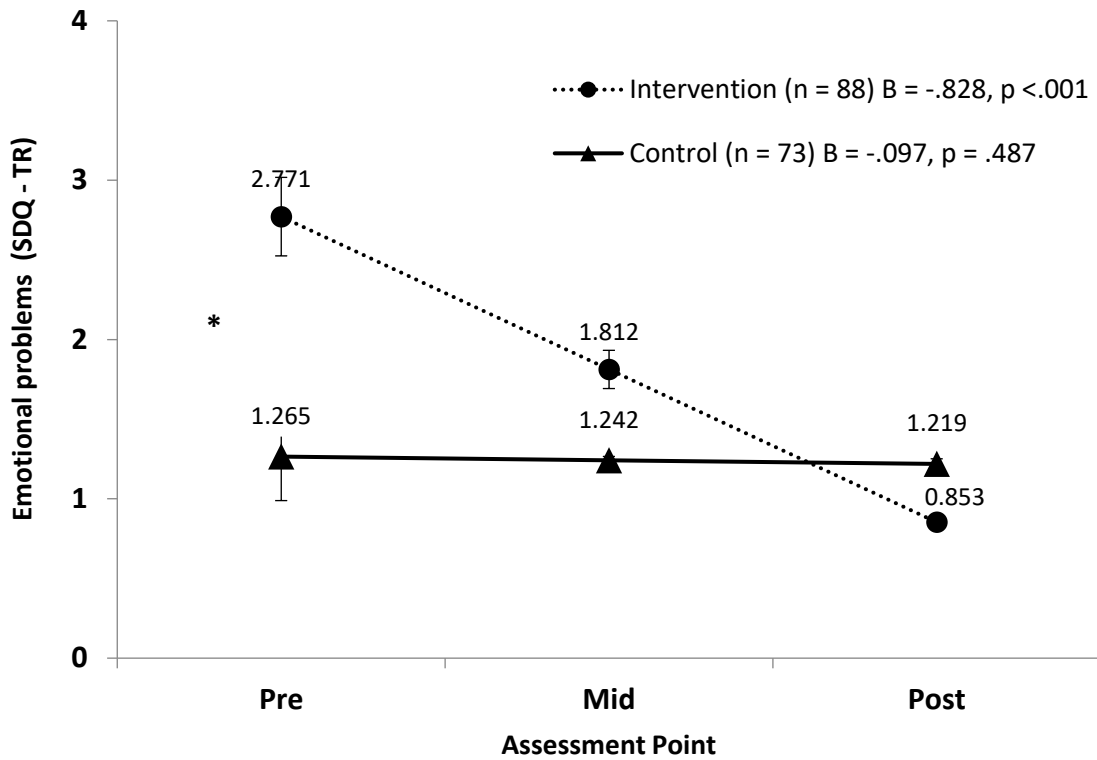


Figure 3. Intervention effects on emotional problems. M and SE based on predicted values.



R SCRIPT

Analyses were conducted using a structural equation approach and specifically using latent growth curve models. Based on needs, negative variances have been fixed at the baseline. Results were replicated using a mixed model approach; findings were stable irrespective of the type of analytic strategy adopted.

(1) For evaluating the main effect of the intervention (LCGM)

```
library(lavaan)
```

```
model <- '
i =~ 1*OUTCOME_BASELINE + 1*OUTCOME_MID + 1*OUTCOME_POST
s =~ 0*OUTCOME_BASELINE + 1*OUTCOME_MID + 2*OUTCOME_POST ##or -1, 0, 1 was used when data were centred at the mid-point and -2, -1, 0 when data were centred at the post-intervention assessment phase)
```

```
OUTCOME_BASELINE ~~ 1*OUTCOME_BASELINE
```

```
i ~ Condition
s ~ Condition
```

```
fit <- growth(model, data=X)
```

(2) For evaluating the interaction effects

```
model <- '
i =~ 1*OUTCOME_BASELINE + 1*OUTCOME_MID + 1*OUTCOME_POST
s =~ 0*OUTCOME_BASELINE + 1*OUTCOME_MID + 2*OUTCOME_POST ##
```

```
i ~ Condition + HSCbaseline + ConditionXHSCbaseline
s ~ Condition + HSCbaseline + ConditionXHSCbaseline
```

```
fit <- growth(model, data=X)
```

(3) For evaluating individual slopes

```
model <- '
i =~ 1*OUTCOME_BASELINE + 1*OUTCOME_MID + 1*OUTCOME_POST
s =~ 0*OUTCOME_BASELINE + 1*OUTCOME_MID + 2*OUTCOME_POST ##
```

```
fit <- growth(model, data=X, group = "Condition")
```