

1

2

3

4 Music education reduces emotional symptoms in children: A quasi-experimental study  
5 in the Guri Program in Brazil

6

7 Graziela Bortz<sup>1\*</sup>, Beatriz Ilari<sup>2¶</sup>, Nayana Di Giuseppe Germano<sup>3¶</sup>, Andrea Parolin

8 Jackowski<sup>4&</sup>, Hugo Cogo-Moreira<sup>5&</sup>, Patrícia Silva Lúcio<sup>6¶</sup>

9

10

11

12 <sup>1</sup> Department of Music, Unesp, Sao Paulo, SP, Brazil

13 <sup>2</sup> Thornton School of Music, USC, Los Angeles, CA, USA

14 <sup>3</sup> Department of Music, Federal University of Santa Maria, Santa Maria, RS, Brazil

15 <sup>4</sup> Department of Psychiatry, Federal University of Sao Paulo, Sao Paulo, SP, Brazil

16 <sup>5</sup> Department of Education, Østfold University College, Østfold, Fredrikstad, Norway

17 <sup>6</sup> Department of Psychology and Psychoanalysis, State University of Londrina, Londrina, PR,  
18 Brazil

19

20 **Short title:** Music education reduces emotional symptoms in children

21

22 \* Corresponding author

23 E-mail: [graziela.bortz@unesp.br](mailto:graziela.bortz@unesp.br)

24

25

26 ¶These authors contributed equally to this work.

27 &These authors also contributed equally to this work.

28

29

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

## 30 Abstract

31 The transferability of music education to cognitive and social skills has been explored recently,  
32 but its causal effects remain debatable. This quasi-experimental study aimed to examine the  
33 impact of a music education program (Guri) on aspects attention, working memory and  
34 socioemotional skills in children from underserved communities of São Paulo, Brazil. The music  
35 group (n = 38, 5-9 years) was recruited from 10 centers (*polos*) distributed across the  
36 metropolitan area of São Paulo and the control group (n = 67) consisted of aged-matched  
37 children who attended public schools surrounding the *polos*. The ABEP SES questionnaire and  
38 the Raven test were used as control variables, the SDQ questionnaire, the Digit-span subtest  
39 from WISC-IV, the PARB-Q questionnaire, and the BPA test were used as subject variables. A  
40 significant effect of the intervention in emotional symptoms in the music group as opposed to  
41 controls was found ( $F(1,89) = 4.562, p = 0.035, \eta^2 = 0.049$ ), indicating benefits for the music  
42 group. Children whose mothers had low levels of education benefited with gains between 0.70  
43 and 0.95 standard deviations in both groups for divided attention, indicating a significant  
44 interaction with maternal education.

45

## 46 Introduction

### 47 Music education and executive functions

48 The hypothesis of the effects of music education practice on cognitive abilities and  
49 brain plasticity, with potential far transferability to other domains, has been the focus of  
50 several studies published in the last decade (1-11). While some studies suggest that abilities  
51 related to executive functions (EFs) are particularly enhanced by musical training (2-4), others  
52 remain inconclusive (11, 12). Yet inhibitory control, one of the central dimensions of EFs,  
53 appears to be a consensus in terms of its transferability from music education/music training,  
54 especially among preschoolers (12).

55 EFs consist of a set of abilities that promote goal-directed and autonomous behavior,  
56 and are crucial to stablish day-to-day goals as well as more complex and long-time planning  
57 (13-17). Although validated models of EFs are still controversial (14, 18, 19), there are  
58 commonalities among proposed theoretical models, including dimensions such as attention,  
59 inhibitory control, working memory (WM), fluid intelligence, problem solving, decision making,  
60 and cognitive flexibility.

61 EFs can be classified as *hot*, or those associated with emotional and social cognitive  
62 aspects relevant to decision-making process, or *cold*, which are linked to mental abilities  
63 involved in inhibitory control, WM, short-term memory (STM), reasoning and planning (20,  
64 21). Cold EF tasks require mental effort and constant top-down supervision of the pre-frontal  
65 cortex (PFC) to prevent automatic actions (16).

66 According to Diamond's model (17), and despite belonging to EF cold domains,  
67 attention, a subdomain of inhibitory control, can be affected by emotional and cognitive self-  
68 regulation (22, 23). Attention requires effort and volition and, therefore, can also be involved  
69 in both hot and cold EFs depending on the type of stimulus to be inhibited (*e.g.*, self-control in  
70 delayed gratification or resistance to a distractor in attentional focus). The two types of  
71 attention (*i.e.*, emotional and cognitive) may not necessarily fall under the same factor, but are  
72 strongly correlated (17, 24). Petersen and Posner (23) updated their model of attention (22) to  
73 include three major networks: alerting, orienting, and executive networks. The latter further  
74 includes focal attention and regulation of processing networks.

75 Attention can also be classified according to the type of task at hand: (1) sustained  
76 attention/focused (*i.e.*, maintaining focus on a single stimulus/task over time), (2) divided  
77 attention (*i.e.*, concentrating on two or more stimuli/tasks simultaneously for a given time),  
78 and (3) shifting/alternated attention (*i.e.*, change the focus of the attention to the perceptual  
79 attributes of the stimuli under relevant stimulus-response maps rules) (25, 26). Attention also  
80 plays an important role in WM, another dimension of cold EF that requires processing and  
81 storage, sharing cognitive resources from both (27). Therefore, it is necessary to maintain a  
82 focus of attention by inhibiting the preponderant stimulus to prioritize information and allow  
83 updates of the environment, and subsequently coordinating them with automated responses  
84 in the use of WM. All this happens while new associations are created in STM.

85 A meta-analysis by Talamini et al. (8) found a large musician advantage for STM and  
86 WM with tonal stimuli, a moderate effect with verbal stimuli, and a non-significant effect with  
87 visuospatial stimuli. The systematic review of Rodriguez-Gomez and Talero-Gutierrez (12)  
88 examining the associations between WM and music training revealed three studies among  
89 eight that showed improvement in one subdomain (*i.e.*, either visual or spatial) following a  
90 music intervention with children. Additionally, in four studies reviewed by the authors, music  
91 education was found to improve selective attention among children, although only one study  
92 showed statistical significance. Furthermore, out of nine studies conducted with preschool-age  
93 children, six presented improvements in inhibitory control following music education. From  
94 these studies, one can assume that verbal WM may be a strong candidate for music education  
95 effects, as seen with selective attention.

## 96 Hot executive functions, social competence, and social skills

97 Hot EFs are far less researched than cold EFs, even though psychometric instruments  
98 that consider level of affection, emotion self-regulation, affect decision making, empathy, and  
99 delayed gratification are often used in studies (15). One of the reasons for the lower incidence  
100 of articles on hot EFs might be the entanglements between the concept of hot EFs and the  
101 constructs of social competence, social skills and self-regulation. For Zimmerman et al. (20),  
102 “hot executive functions entail cognitive abilities supported by emotional awareness and social  
103 perception (*e.g.* social cognition).” The authors adapted their concept from the models of Chan  
104 et al. (28), and McDonald (29), placing social cognition on one side, and emotion cognition and  
105 decision making on the other, and subsequently positioning all under the umbrella of hot EFs.  
106 Social cognition, in turn, was also further divided into cold (*i.e.*, cognitive empathy and theory  
107 of mind) and hot (*i.e.* affective empathy and theory of mind, and emotion recognition) aspects  
108 (20). For example, while one person may make decisions based on *understanding* other

109 people's cues, others may arrive at decisions based on their *affections*. Unsurprisingly, both  
110 cold and hot social cognition have neural correlates (30-32).

111 Social competence has been referred to as the ability to maintain positive relationships  
112 while pursuing personal goals (33). For Hukkelberg et al. (34), whereas social competence is a  
113 broad term, social skills can be situated within more specific behaviors, as the latter can be  
114 examined through observable variables that compound a social competence factor (35-38).

115 Both social skills and social competence are positively correlated with academic  
116 achievement and problem solving skills (35, 38, 39), and have been associated with positive  
117 mental health (35, 38). Social skills can be measured with scales and subscales that address  
118 cooperation, self-control, internalizing and externalizing problem behaviors, empathy and  
119 hyperactivity (35, 38, 39).

## 120 Social skills and music

121 Associations between emotion self-regulation, prosocial skills, and music are still under  
122 scrutiny. While some studies observed effects of musical interventions on prosocial skills in  
123 preschoolers and school-aged children (40-42), others found moderate to null effects (11, 43,  
124 44). Still other studies found effects of music education only on selected social skills of children  
125 and youth. Villanueva et al. (45), for example, found effects of formal music instruction on  
126 children's sharing behaviors, theory of mind, and state empathy. Alemán et al. (11), in turn,  
127 found a significant improvement in behavioral difficulties detected via parental questionnaires  
128 and scales, especially among male children who had experienced violence and had mothers  
129 with low levels of formal education. Also using a parental scale, Ilari et al. (46) found a  
130 significant effect of four years of music education on children's aggressiveness and  
131 hyperactivity. Similarly, Boal-Palheiros and Ilari (38) found effects of one year of music  
132 education on externalizing behavioral problems in Portuguese children from underserved

133 communities. Thus, current evidence is mixed and needs further substantiation, and this is  
134 particularly true in the case of children from non-WEIRD (i.e., non-White, from English  
135 speaking, industrialized, rich and democratic) populations (47) and those at risk for behavioral  
136 problems.

137

## 138 The present study

139 This study aims to contribute towards the area of music education through an  
140 examination of its impact on specific cognitive and socioemotional outcomes among children  
141 at risk for behavioral problems. This quasi-experimental investigation adds to the body of  
142 knowledge derived from cross-sectional and correlational research that found evidence  
143 supporting beneficial effects of music education in the target outcomes. Particularly, we focus  
144 on aspects of attention (shifting, divided, and sustained) and aggressive behavior, by  
145 comparing test results from a group of children who participated in formal music classes to  
146 matched controls.

147 As noted, most of the studies examining the associations between music, EFs, and  
148 social skills were conducted in the global North, predominantly in North America and Western  
149 Europe and with WEIRD populations (47). This study<sup>1</sup> aimed to investigate whether  
150 participation in a musical program called *Projeto Guri* (48) designed to serve low-income  
151 communities in the city of São Paulo, Brazil, conferred any benefits to children's cognitive and  
152 social skills, including those who were at risk for behavioral problems.

153

---

<sup>1</sup> The data discussed here were first presented as a thesis of senior professor exam in 2023 by the corresponding author. It remains unpublished, and the results were reviewed in this article.

## 154 Method

### 155 Ethics

156 All research protocols were approved by the local Ethical Committee *Plataforma Brasil*  
157 n. 15049919.8.0000.5420, and the Rebec n. U1111-1243-8071. Written consent was also  
158 obtained from the municipality of São Paulo. A Memorandum of Understanding was created  
159 between the State University of São Paulo (Unesp) and the Guri Program, outlining the  
160 specifics of the research, expectations, and obligations from all involved parts.

### 161 Sample and recruitment

162 Children and their families were invited to participate in the study through their music  
163 programs (Guri *polos*) and local schools. Due to ethical reasons related to the policies of  
164 enrollment in the Guri Program, it was not possible to randomize child participants into music  
165 and control groups.<sup>2</sup> Therefore, we performed an intervention study based on a quasi-  
166 experimental design, which took place in 2022 during the Brazilian academic year—February  
167 through December.<sup>3</sup> The researchers responsible for the statistical analyzes were blinded to  
168 the participants. The examiners were also blinded to those who analyzed the outcomes. The  
169 target music group consisted of children aged 5 to 9 years,<sup>4</sup> who were about to start music  
170 classes for the first time in the Guri Program (48). This Program consists of a network of  
171 various centers –called *polos*– distributed across the metropolitan area of São Paulo, also  
172 known as *Grande São Paulo*. Most of the *polos* are located in the outskirts of the city and  
173 function within a public system of integrated schools called *Centros Educacionais Unificados*

---

<sup>2</sup> The children also underwent MRI scans at baseline and follow-up. These data are not included in this article. For adherence, the children then received a gift, a snack, a certificate of contribution to scientific research and their families received a refund for transportation.

<sup>3</sup> The original schedule proposed in the trial study protocol predicted that the intervention would take place during the 2020 academic year, preceded by a pilot study in 2019. Due to the Covid-19 pandemic, the intervention had to be interrupted during the initial data collection, in March 2020, and the study was resumed in 2022 with new participants.

<sup>4</sup> In the trial study protocol, we intended to test children aged between 6 and 7, but as it was difficult to achieve the minimum sample size desired in this group, the range was extended.

174 (Unified Educational Centers) or CEUs. Children in the music group were recruited from 10  
175 different *polos*: 7 from the extreme east part of the city, which is also the most populated, 3  
176 from the extreme south, and 2 from the extreme north. Beginning music classes take place  
177 weekly for one hour through a modality known as *musicalização* (49), a form of general music  
178 that is carried out through collective singing, movement, and performance of rhythms and  
179 melodies, using small percussion instruments, recorders, and Orff xylophones and  
180 metallophones.

181 The control group consisted of aged-matched children who attended regular schools  
182 surrounding the Guri *polos*. Exclusion criteria for sample recruitment in both groups were: 1)  
183 children had attended regular school music classes or instrumental/voice lessons before, and  
184 2) children had received a diagnosis of ASD (Autism Spectrum Disorders) or ADHD (Attention  
185 Deficit Hyperactivity Disorder). Children in the experimental group were recruited at the time  
186 of enrollment in the Guri Program, when parents/guardians were informed about the study  
187 and given consent forms. After signing consent forms, parents were invited to complete the  
188 SDQ–parent version (50, 51) and the ABEP questionnaire (52) that assesses socioeconomic  
189 status. This procedure constituted the first screening phase and aimed to identify children who  
190 would attend the Guri Program in 2022, and who were at risk for behavioral problems in four  
191 of the five areas evaluated by the SDQ, namely, emotional symptoms, conduct problems,  
192 hyperactivity, and relationship problems.

193 Recruitment was done upon securing authorization from the school boards to  
194 participate in the research. The researchers met with parents at the schools to explain the  
195 study, get consent, and subsequently distribute questionnaires. When meeting at schools was  
196 not possible, questionnaires were sent home by children and subsequently collected. A total of  
197 255 parent protocols were collected at the Guri Program and in schools surrounding the  
198 programs for the control group, out of which 137 protocols were excluded. After further  
199 screening with the Brazilian version of the Raven’s test (53, 54), children at risk for intellectual



200 difficulties were excluded from the sample to avoid confounders (n = 10). One child was  
201 excluded for having previous music lessons and another was diagnosed with autism, totaling  
202 149 exclusions. The total of participants (n = 106) consisted in the intervention group –music  
203 group (n = 45)– and the control group (n = 61). The music group lost 7 children, who dropped  
204 out of the music program at the follow up (FUP). Therefore, 38 children remained in the music  
205 group. One child was excluded of the control group at the FUP for late autism diagnosis.  
206 Baseline data were collected in March and April of 2022, and the FUP tests and questionnaires  
207 were taken seven months later<sup>5</sup>. Because children differed significantly in both groups (see  
208 Sample Matching below), 35 children from the music group and 57 children from the control  
209 group were included in the final statistical analysis after matching (Fig 1).

210 **Fig 1. Flowchart of participants through each stage of the trial.**

211

212

213 Children in the music group were 5-9 years old (mean = 6,61; S.D. = 0,72) and 52,6%  
214 boys at the baseline. In the control group, children were mostly boys (70%) and were 5-7 years  
215 old (mean = 6,28; S.D. = 0,49). Seven children from the music group dropped out from the  
216 research. Children who dropped out did not differ from those who completed the FUP in terms  
217 of sex ( $\chi^2(1) = 0,007, p = 1,000$ ) or group assignment ( $\chi^2(1) = 2,321, p = 0,148$ ), but they tended  
218 to be older at the baseline ( $t(96) = 2,589, p = 0,011$ ). Two children were later diagnosed, one  
219 from the music group and another from the control group.

---

<sup>5</sup> This study started at 2020, but when the intervention was about to begin, the peak of the coronavirus pandemic come. By that time, although the baseline assessment had already finished, the children of the experimental group were taking remote music classes and the research had to be cancelled. Therefore, this sample is different from that in their pre and post-test assessments.

## 220 Instruments

### 221 Control variables

- 222 • *Maternal schooling*: We used the Brazilian questionnaire of economic classification criteria  
223 –*Associação Brasileira de Empresas de Pesquisa* (52), created to assess socioeconomic status  
224 based on samples of households. This questionnaire contains nine items that assess: a)  
225 possession of durable consumer goods; b) type of water supply system and street paving; c)  
226 number of residents in the household; d) family composition; and e) level of education of the  
227 head of the family. The result of the questionnaire is a measure stratified into five  
228 socioeconomic classes, namely: A (subdivided into A1 and A2), B (also subdivided into B1 and  
229 B2), C, D and E. For the purpose of this study, we used the mother’s years of schooling  
230 measured by this questionnaire.
- 231 • *The Raven's Colored Progressive Matrices* (CPM)–Special Scale (53, 54). We used this test to  
232 assess non-verbal fluid intelligence, analogical reasoning, and the ability to generate responses  
233 from non-verbal stimuli in participating children. This test is composed by 36 items  
234 dichotomously corrected (score of 0 for incorrect and 1 for correct answers), with a total score  
235 of 36. Normative data is given in percentiles, attributed by age group. Children with non-verbal  
236 intelligence falling under percentile 25 were excluded from the sample.

### 237 Subject variables

- 238 • *Strengths and Difficulties Questionnaire* (SDQ)–parent version (50, 51), Brazilian version  
239 (55). The SQD provides an assessment of behavior, emotions, and relationships of children and  
240 adolescents. Questions are grouped into five dimensions: emotional symptoms, conduct  
241 problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior. Each  
242 item is scored between 0 to 2, according to respondents’ agreement with the statements (*e.g.*,  
243 if the mother/parent thinks that the statement “Restless, overactive, cannot stay still for long”  
244 is certainly true about his/her kid, the item is scored with 2 for the hyperactivity-inattention

245 dimension). Some items present opposite directional relationship with the construct and must  
246 be reversed before scoring. Each dimension is scored on a scale of 0-10. The higher the score,  
247 the higher the behavioral difficulty or prosocial behavior. Difficulty items are summed up into a  
248 difficulty score varying from 0 to 40. For selecting children at behavioral risk, we used the  
249 cutoffs of difficulties scores  $\geq 10$ , but the prosocial behavior scores varied across the sample.

- 250 • *Digit-span subtest for children from the WISC-IV* (56, 57). In this test, children are presented  
251 with increasingly long strings of numbers, composed by 2 to 9 items each (*e.g.*, 2-9-7). Upon  
252 hearing the sequence, the child must repeat the digits in the same order given (short term  
253 memory task) or in inverse order (working memory task). Each sequence size presents two  
254 different trials, and the child receives a score of 1 for each sequence correctly uttered. The  
255 total score in each task is 16. The test is discontinued when the child makes an error on two  
256 trials in a same string of numbers. In this study, we used the span score in reverse order, or the  
257 longest sequence numbers kept by the child in memory, as this subtest has been recognized as  
258 a robust measure of working memory (58).
- 259 • *Peer Aggressive and Reactive Behaviors Questionnaire* (PARB-Q/QCARP) (59), Brazilian  
260 version (60). This self-report questionnaire measures behaviors in terms of active and reactive  
261 aggressions among pairs (59). This instrument contains 20 questions and contains two scales:  
262 the Peer Aggression Scale (PA) and the Reaction to Peer Aggression Scale (RPA). The PA  
263 consists of five items that assess proactive (or instrumental) physical and verbal aggression  
264 (*e.g.*, kicking or slapping) and three social-desirability bias items that are not scored (*e.g.*,  
265 telling jokes). The questions begin with “how often do you...,” followed by the circumstances  
266 to be answered. The RPA consists of 12 sentences containing different ways reported by  
267 children about how they react to aggression from their peers. Questions like “when a  
268 colleague of yours...,” are followed by statements, which may signal reactive aggression (RA; 6  
269 items, *e.g.*, “hit or push you,” “do you hit your colleague?”), seeking teacher support (STS; 3  
270 items, *e.g.*, “hit or push you, do you tell the teacher?”), as well as internalized emotional

271 reactions (IR; 3 items, e.g., “hit or push you, do you cry and sulk?”). Each question is rated on a  
272 4-point Likert scale (every day, sometimes, a few times, and never). PA scores range from 0 to  
273 15, and RPA scores range from 0-18 in the RA scale and 0-9 in the STS and IR scales.

274 • *Bateria Psicológica para Avaliação da Atenção* (BPA; Psychological Battery for Attention  
275 Assessment) (61). The BPA is a pencil-and-paper test divided into three subtests that assess  
276 different domains of attention: sustained attention (in one object among distractors), divided  
277 attention (focus in three objects among distractors), and alternated attention (alternating  
278 focus between different objects among distractors).

279 The entire test battery was completed in 40-50 minutes, depending on each child's  
280 performance. Children were allowed to take breaks during the testing sessions.

## 281 Sample matching

282 For the baseline measures, the groups differed significantly in terms of age, reactive aggression  
283 (QCARP), non-verbal intelligence, and mother’s years of schooling. As depicted in S1 Table,  
284 children in the music group tended to be older, presented higher levels of reactive aggression  
285 and non-verbal intelligence, and their mothers had higher levels of education. Moreover, even  
286 though sex did not significantly differ between the groups, the proportion of boys was higher  
287 in the control group. As the groups differed in terms of these characteristics, comparisons of  
288 the effects of music on children’s behavior would be biased. Therefore, we managed the  
289 sample characteristics to match study groups.

290 Sample matching was carried out based on the removal of children from both groups  
291 considering simultaneously the following criteria, respectively, in increasing order of  
292 importance: (1) older children from the experimental and younger from the control group; (2)  
293 higher levels of reaction to aggression in the experimental group and lower in the control  
294 group; (3) superior and average superior performance in the Raven test in the experimental  
295 group and lower performance in the Raven test in the control group. Given that mother’s years

296 of education is not a subject variable but rather related to it, we opted for removing children  
 297 whose mothers had more years of schooling from the experimental group, and the opposite in  
 298 the control group. Following this procedure, we removed 6 children (3 from each group) and  
 299 their results in these variables are depicted in S2 Table. With the matching, the final sample  
 300 was composed by 92 children (experimental = 35; control = 57). Table 1 shows the comparison  
 301 between groups in the baseline measures after matching.

302  
 303

304 **Table 1. Comparison between groups in terms of baseline characteristics after matching ( $n = 92$ ).**

Variables	Group		Comparison
	Experimental	Control	
<i>N</i>	35	57	-
Age (years)	5-8 (M = 6.51; sd = 0,61)	6-7 (M = 6.32; sd = 0.50)	$t(90) = 1.752, p = 0.083$
SDQ			
Difficulties (total score)	15.49 (4.53)	14.70 (4.39)	$t(90) = 0.821, p = 0.400$
Hyperactivity/Inattention	5.66 (2.25)	5.25 (2.28)	$t(90) = 0.845, p = 0.083$
Peer relationship problems	2,57 (1.70)	2.49 (1.65)	$t(90) = 0.224, p = 0.823$
Conduct problems	2.71 (1.62)	2.88 (2.01)	$t(90) = -0.404, p = 0.687$
Emotional Symptoms	4.54 (2.53)	4.09 (2.06)	$t(90) = 0.942, p = 0.349$
Prosocial behavior (free variation)	8.11 (1.55)	8.18 (0.47)	$t(90) = -0.164, p = 0.870$
QCARP			
Direct aggression	13.22 (2.25)	12.26 (2.91)	$t(90) = 1.676, p = 0.097$
Reactive aggression	15.54 (3.76)	14.00 (4.26)	$t(90) = 1.806, p = 0.074$
Internalizing behavior	5.09 (3.02)	4.26 (3.09)	$t(90) = 1.251, p = 0.214$
Seeking teacher support	2.71 (2.36)	3.16 (3.34)	$t(90) = -0.687, p = 0.494$
WM (reverse span)	1.97 (0.99)	1.63 (1.28)	$t(90) = 1.347, p = 0.181$
Attention			
Sustained Attention	30.54 (27.43)	24.31 (19.28)	$t(90) = 1.227, p = 0.205$
Divided Attention	14.17 (34.78)	7.01 (40.20)	$t(90) = 0.871, p = 0.386$
Alternated Attention	27.34 (37.10)	24.31 (18.11)	$t(90) = 0.525, p = 0.601$
Sex (boys)	54.29%	68.42%	$\chi^2(1) = 0.173, p = 0.127$
Raven (classification)			$\chi^2(3) = 7.600, p = 0.055$
Superior	14.29%	3.51%	
Medium superior	31.43%	17.54%	
Medium	37.14%	45.61%	
Medium Inferior	14.14%	33.33%	$\chi^2(4) = 25.664, p < 0.001$
Mother's schooling			
Illiterate	2.86%	1.75%	
Elementary school (incomplete/complete)	0.0%	12.28%	

Variables	Group		Comparison
	Experimental	Control	
High school (not complete)	8.57%	17.54%	
High school (complete)	34.29%	59.65%	
Undergraduate (complete)	54.29%	8.77%	

305 Standard deviations are in parentheses. Chi-square ( $\chi^2$ ) was used for comparing nominal/ordinal  
306 variables and *t*-test for comparing metric variables. SDQ = strengths and difficulties questionnaire.  
307 QCARP = Peer Aggressive and Reactive Behaviors Questionnaire. WM = working memory.

308

309

## 310 Statistical analysis

311 To deal with missing data, imputation was performed. Percent missing for each  
312 individual variable varied from 7.1% (times of execution of the Raven's test; not treated here)  
313 from 15.31 (SDQ at the post-test). We replaced missing data using the expectation  
314 maximization method with 25 iterations. Little's MCAR test suggested that missing data was at  
315 random ( $\chi^2(147) = 133.214, p = 0.786$ ). We paired the groups using pre-test data (before  
316 running the analyses, using data from all children who were paired). By imputing missing data,  
317 we did not remove any subjects. Everyone who underwent "treatment" –music classes–,  
318 whether absent or present, was part of the clinical group (intention to treat–ITT).

319 Descriptive statistics—means, standard deviations, and mean differences (between FUP  
320 and baseline measures) as well as statistical graphic representations are reported. Inferential  
321 analyses are reported by ANOVAS with repeated measures, using group as a fixed factor.  
322 Because maternal levels of schooling differed between the groups, it was included as  
323 covariate. Interactions between factors, groups, and maternal schooling levels were  
324 performed. Data was analyzed using SPSS 23.0 for windows and we consider significant the  
325 results with  $p < 0.05$ . For effect size,  $\eta^2$  is reported, in which: small ( $\eta^2 < 0.01$ ), medium ( $\eta^2$   
326 between 0.02 and 0.06) and large ( $\eta^2 > 0.14$ ) effects (Cohen, 2013 [1988]).

327

## 328 Results

### 329 Behavioral and cognitive results

330 Table 2 presents the descriptive statistics in the FUP (mean, standard deviation, and  
 331 means differences between FUP and baseline results). Polled results are also presented  
 332 (imputation).

333

334

335 **Table 2. Descriptive statistics of the follow-up measures (pooled means, standard deviation, means**  
 336 **differences from baseline)**

Variables	Experimental		Control	
	Follow-up	Mean difference (T2-T1)	Follow-up	Mean difference (T2 -T1)
Age (years)	5-9 (M = 6.51; sd = 0.61)	0.475	6-8 (M = 6.99; sd = 0.75)	0.437
SDQ				
Difficulties (total score)	12.40 (4.85)	-3.08	13.66 (6.58)	-1.05
Hyperactivity/Inattention	5.66 (2.25)	-1.35	4.53 (2.63)	-0.71
Peer relationship problems	2.13 (1.89)	-0.44	2.16 (1.76)	-0.33
Conduct problems	2.27 (1.41)	-0.44	2.66 (2.02)	-0.22
Emotional Symptoms	3.70 (2.15)	-0.85	4.30 (2.53)	0.22
Prosocial behavior (free variation)	8.06 (1.75)	-0.05	8.31 (2.13)	0.14
QCARP				
Direct aggression	13.66 (1.66)	0.43	13.45 (2.41)	1.21
Reactive aggression	16.12 (2.58)	0.58	15.83 (3.86)	1.83
Internalizing behavior	4.98 (2.91)	-0.10	4.40 (2.71)	0.13
Seeking teacher support	2.12 (2.55)	-0.59	2.27 (2.18)	-0.88
WM (reverse span)	2.69 (0.81)	0.72	2.40 (0.97)	0.77
Attention				
Sustained Attention	40.63 (19.30)	10.09	35.84 (20.82)	11.52
Divided Attention	33.69 (27.80)	19.52	23.86 (38.00)	16.84
Alternated Attention	47.16 (16.07)	19.82	38.77 (15.98)	14.46

337 Standard deviations are in parentheses. T2 represents follow-up results and T1 baseline results. SDQ =  
 338 strengths and difficulties questionnaire. QCARP = Peer Aggressive and Reactive Behaviors Questionnaire.  
 339 WM = working memory.

340

341

342 A repeated-measures ANOVA was conducted to verify the impact of the musical  
 343 intervention on different variables of the study, considering as factor among subjects the  
 344 group (control, experimental), schooling as covariate, and observing interactions. For SDQ  
 345 analyzes, difficulties (*i.e.*, hyperactivity, relationship problems, conduct problems, emotional  
 346 symptoms, difficulties in general) and prosocial behavior were considered as dependent  
 347 variables. A reduction in difficulties and an increase in prosocial behavior were the expected  
 348 outcomes to verify the impact of music education. Baseline and FUP comparisons (ANOVAS)  
 349 are shown in Table 3.

350

351 **Table 3. Repeated measures comparisons (main effect and interactions) for the behavioral measures**  
 352 **(SDQ)<sup>a</sup>**

Variable (SDQ)	Main effect	Interaction (mother's schooling)	Interaction (group)
Difficulties (total score)	$F(1,89) = 1.062, p = 0.306, \eta^2 = 0.012$	$F(1,89) = 0.047, p = 0.829, \eta^2 = 0.001$	$F(1,89) = 3.179, p = 0.078, \eta^2 = 0.034$
Hyperactivity/Inattention	$F(1,89) = 0.461, p = 0.499, \eta^2 = 0.005$	$F(1,89) = 0.062, p = 0.084, \eta^2 = 0.001$	$F(1,89) = 1.185, p = 0.279, \eta^2 = 0.013$
Peer relationship problems	<b><math>F(1,89) = 6.615, p = 0.012, \eta^2 = 0.069</math></b>	<b><math>F(1,89) = 4.431, p = 0.0338, \eta^2 = 0.047</math></b>	$F(1,89) = 1.280, p = 0.261, \eta^2 = 0.014$
Conduct problems	$F(1,89) = 0.048, p = 0.828, \eta^2 = 0.001$	$F(1,89) = 0.376, p = 0.542, \eta^2 = 0.004$	$F(1,89) = 0.108, p = 0.744, \eta^2 = 0.001$
Emotional Symptoms	$F(1,89) = 0.014, p = 0.907, \eta^2 = 0.000$	$F(1,89) = 0.040, p = 0.841, \eta^2 = 0.000$	<b><math>F(1,89) = 4.562, p = 0.035, \eta^2 = 0.049</math></b>
Prosocial behavior (free variation)	$F(1,89) = 0.164, p = 0.687, \eta^2 = 0.002$	$F(1,89) = 0.230, p = 0.632, \eta^2 = 0.003$	$F(1,89) = 0.568, p = 0.442, \eta^2 = 0.007$

353 <sup>a</sup>SDQ = strengths and difficulties questionnaire.

354

355

356 As seen in Table 3, the main effect was significant for peer relationship problems,  
 357 indicating that both groups reduced these symptoms between baseline and FUP testing (see  
 358 difference means in Table 2). There was a significant interaction between peer relationship  
 359 problems and maternal schooling levels; the decrease in relationship problems was greater  
 360 depending on the schooling level of the mother, irrespective of group. Post hoc comparisons  
 361 showed that children whose mothers had medium to high years of education tended to start



362 the study with higher levels of peer relationship problems, compared to children whose  
 363 mothers had lower levels of schooling. Yet, at FUP their scores were equivalent to those of  
 364 children with less educated mothers (see S3 Table). The decrease in scores for relationship  
 365 problems was greater for children with mothers who did not complete high school, although  
 366 the lowest score for relationship problems (mean 1.19) came from children with mothers who  
 367 had attended elementary school. In terms of group interactions, significance was found only  
 368 for emotional symptoms. While mean scores for emotional symptoms decreased for the  
 369 experimental group at FUP, they had a slight increase for controls. As seen in Table 2  
 370 (difference means from baseline to FUP) and Fig 2, emotional symptoms fell among the  
 371 experimental group and tended to increase slightly in the control group.

372 **Fig 2. Comparative charts of SDQ symptoms and behaviors in the baseline and follow up.** Solid lines  
 373 represent the experimental group and dashed lines the control group. Scales are not standardized and  
 374 represents estimated means generated for the program by default.

375

376

377 Comparisons between groups for direct aggressive behavior and reaction to aggression  
 378 assessed by the QCARP are shown in Table 4. No significant effects were found for any of the  
 379 aggression measures.

380 **Table 4. Repeated measures comparisons (main effect and interactions) for the behavioral measures**  
 381 **(aggression – QCARP)**

Variable (QCARP)	Main effect	Interaction (mother's schooling)	Interaction (group)
Direct aggression	$F(1,89) = 0.229, p = 0.633, \eta^2 = 0.003$	$F(1,89) = 0.014, p = 0.905, \eta^2 = 0.000$	$F(1,89) = 1.529, p = 0.220, \eta^2 = 0.017$
Reactive aggression	$F(1,89) = 1.735, p = 0.633, \eta^2 = 0.003$	$F(1,89) = 0.014, p = 0.905, \eta^2 = 0.000$	$F(1,89) = 1.529, p = 0.220, \eta^2 = 0.017$
Internalizing behavior	$F(1,89) = 0.092, p = 0.762, \eta^2 = 0.001$	$F(1,89) = 0.102, p = 0.750, \eta^2 = 0.001$	$F(1,89) = 0.173, p = 0.678, \eta^2 = 0.002$
Seeking teacher support	$F(1,89) = 1.061, p = 0.306, \eta^2 = 0.012$	$F(1,89) = 0.358, p = 0.551, \eta^2 = 0.004$	$F(1,89) = 0.021, p = 0.884, \eta^2 = 0.000$

382 QCARP = Peer Aggressive and Reactive Behaviors Questionnaire.

383

384

385 Fig 3 depicts the baseline and FUP results for these measures.

386 **Fig 3. Comparative charts of aggressive behavior and reaction to aggression evaluated by the QCARP**  
 387 **in the baseline and follow-up.** Solid lines represent the experimental group and dashed lines the control  
 388 group. Scales are not standardized and represents estimated means generated for the program by  
 389 default.

390

391

392 The results for verbal working memory assessed by digits span (reverse order) and the  
 393 attentional measures are shown in Table 5 (and their respective descriptive statistics in Table  
 394 2).

395 **Table 5. Repeated measures comparisons (main effect and interactions) for the cognitive measures**

Variable	Main effect	Interaction (mother's schooling)	Interaction (group)
WM <sup>b</sup> (reverse span)	$F(1,89) = 2.443, p = 0.122, \eta^2 = 0.027$	$F(1,89) = 0.072, p = 0.789, \eta^2 = 0.001$	$F(1,89) = 0.006, p = 0.936, \eta^2 = 0.000$
Sustained Attention	$F(1,89) = 2.985, p = 0.088, \eta^2 = 0.032$	$F(1,89) = 0.885, p = 0.349, \eta^2 = 0.010$	$F(1,89) = 0.020, p = 0.888, \eta^2 = 0.000$
Divided Attention	<b><math>F(1,89) = 14.221, p &lt; 0.001, \eta^2 = 0.032</math></b>	<b><math>F(1,89) = 5.476, p = 0.006, \eta^2 = 0.083</math></b>	$F(1,89) = 2.000, p = 0.161, \eta^2 = 0.022$
Alternated Attention	<b><math>F(1,89) = 9.853, p = 0.002, \eta^2 = 0.100</math></b>	$F(1,89) = 2.605, p = 0.110, \eta^2 = 0.028$	$F(1,89) = 2.747, p = 0.101, \eta^2 = 0.030$

396 <sup>b</sup>WM = working memory.

397

398

399 As shown in Table 5, there were significant main effects for divided and  
 400 alternated attention. For both groups, attention scores increased from baseline to  
 401 FUP. There was a significant interaction between maternal schooling levels and divided  
 402 attention. Post hoc analyses showed that children with mothers who had low levels of  
 403 education benefited more from the development of divided attention, with gains  
 404 varying around 0.70 to 0.95 standard deviations (S4 Table). No other significant  
 405 interactions emerged for any of the cognitive variables (Fig 4).

406 **Fig 4. Comparative charts for the cognitive measures (working memory and attention) in the baseline**  
407 **and follow-up.** Solid represent the experimental group and dashed lines the control group. Scales are  
408 not standardized and represents estimated means generated for the program by default.

409

410

## 411 Discussion

412 The aim of this quasi-experimental study was to examine the impact of a music  
413 education program on aspects of attention—shifting, divided, and sustained—, socioemotional  
414 abilities, and reactive and aggressive behaviors in children aged 5-9 from underserved  
415 communities of São Paulo. Using standardized tests, we compared results from a group of  
416 children who participated in a community-based music program to matched controls at two  
417 time points: before the intervention and seven months later. Because maternal education is  
418 known to have an effect on children’s cognitive and socioemotional development (62-64), we  
419 also analyzed maternal education data reported by the children’s parents through the ABEP  
420 socioeconomic status questionnaire (52). To our knowledge, this is likely the first quantitative  
421 study carried out on the impact of sociomusical programs in Brazil.

422 A main finding that emerged from our study was a significant effect of the intervention  
423 on the music group in the FUP. Our data revealed a more pronounced decrease in emotional  
424 symptoms in the music group as opposed to controls. This finding is consistent with earlier  
425 theoretical (65) and empirical studies on the role of music education in children’s mood and  
426 emotional regulation (66, 67). We also found a developmental effect for emotional symptoms,  
427 with children in both groups showing improved scores from baseline to FUP. Emotional  
428 symptoms including depression, anxiety, perceived peer rejection, victimization, and feelings  
429 of loneliness, have been associated with motor competence and to one’s ability to move with  
430 rhythm (67-69). It has been established that motor coordination disorders are risk factors for  
431 mental health problems, such as eating disorders (68, 70, 71). Considering that rhythmic

432 training is intrinsic to music education in the Guri's program, we suspect that it may have been  
433 a protective factor for children's emotional symptoms.

434 In this study, maternal schooling was found to interact with relationship problems and  
435 emotional symptoms in different ways. Children of more educated mothers had more  
436 relationship problems at the start of the study in both groups, but these balanced out after 7  
437 months. Emotional symptoms decreased more markedly in children whose mothers had less  
438 education. Why would maternal education impact these two socioemotional areas in such  
439 distinct ways? Previous studies (63) have found higher maternal education levels (*i.e.*, at least  
440 some college education) to be associated with fewer behavioral problems in children, although  
441 earlier comparisons of different levels of maternal education (*i.e.*, without some college  
442 education) revealed similar results from those of our study. Maternal education has been  
443 linked to numerous practices that grant children cognitive advantages, such as the use of more  
444 complex language (63). As socioemotional symptoms were assessed through parental reports,  
445 it is possible that children with more educated mothers were also more prone to verbalize the  
446 impact of formal learning environments in public schools as opposed to the home  
447 environment in the aftermath of the pandemic. Our study took place immediately after the  
448 social restrictions imposed by the coronavirus pandemic began to be lifted. In Brazil, there also  
449 were marked differences in access to education between families during the COVID-19  
450 quarantine. This is, of course, speculative. By contrast, the finding of a significant interaction  
451 between maternal schooling and children's emotional symptoms is consistent with earlier  
452 studies that found a close relationship between maternal education and children's mental  
453 health (62). Maternal education has been directly associated with low SES and social status,  
454 another factor to have an impact on children's mental health and socioemotional wellbeing  
455 (72), particularly in early childhood (62).

456 In terms of cognitive outcomes, divided attention had a developmental effect,  
457 increasing equally between groups, from baseline to FUP. We also found divided attention to  
458 increase more in children whose mothers had lower levels of education. A possible explanation  
459 for this finding is that in formal educational settings, children from our study were more able  
460 to learn and retain new information, keeping up with other children's skills. During the  
461 pandemic, many children spent long periods of time in front of screens (*e.g.*, TV, tablets,  
462 games, and cell phones), which is known to reduce their attention span (73). Furthermore,  
463 children from low-SES families in Brazil were often left at home unsupervised (74). Alternated  
464 attention, in turn, was higher for the music group, although no causal relationship was  
465 observed. Contrary to earlier studies, we did not find effects of music instruction on children's  
466 attention, possibly due to our small sample size, the short duration of the intervention (*i.e.*, 7  
467 months) and music classes (*i.e.*, only one hour per week), and the unprecedented pandemic  
468 variable.

469

## 470 Conclusion & Limitations

471 Parents of the children selected for this study reported high difficulties in the SDQ,  
472 which suggests a risk of exposure to mental health vulnerabilities. The surveys were carried  
473 out in peripheral regions of São Paulo, Brazil, where services such as public security, health,  
474 education, and transportation are scarce and oftentimes precarious where they exist. The  
475 literature indicates such elements as predictors of emotional symptoms, health problems  
476 related to emotional stress, and behavioral problems in children (75).

477 Given the harmful effect of scarcity on mental health of underserved communities, we  
478 believe it is crucial that public policies pay attention to the potential benefits of music  
479 education on socioemotional symptoms. It is essential that the State, public policy

480 administrators, as well as the third sector represented by social organizations and non-  
481 governmental organizations (NGOs), pay attention to the responsibility of expanding the scope  
482 of their activities to these vulnerable populations. Access to music education can potentially  
483 serve as a protective factor for children exposed to violence and cognitive vulnerabilities.

484 As limitations of this study, the absence of statistically significant interaction between  
485 the groups in some cases may be due to the low sample size and short intervention time. Also,  
486 a question remains unanswered is: Would different curricular approaches to music education  
487 yield different results? Previous studies suggest that curated curricula may yield more  
488 substantial findings in terms of both EFs (76) and socioemotional skills (77). Additionally, the  
489 length of programs is important; the literature points to the potentiation of long-term  
490 exposure to music practice (more than two years) on the impact of behavioral measures and  
491 brain structures (3, 8, 78-80). Our study collected data over a school year, which, in practice,  
492 had an interval between the baseline and FUP of only about seven months, and included an  
493 hour of music class per week, and in the aftermath of a global pandemic. We hope that the  
494 tendency observed here can contribute to future research. More studies are obviously needed.

495

## 496 Acknowledgements

497 We thank all the families and their children who volunteered in this research, as well as the  
498 assistants who collected the data, the Projeto Guri (OS Santa Marcelina), Secretaria Municipal  
499 de Educação de São Paulo, and Secretaria da Cultura e Economia Criativa do Estado de São  
500 Paulo for granting permission for the study to take place.

501

502

## 503 References

504 1. Talamini F, Blain S, Ginzburg J, Houix O, Bouchet P, Grassi M, et al. Auditory and visual short-  
505 term memory: influence of material type, contour, and musical expertise. *Psychological*  
506 *Research*. 2022; 86(2): 421-42.

- 507 2. Ilari B, Helfter S, Huynh T, Bowmer A, Mason K, Knight J, et al. Musical Activities, Prosocial  
508 Behaviors, and Executive Function Skills of Kindergarten Children. *Music & Science*. 2021; 4:  
509 20592043211054829.
- 510 3. Holochwost SJ, Propper CB, Wolf DP, Willoughby MT, Fisher KR, Kolacz J, et al. Music  
511 education, academic achievement, and executive functions. *Psychology of Aesthetics,*  
512 *Creativity, and the Arts*. 2017;11:147-66.
- 513 4. Hennessy S, Sachs M, Ilari B, Habibi A. Effects of Music Training on Inhibitory Control and  
514 Associated Neural Networks in School-Aged Children: A Longitudinal Study. *Frontiers in*  
515 *Neuroscience*. 2019: art. 1080.
- 516 5. Cogo-Moreira H, Ávila CRBd, Ploubidis GB, Mari JdJ. Pathway Evidence of How Musical  
517 Perception Predicts Word-Level Reading Ability in Children with Reading Difficulties. *PLoS One*.  
518 2013; 8(12): 1-10.
- 519 6. Zuk J, Benjamin C, Kenyon A, Gaab N. Behavioral and neural correlates of executive  
520 functioning in musicians and non-musicians. *PLoS One*. 2014; 9(6): e99868.
- 521 7. Talamini F, Carretti B, Grassi M. The Working Memory of Musicians and Nonmusicians.  
522 *Music Perception*. 2016; 34(2): 183-91.
- 523 8. Talamini F, Altoè G, Carretti B, Grassi M. Musicians have better memory than nonmusicians:  
524 A meta-analysis. *PLoS One*. 2017; 12(10): e0186773.
- 525 9. D'Souza AA, Moradzadeh L, Wiseheart M. Musical training, bilingualism, and executive  
526 function: working memory and inhibitory control. *Cogn Res Princ Implic*. 2018; 3(1): 11.
- 527 10. Chen JJ, Scheller M, Wu CY, Hu BY, Peng R, Liu CH, et al. The relationship between early  
528 musical training and executive functions: Validation of effects of the sensitive period.  
529 *Psychology of Music*. 2022; 50(1): 86-99.
- 530 11. Alemán X, Duryea S, Guerra NG, McEwan PJ, Muñoz R, Stampini M, et al. The Effects of  
531 Musical Training on Child Development: a Randomized Trial of El Sistema in Venezuela.  
532 *Prevention Science*. 2017; 18: 865–78.
- 533 12. Rodriguez-Gomez DA, Talero-Gutierrez C. Effects of music training in executive function  
534 performance in children: A systematic review. *Front Psychol*. 2022; 13:968144.
- 535 13. Miller E, Cohen J. An Integrative Theory of Prefrontal Cortex Function. *Annual review of*  
536 *neuroscience*. 2001; 24: 167-202.
- 537 14. Miller DC, Maricle DE, Kaufman AS, Kaufman NL. *Essentials of School Neuropsychological*  
538 *Assessment*: Wiley; 2019.
- 539 15. Mehse V, Morag L, Chesta S, Cleaton K, Burgos H. Hot Executive Function Assessment  
540 Instruments in Preschool Children: A Systematic Review. *Int J Environ Res Public Health*. 2021;  
541 19(1).
- 542 16. Lezak MD, Howieson DB, Bigler ED, Tranel D. *Neuropsychological Assessment*. Oxford:  
543 Oxford University Press; 2012.
- 544 17. Diamond A. Executive functions. *Annual Review of Psychology*. 2013; 64: 135-68.
- 545 18. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and  
546 diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a  
547 latent variable analysis. *Cogn Psychol*. 2000; 41(1): 49-100.

- 548 19. Downing KM. An examination of three theoretical models of executive functioning.  
549 Denton, Texas: Texas Woman's University; 2015.
- 550 20. Zimmerman DL, Owsnworth T, O'Donovan A, Roberts J, Gullo MJ. Independence of Hot and  
551 Cold Executive Function Deficits in High-Functioning Adults with Autism Spectrum Disorder.  
552 *Frontiers in Human Neuroscience*. 2016; 10: art. 24.
- 553 21. Grafman J, Litvan I. Importance of deficits in executive functions. *Lancet*. 2000; 354: 1921-  
554 3.
- 555 22. Posner MI, Petersen SE. The Attention System of the Human Brain. *Annual Review of*  
556 *Neuroscience*. 1990; 13(1): 25-42.
- 557 23. Petersen SE, Posner MI. The Attention System of the Human Brain: 20 Years After. *Annual*  
558 *Review of Neuroscience*. 2012; 35(1): 73-89.
- 559 24. Christensen JF. Pleasure junkies all around! Why it matters and why 'the arts' might be the  
560 answer: a biopsychological perspective. *Proc Biol Sci*. 2017; 284(1854).
- 561 25. Wager TD, Jonides J, Reading S. Neuroimaging studies of shifting attention: a meta-  
562 analysis. *Neuroimage*. 2004; 22(4): 1679-93.
- 563 26. Esterman M, Rothlein D. Models of sustained attention. *Curr Opin Psychol*. 2019; 29: 174-  
564 80.
- 565 27. Cowan N. Working memory capacity. New York, NY, US: Psychology Press; 2005. 246 p.
- 566 28. Chan RCK, Shum D, Touloupoulou T, Chen EYH. Assessment of executive functions: Review  
567 of instruments and identification of critical issues. *Arch Clin Neuropsych*. 2008; 23(2): 201-16.
- 568 29. McDonald S. Impairments in social cognition following severe traumatic brain injury. *J Int*  
569 *Neuropsychol Soc*. 2013; 19(3): 231-46.
- 570 30. Dixon ML, Thiruchselvam R, Todd R, Christoff K. Emotion and the prefrontal cortex: An  
571 integrative review. *Psychol Bull*. 2017; 143(10): 1033-81.
- 572 31. Bechara A, Damasio H, Damasio AR, Lee GP. Different Contributions of the Human  
573 Amygdala and Ventromedial Prefrontal Cortex to Decision-Making. *The Journal of*  
574 *Neuroscience*. 1999; 19(13): 5473.
- 575 32. Bechara A, Damasio H, Damasio AR. Emotion, Decision Making and the Orbitofrontal  
576 Cortex. *Cerebral Cortex*. 2000; 10(3): 295-307.
- 577 33. Rubin KH, Rose-Krasnor L. Interpersonal problem solving and social competence in  
578 children. In: Hersen VBVM, editor: Plenum Press; 1992. p. 283-323.
- 579 34. Hukkelberg S, Keles S, Ogden T, Hammerstrøm K. The relation between behavioral  
580 problems and social competence: A correlational Meta-analysis. *BMC Psychiatry*. 2019; 19(1):  
581 354.
- 582 35. Milligan K, Sibalis A, Morgan A, Phillips M. Social competence: consideration of  
583 behavioural, cognitive, and emotional factors. Matson J, editor: Cham: Springer Cham; 2017.  
584 301-20 p.
- 585 36. Sørli MA, Hagen K, Nordahl KB. Development of social skills during middle childhood:  
586 growth trajectories and school-related predictors. *Int J Sch Educ Psychol*. 2021; 9: S69-S87.
- 587 37. Little SG, Swangler J, Akin-Little A. Defining social skills. In: Matson J, editor.: Cham:  
588 Springer; 2017. p. 9-17.



- 589 38. Boal-Palheiros G, Ilari B. Music, drama, and social development in Portuguese children.  
590 Front Psychol. 2023; 14: 1093832.
- 591 39. Gresham FM, Elliott SN, Vance MJ, Cook CR. Comparability of the social skills rating system  
592 to the social skills improvement system: content and psychometric comparisons across  
593 elementary and secondary age levels. Sch Psychol. 2011; Q. 26: 27–44.
- 594 40. Schellenberg EG, Corrigan KA, Dys SP, Malti T. Group Music Training and Children's  
595 Prosocial Skills. PLOS ONE. 2015; 10(10): e0141449.
- 596 41. Buren V, Degé F, Schwarzer G. Active music making facilitates prosocial behaviour in 18-  
597 month-old children. Musicae Scientiae. 2019; 25(4): 449-64.
- 598 42. Beck SL, Rieser J. Non-random acts of kindness: Joint music making increases preschoolers'  
599 helping and sharing with an adult. Psychology of Music. 2020; 50(1): 17-33.
- 600 43. Kirschner S, Ilari B. Joint Drumming in Brazilian and German Preschool Children: Cultural  
601 Differences in Rhythmic Entrainment, but No Prosocial Effects. Journal of Cross-Cultural  
602 Psychology. 2013; 45(1): 137-66.
- 603 44. Ilari B, Fesjian C, Habibi A. Entrainment, theory of mind, and prosociality in child musicians.  
604 Music & Science. 2018; 1: 2059204317753153.
- 605 45. Villanueva J, Ilari B, Habibi A. Long-term music instruction is partially associated with the  
606 development of socioemotional skills. PLOS ONE. 2024; 19(7): e0307373.
- 607 46. Ilari B, Perez P, Wood A, Habibi A. The role of community-based music and sports  
608 programmes in parental views of children's social skills and personality. International Journal  
609 of Community Music. 2019; 12(1): 35-56.
- 610 47. Henrich J. The WEIRD People in the World: How the West Became Psychologically  
611 Peculiar and Particularly Prosperous: Farrar, Straus and Giroux; 2020.
- 612 48. Guri. Music and Culture Inspiring Lives: Musical education, cultural development and  
613 human formation for children and young people in the state of São Paulo. 2024 [Available  
614 from: <https://www.projetoguri.org.br>.
- 615 49. Cristal QR, editor O Processo da Musicalização: concepções e implicações práticas. Anais  
616 do XIV Encontro Regional Nordeste da Associação Brasileira de Educação Musical - ABEM; 2018  
617 Sept., 2018; Salvador, BA: ABEM/UFBA/Nova UCSAL.
- 618 50. Goodman R. The Strengths and Difficulties Questionnaire: a research note. J Child Psychol  
619 Psychiatry. 1997; 38(5): 581-6.
- 620 51. Goodman A, Goodman R. Population mean scores predict child mental disorder rates:  
621 validating SDQ prevalence estimators in Britain. J Child Psychol Psychiatry. 2011; 52(1): 100-8.
- 622 52. ABEP ABdEdP-. Critério de classificação econômica Brasil 2020 [updated Acesso em 9 de  
623 outubro de 2020. Available from: <http://www.abep.org/criterio-brasil>
- 624 53. Raven J, Raven JC, Court JH. Matrizes Progressivas Coloridas de Raven: Manual. São  
625 Paulo 2003 2018.
- 626 54. Raven J, Raven J. Raven Progressive Matrices. Handbook of nonverbal assessment. New  
627 York, NY, US: Kluwer Academic/Plenum Publishers; 2003. p. 223-37.
- 628 55. Stivanin L, Scheuer CI, Assumpção Jr FB. SDQ (Strengths and Difficulties Questionnaire):  
629 identificação de características comportamentais de crianças leitoras. Psicologia: Teoria e  
630 Pesquisa. 2008; 24(4): 407-13.

- 631 56. Wechsler D. Escala Wechsler de inteligência para crianças: WISC-IV. Manual Técnico. 4a  
632 ed. São Paulo: Casa do Psicólogo; 2013.
- 633 57. Wechsler D. Wechsler Intelligence Scale for Children (WISC IV). 4a ed. San Antonio: Pro-Ed,  
634 Inc.; 2003.
- 635 58. Egeland J. Measuring Working Memory With Digit Span and the Letter-Number Sequencing  
636 Subtests From the WAIS-IV: Too Low Manipulation Load and Risk for Underestimating Modality  
637 Effects. *Appl Neuropsychol Adult*. 2015; 22(6): 445-51.
- 638 59. Gremigni P. Validazione di un nuovo Questionario Per L'eta' Evolutiva Il Cap  
639 Comportamento Aggressivo Tra Pari Report di Ricerca. . In: *Psicologia*. UDSdBDD, editor. 2006.
- 640 60. Borsa JC, Bandeira DR. Adaptação transcultural do questionário de comportamentos  
641 agressivos e reativos entre pares no Brasil. *PsicoUSF*. 2014; 19(2): 287-96.
- 642 61. Rueda FJM. Coleção BPA - Bateria Psicológica de Atenção. São Paulo: Vetor; 2013.
- 643 62. Meyrose AK, Klasen F, Otto C, Gniewosz G, Lampert T, Ravens-Sieberer U. Benefits of  
644 maternal education for mental health trajectories across childhood and adolescence. *Soc Sci*  
645 *Med*. 2018; 202: 170-8.
- 646 63. Harding JF. Increases in maternal education and low-income children's cognitive and  
647 behavioral outcomes. *Dev Psychol*. 2015; 51(5): 583-99.
- 648 64. Harding JF, Morris P, Hughes D. The relationship between maternal education and  
649 children's academic outcomes: a theoretical framework. *Journal of Marriage and Family*. 2015;  
650 77(1): 60-76.
- 651 65. Ilari BS, Keller P, Damasio H, Habibi A. The Development of Musical Skills of  
652 Underprivileged Children Over the Course of 1 Year: A Study in the Context of an El Sistema-  
653 Inspired Program. *Frontiers in Psychology*. 2016; 7.
- 654 66. Gaudette-Leblanc A, Boucher H, Bédard-Bruyère F, Bolduc J, Pearson J, Tarabulsy G.  
655 Participation in an early childhood music programme and socioemotional development: A  
656 meta-analysis. *International Journal of Music in Early Childhood*. 2022; 16: 131-53.
- 657 67. Williams KE. Moving to the Beat: Using Music, Rhythm, and Movement to Enhance Self-  
658 Regulation in Early Childhood Classrooms. *Int J Early Child*. 2018; 50(1): 85-100.
- 659 68. Medeiros P, Zequinão MA, Silva WR, Maciel IR, Cardoso FL. Behavioral and emotional  
660 problems in children with poor motor skills: According to parental emic perception. *Journal of*  
661 *Physical Education*. 2022; 33.
- 662 69. Green D, Baird G, Sugden D. A pilot study of psychopathology in Developmental  
663 Coordination Disorder. *Child Care Health Dev*. 2006; 32(6): 741-50.
- 664 70. Li YC, Kwan MYW, Cairney J. Motor coordination problems and psychological distress in  
665 young adults: A test of the Environmental Stress Hypothesis. *Res Dev Disabil*. 2019; 84: 112-21.
- 666 71. Cairney J, Rigoli D, Piek J. Developmental coordination disorder and internalizing problems  
667 in children: The environmental stress hypothesis elaborated. *Developmental Review*. 2013;  
668 33(3): 224-38.
- 669 72. Noonan RJ, Fairclough SJ. Social Disadvantage, Maternal Psychological Distress, and  
670 Difficulties in Children's Social-Emotional Well-Being. *Behav Sci (Basel)*. 2018; 8(11).

- 671 73. Moller N, Gomes Junior SC, Marano D, Zin A. Survey of the Adequacy of Brazilian Children  
672 and Adolescents to the 24-Hour Movement Guidelines before and during the COVID-19  
673 Pandemic. *Int J Env Res Pub He.* 2023; 20(9): 5737.
- 674 74. Brito PKH, Soares AR, Bezerra ICdS, Reichert LP, Santos NCCdB, Collet N, et al. Impact of the  
675 Covid-19 pandemic on the use of screens in early childhood. *Revista Gaúcha de Enfermagem.*  
676 2023; 44.
- 677 75. Borsa JC, Nunes MLT. Prevalência de problemas de comportamento em uma amostra de  
678 crianças em idade escolar da cidade de Porto Alegre. *Aletheia.* 2011: 32-46.
- 679 76. Bowmer A, Mason K, Knight J, Welch G. Investigating the Impact of a Musical Intervention  
680 on Preschool Children's Executive Function. *Front Psychol.* 2018; 9: 2389.
- 681 77. Rabinowitch T-C, Cross I, Burnard P. Long-term musical group interaction has a positive  
682 influence on empathy in children. *Psychology of Music.* 2012; 41(4): 484-98.
- 683 78. Hyde KL, Lerch J, Norton A, Forgeard M, Winner E, Evans AC, et al. Musical Training Shapes  
684 Structural Brain Development. *The Journal of Neuroscience.* 2009; 29(10): 3019-25.
- 685 79. Habibi A, Damasio A, Ilari B, Elliott Sachs M, Damasio H. Music training and child  
686 development: a review of recent findings from a longitudinal study. *Ann N Y Acad Sci.* 2018.
- 687 80. Cheng L-K, Chiu Y-H, Lin Y-C, Li W-C, Hong T-Y, Yang C-J, et al. Long-term musical training  
688 induces white matter plasticity in emotion and language networks. *Human Brain Mapping.*  
689 2023; 44(1): 5-17.

690

691

## 692 Supporting information

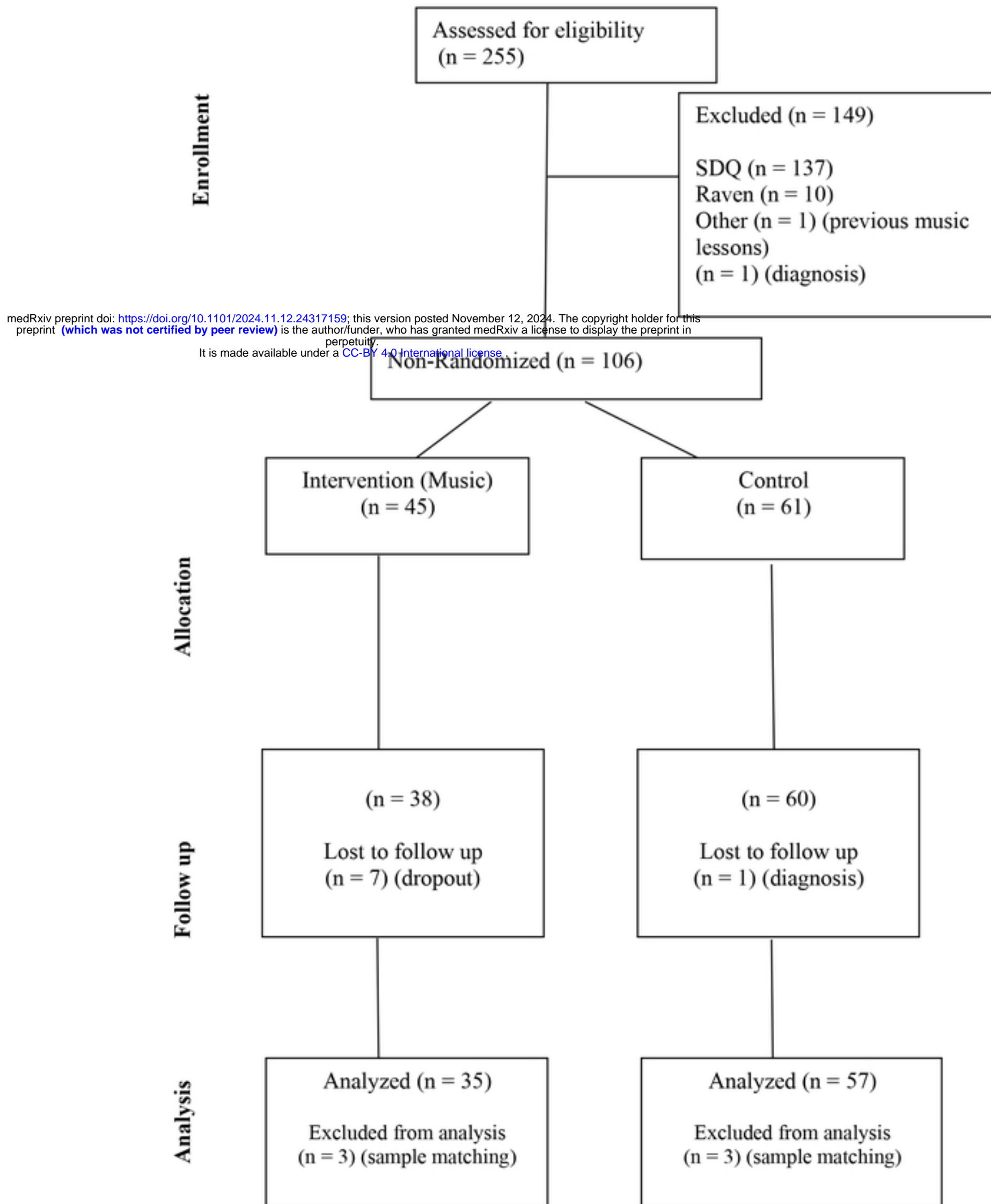
693 **S1 Table. Comparison of groups in terms of baseline characteristics.** *Note.* Standard deviation in  
694 parentheses. Chi-square ( $\chi^2$ ) for comparing nominal/ordinal variables and t-test for comparing metric  
695 variables. SDQ = strengths and difficulties questionnaire. Qcarp = Peer Aggressive and Reactive  
696 Behaviors Questionnaire. PA = active aggression; RA

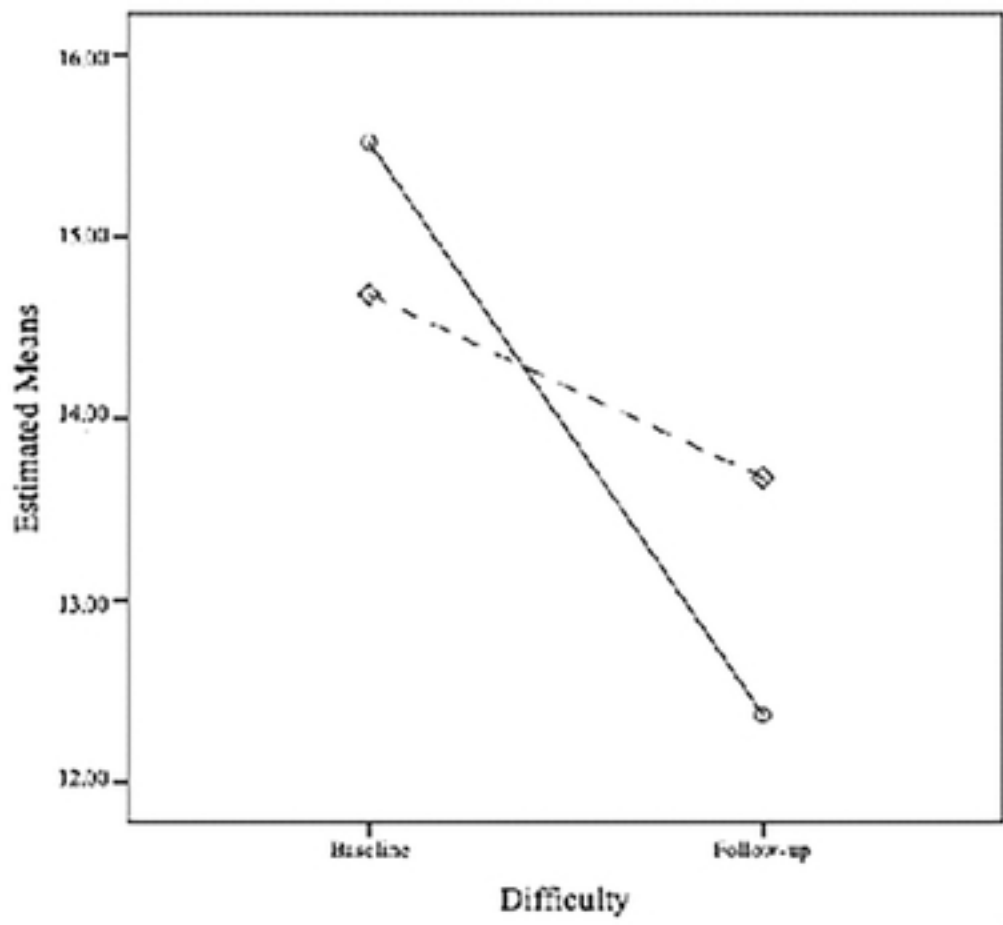
697 **S2 Table. Matching according to age, levels of reaction to aggression and performance in the**  
698 **Raven test.** *Note.* RA = reactive aggression in Q-carp questionnaire. Intelligence in Raven test  
699 was classified as follows: 1 = superior; 2 = Medium superior; 3 = Medium; 4 = Medium Inferior;  
700 5 = Inferior (not present in the sample; removed from inclusion criteria). Mother's schooling  
701 was classified as follows: 1= illiterate; 2 = Elementary school (incomplete/complete); 3 = High  
702 school (not complete); 4 = High school (complete); 5 = Undergraduate (complete).

703 **S3 Table. Descriptive statistics for the peer relationship problems considering mother's level**  
704 **of education.** *d* = Cohen's *d* (0.30 = low; 0.5 moderate; 0.80 high).

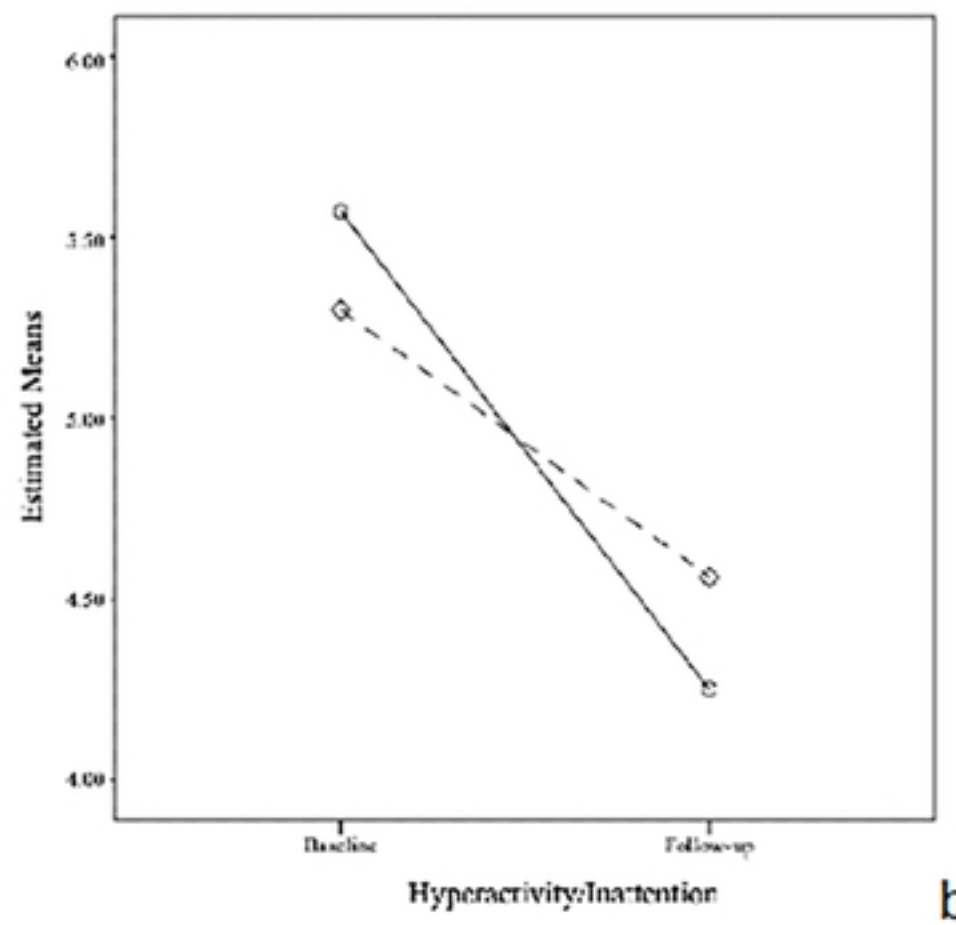
705 **S4 Table. Descriptive statistics for the peer relationship problems considering mother's level**  
706 **of education.** *d* = Cohen's *d* (0.30 = low; 0.5 moderate; 0.80 high).

**Flow of participants through each stage of the trial.**

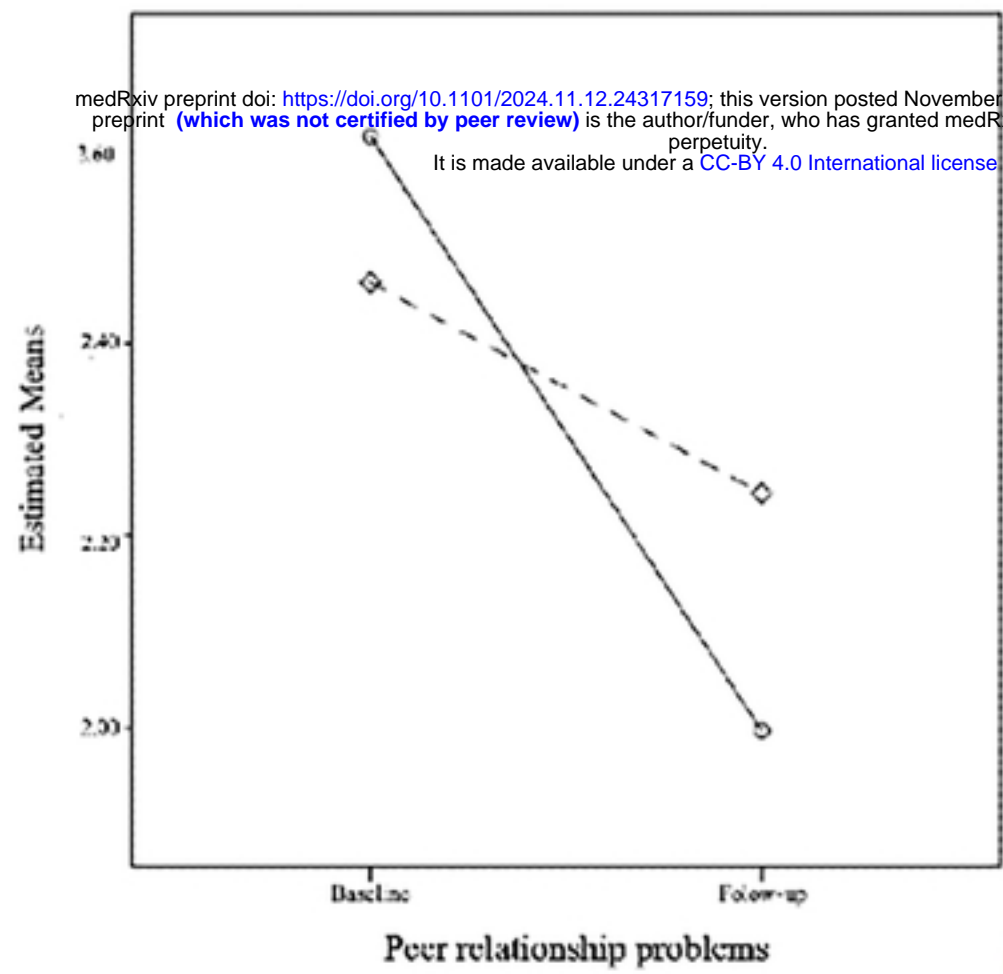




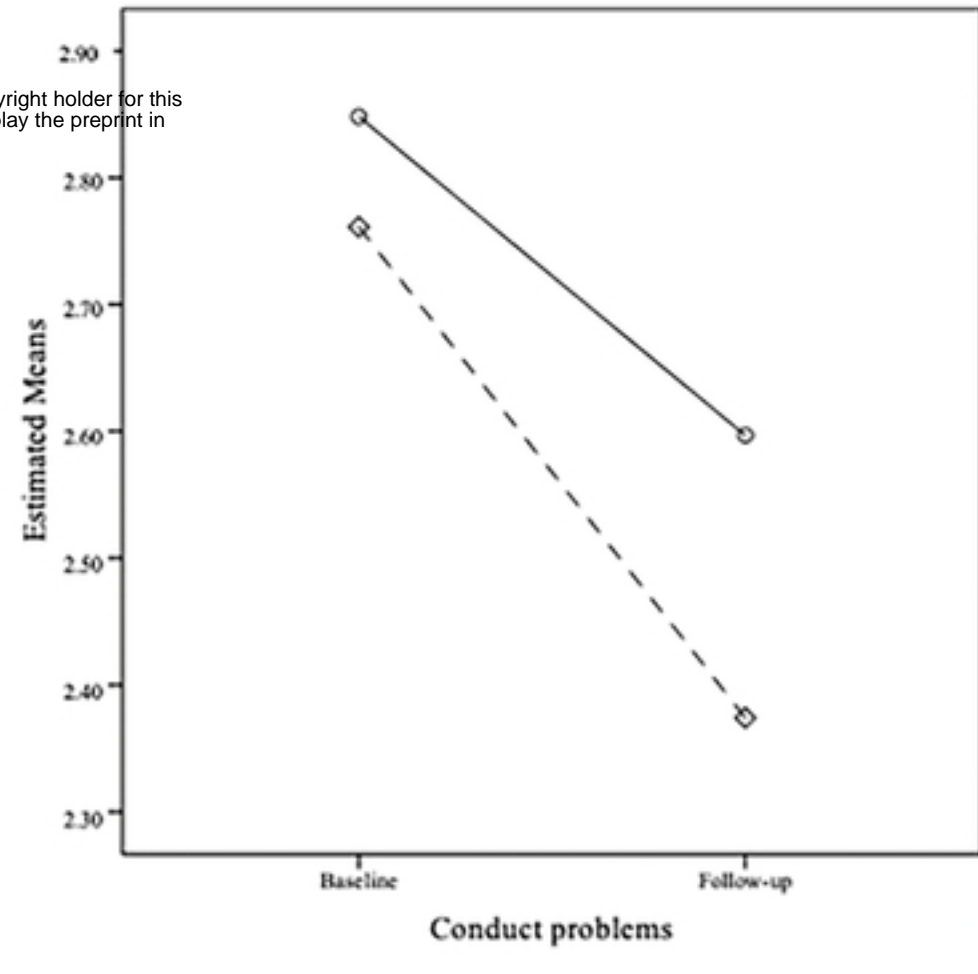
a



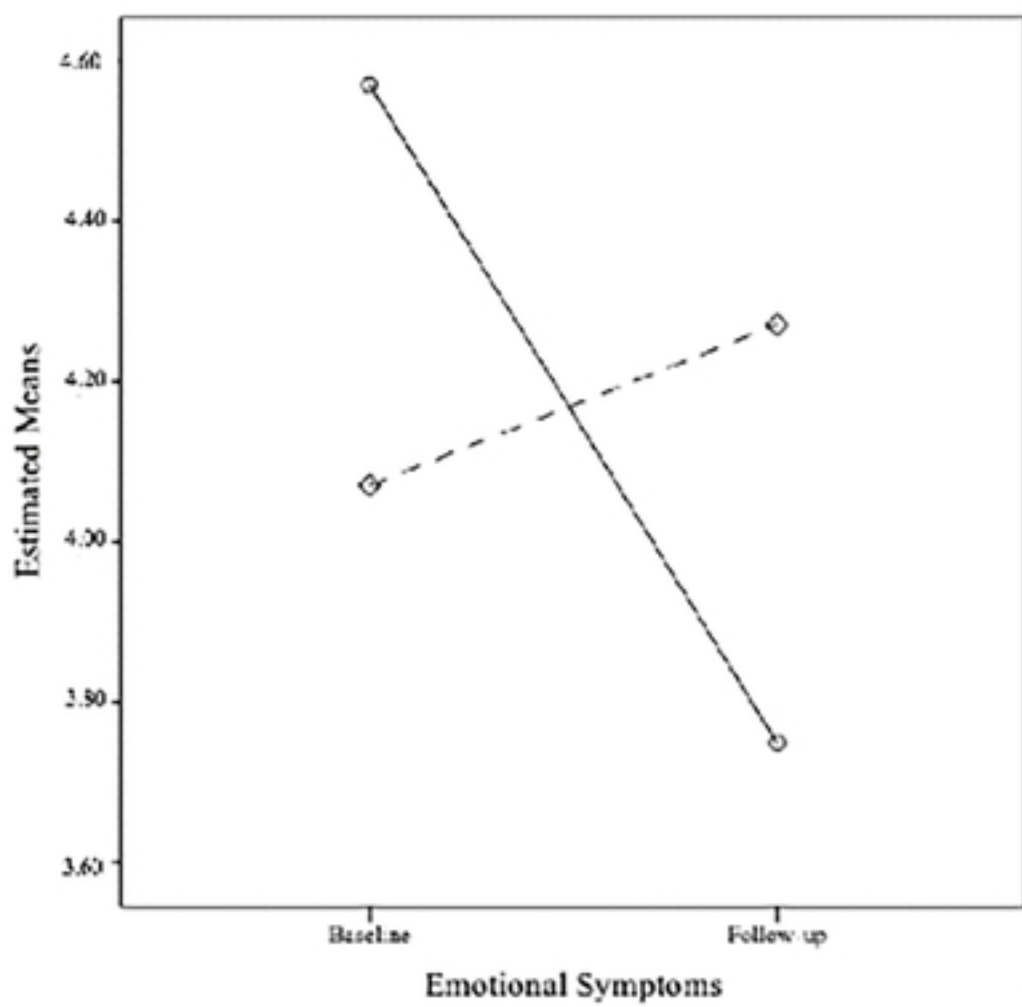
b



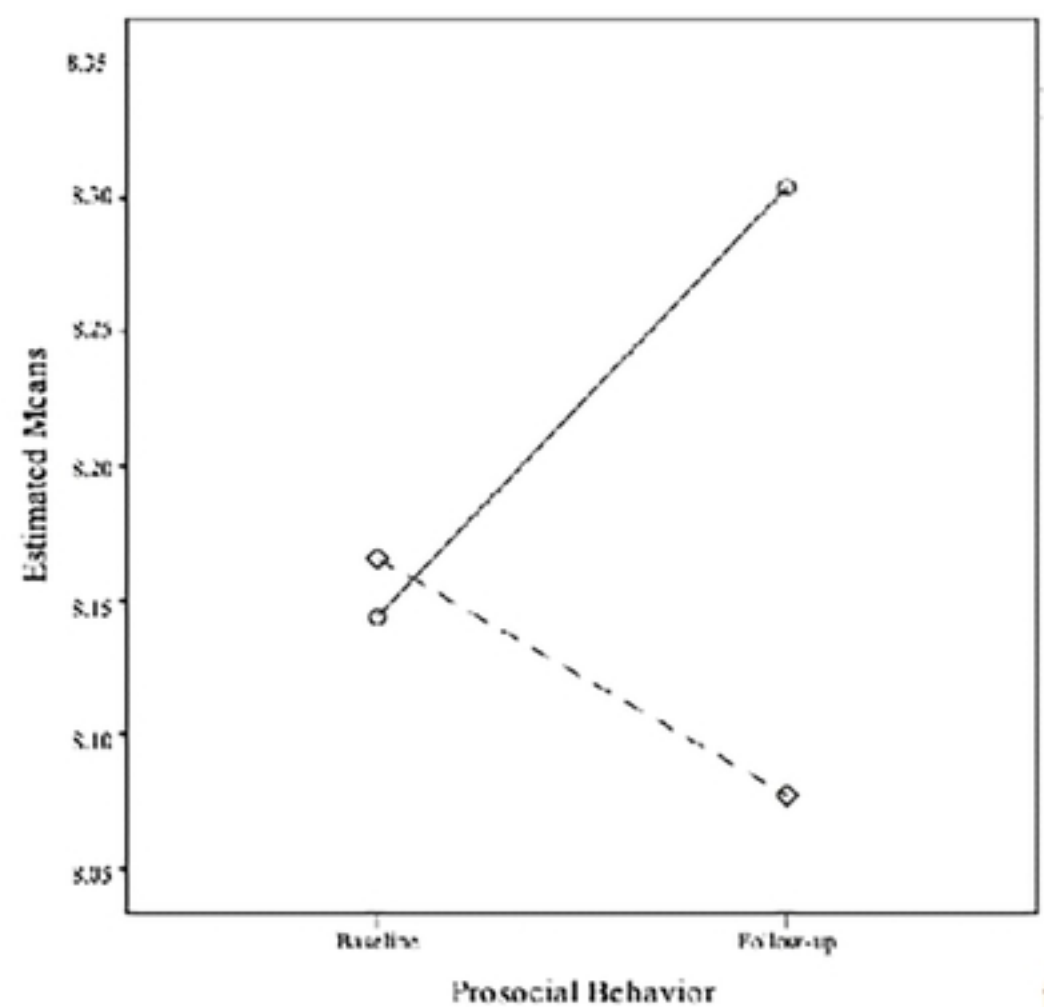
c



d



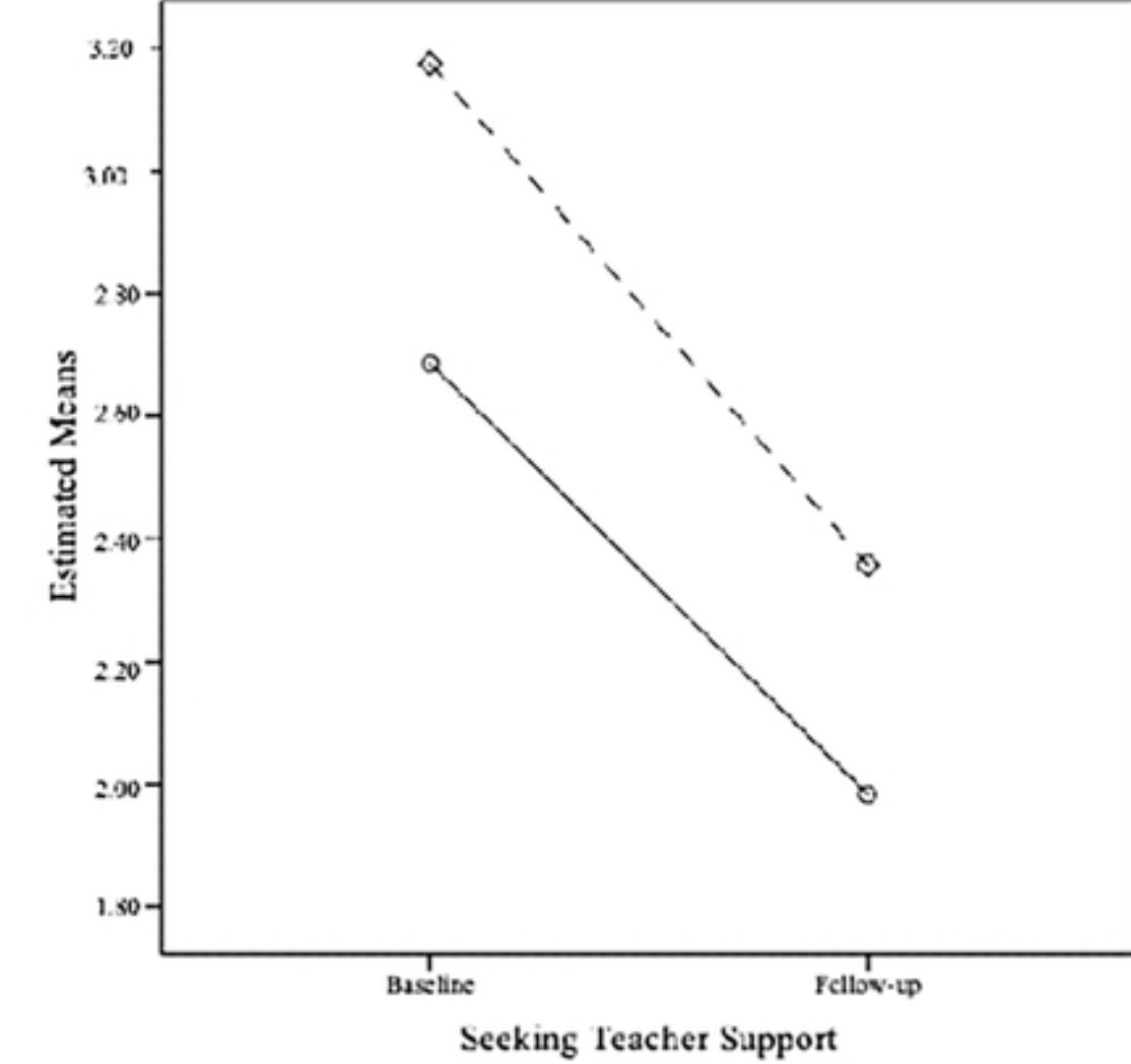
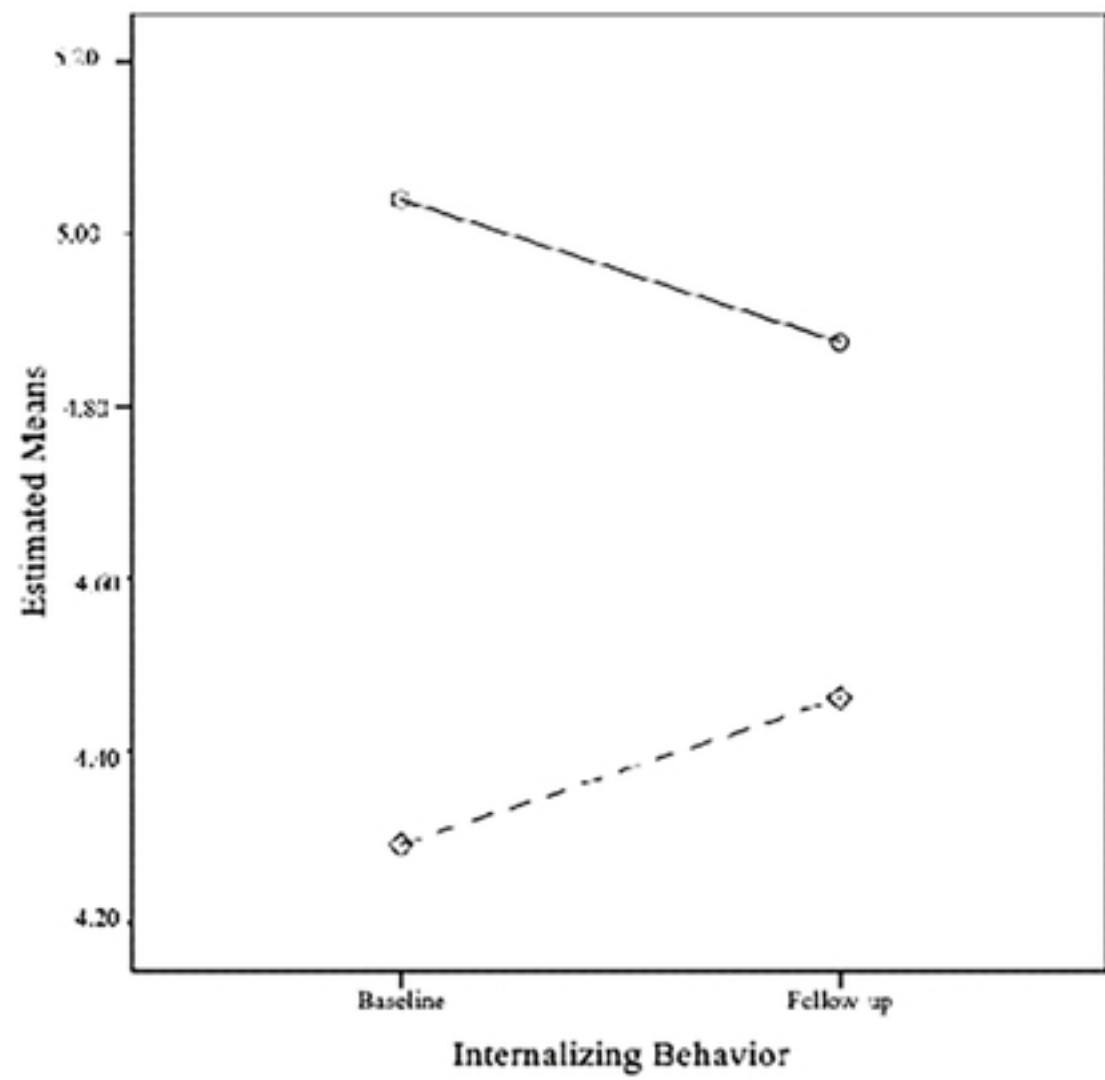
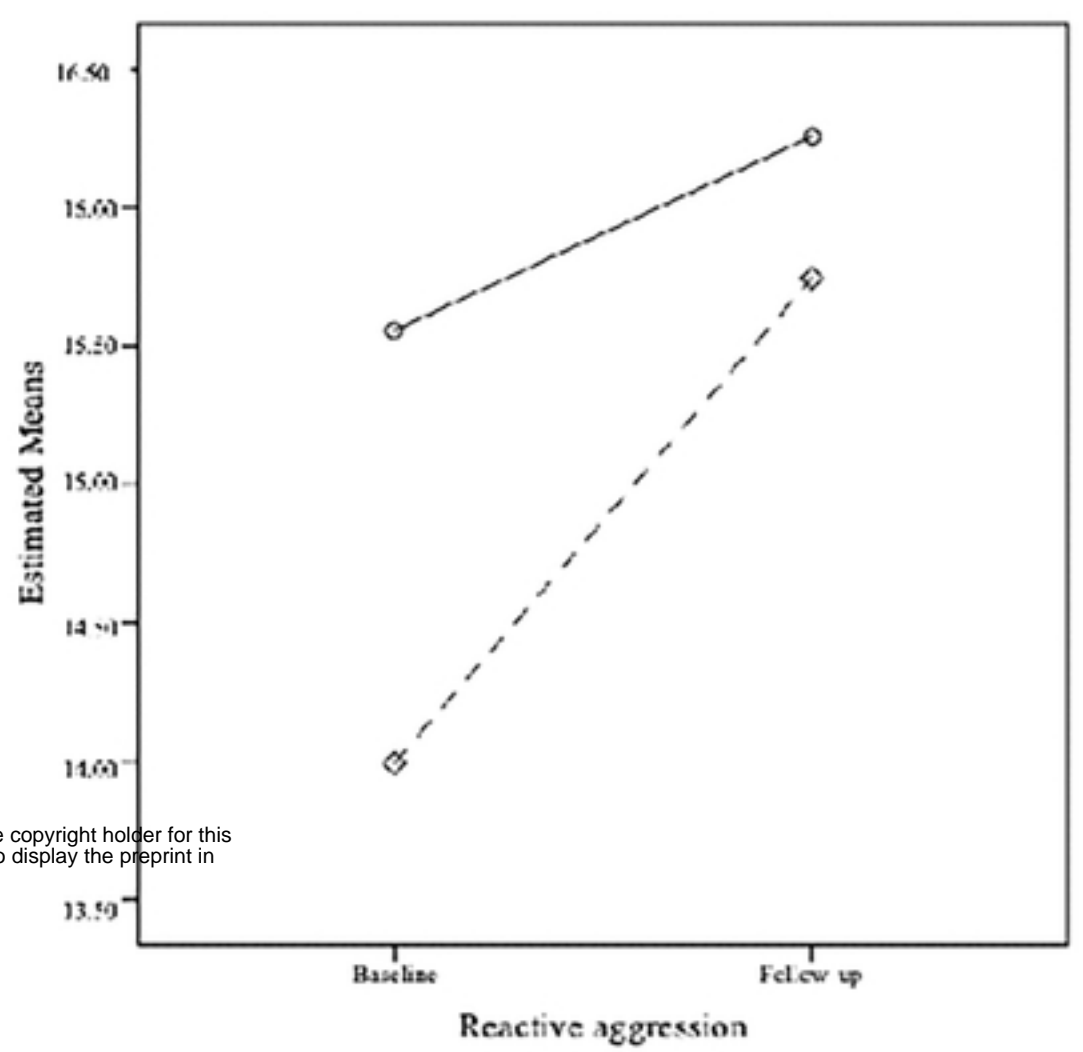
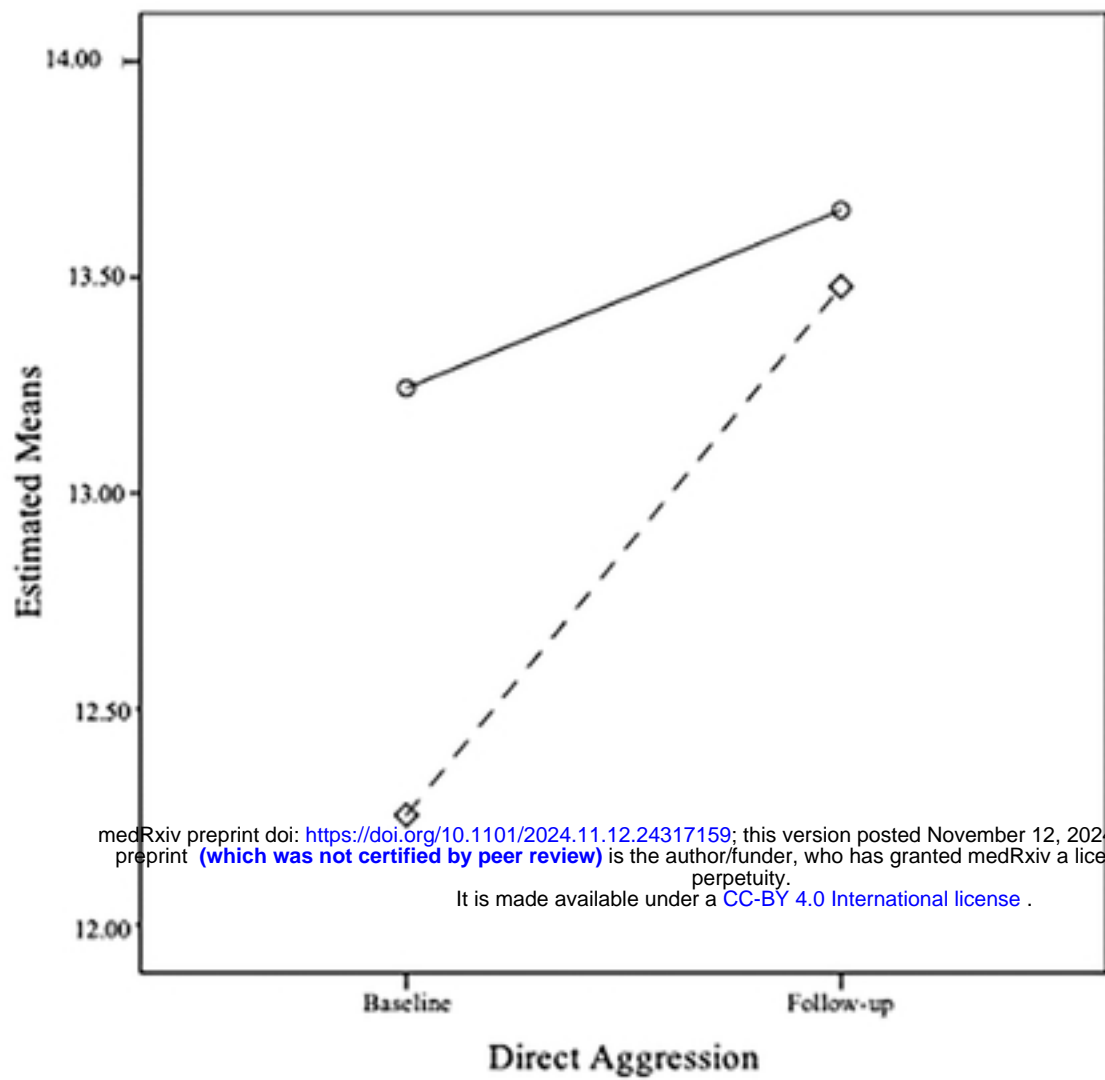
e



f

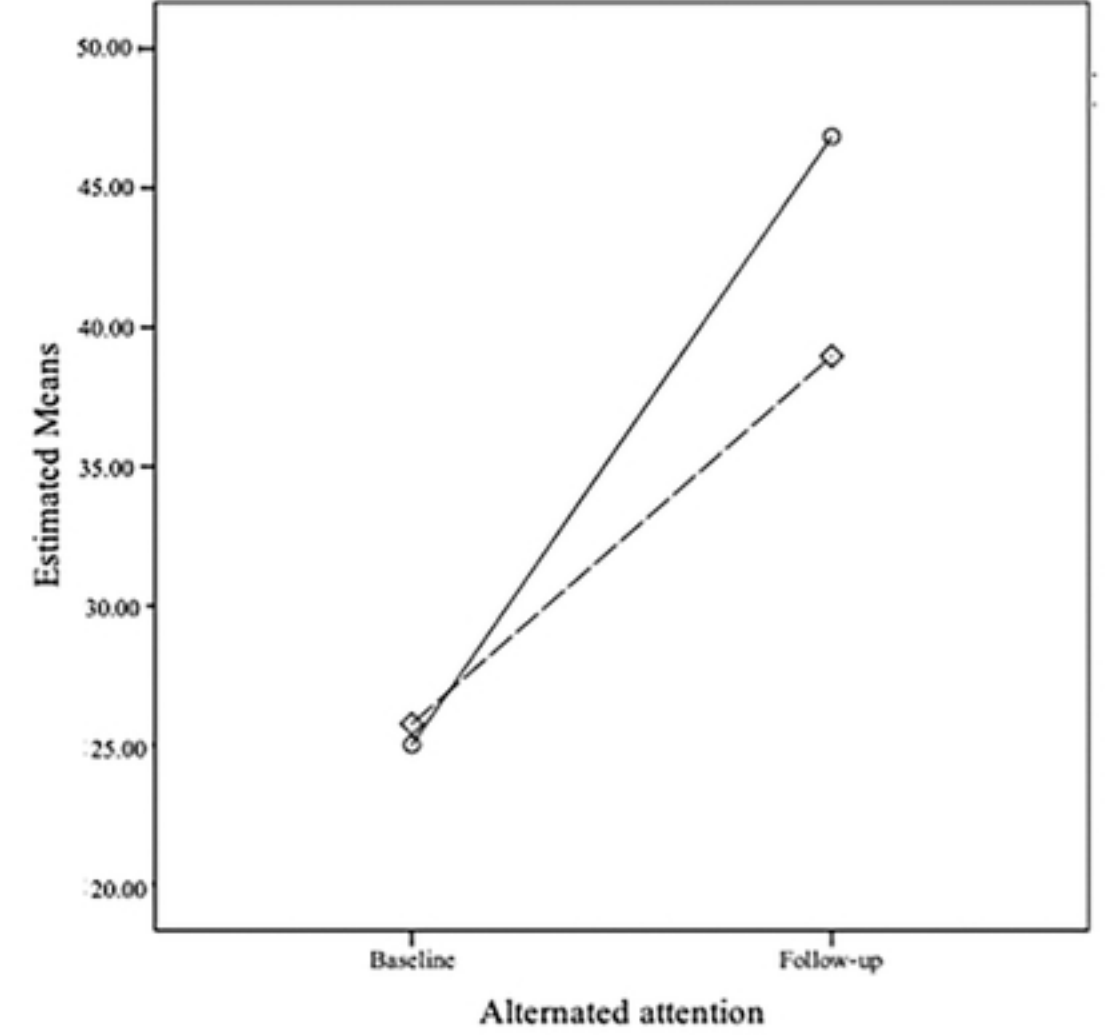
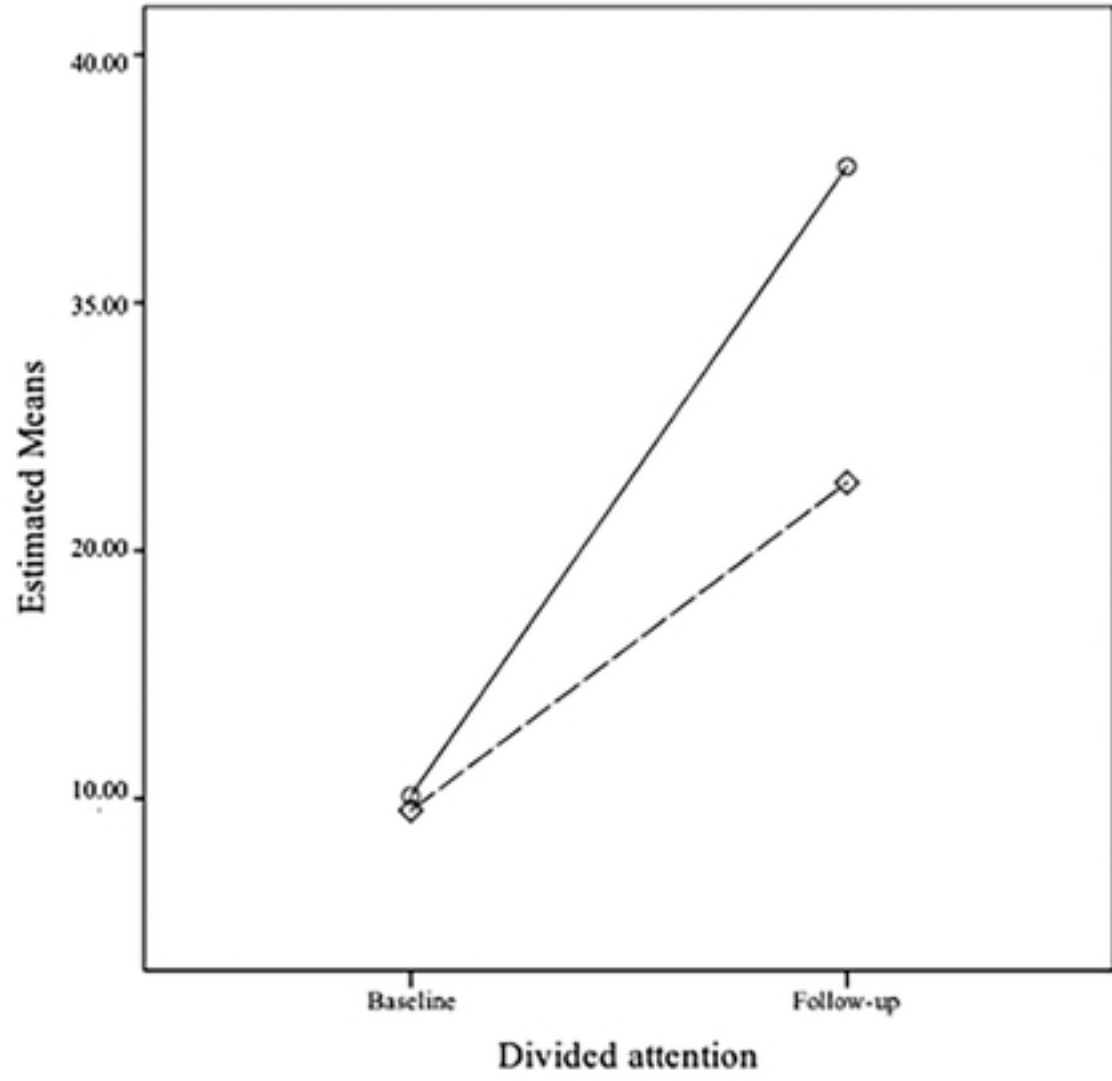
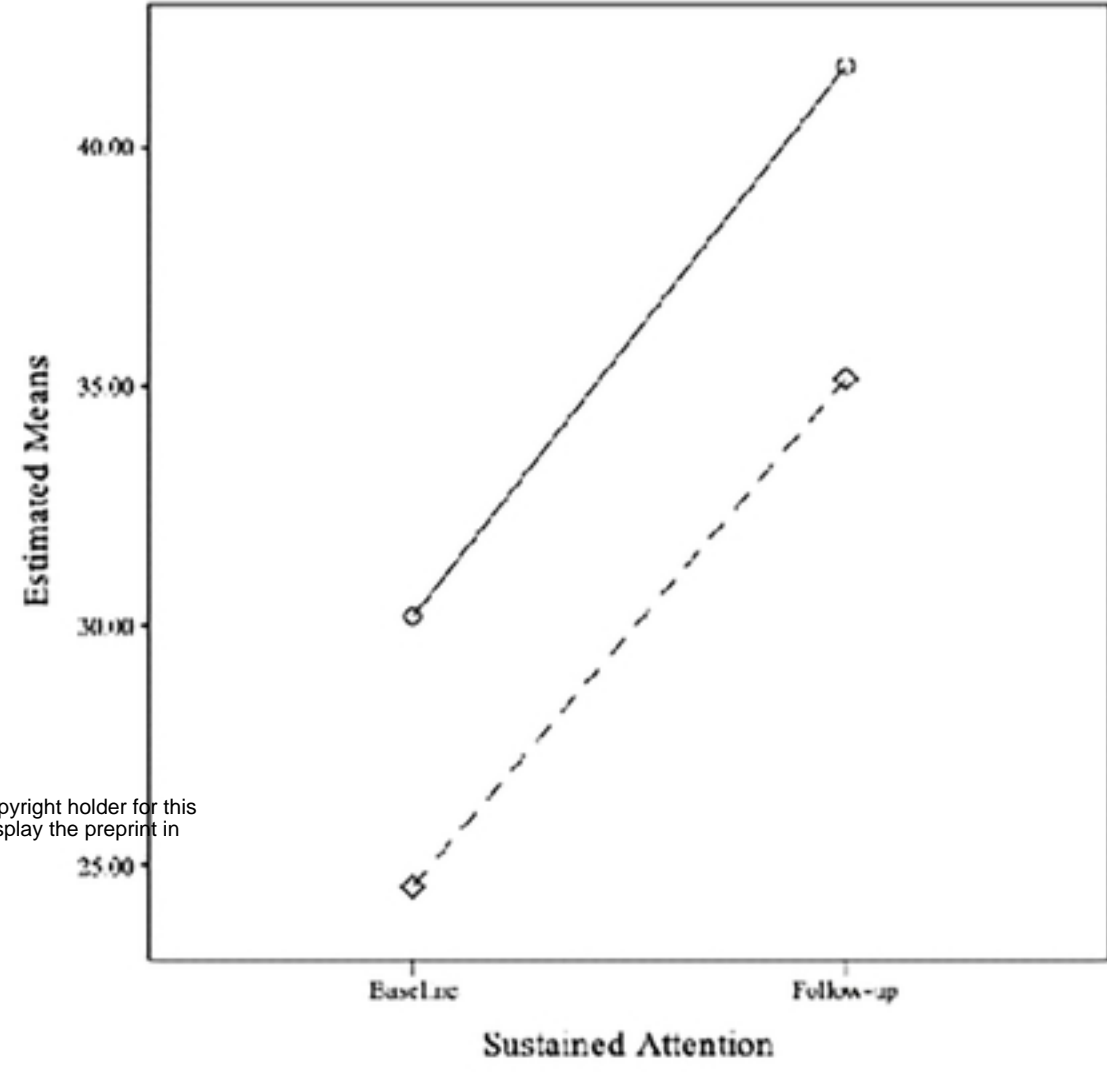
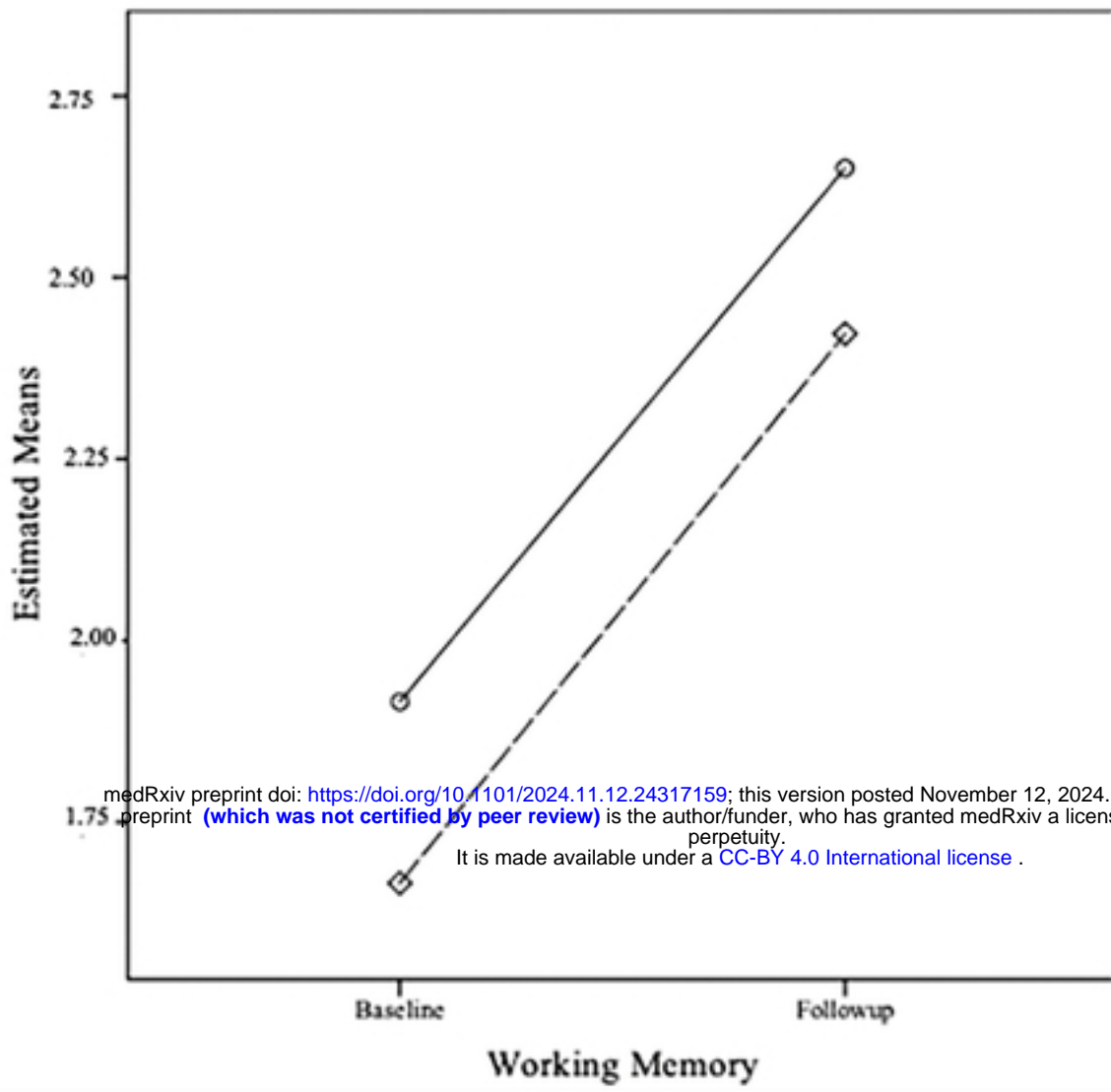
medRxiv preprint doi: <https://doi.org/10.1101/2024.11.12.24317159>; this version posted November 12, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Figure



Figure





Figure