Psychiatry Research 293 (2020) 113377

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

A randomized controlled trial of 25 sessions comparing music therapy and music listening for children with autism spectrum disorder

Thomas Rabeyron[a,*], Juan-Pablo Robledo del Canto[a,d], Emmanuelle Carasco[b], Vanessa Bisson[c], Nicolas Bodeau[c], François-Xavier Vrait[b], Fabrice Berna[e,f], Olivier Bonnot[a,c]

[a] Laboratoire Interpsy, Psyclip, Université de Lorraine, France
[b] Institut de Musicothérapie, Université de Nantes, France
[c] Centre Hospitalo-Universitaire de Nantes, LPPL, France
[d] Centre for Music and Science, University of Cambridge, United Kingdom
[e] Université de Strasbourg, Hôpitaux Universitaires de Strasbourg, Inserm U1114, France
[f] Collègue Universitaire des Médecines Intégratives et Complémentaires (CUMIC), Nantes, France

ARTICLE INFO

Keywords:
Autism spectrum disorders
Art therapy
Music
Non-verbal communication

ABSTRACT

Background: Music therapy is based on the use of musical elements by a trained and qualified therapist. Clinical researches have suggested that children with Autism Spectrum Disorders (ASD) may benefit from MT. In this regard, this study examines if MT is more effective than simply listening to music for children with ASD. *Method:* A 8-month RCT has been carried out comparing music therapy (MT) to music listening (ML) for children with ASD aged from 4 to 7 years old. Thirty-seven participants were randomly assigned to one of the two groups (MT vs. ML). The outcome measures were the Clinical Global Impression (CGI), the Childhood Autism Rating Scale (CARS) and the Aberrant Behavior Checklist (ABC) in each condition (MT and ML). *Results:* CGI scores decreased more for participants in the MT than in the ML condition. This clinical improvement was associated with an improvement of autistic symptoms on lethargy and stereotypy ABC subscales. *Conclusion:* Our findings suggest that music therapy is more efficient than music listening for children with ASD. The present study thus supports the consideration of MT as a rightful add-on to ASD healthcare programs.

1. Introduction

Music therapy (MT) is based on the use of music or musical elements by a trained and qualified therapist in order to promote health and psychological change. The theory behind MT is heterogeneous and based on several theoretical approaches, including behavioral, cognitive, phenomenological, humanistic and psychoanalytic schools (Simpson and Keen, 2011; Rabeyron et al. 2019). Empirical works and clinical research from 1940 to the early 1990s provide extensive arguments to suggest that individuals with Autism Spectrum Disorders (ASD) may benefit from MT, particularly with regard to communication impairments, social skills, imitation, emotion sharing and recognition as well as general behavior (Reschke-Hernandez, 2011). Children with ASD are indeed often very sensitive to music and respond better to music than words. Consequently, the musicality of sounds could represent a nonverbal language which help them to communicate more easily during therapeutic sessions.

In the last 20 years, the development of Evidence-Based Medicine (EBM) has led to more rigorous studies whose aim is to support the efficacy of MT in the improvement of ASD symptoms. Recent works confirm that MT may indeed help children with ASD (Brondino et al., 2015; Kern et al., 2013) and Geretsegger et al. (2014) identified in a recent review 10 studies providing evidence that MT improves the symptoms of children with ASD for primary outcome areas. Nevertheless, the aforementioned studies present at least one out of two sub-optimal features: they either comprise small samples or short durations. Concerning sample sizes, several studies had fewer than 10 participants (Brownell, 2002; Buday, 1995; Farmer, 2002; Kim et al., 2008), other have included 22 to 24 individuals (Lim and Draper, 2011; Thompson, 2012; Gattino et al., 2011) and one a bigger sample of 50 participants (Lim, 2010). Concerning the duration of the studies mentioned above, most of them implicated evaluations after only a few days or a few weeks. A recent randomized controlled trial study was conducted to test the efficacy of music therapy on 105 children and adolescents with behavioral and emotional problems after 13 and 26 weeks, ASD included (Porter et al. 2017), but only 20 participants had

* Corresponding author at: 23 Boulevard Albert premier, 54000, France.
E-mail address: thomas.rabeyron@univ-lorraine.fr (T. Rabeyron).

https://doi.org/10.1016/j.psychres.2020.113377
Received 16 June 2020; Received in revised form 7 August 2020; Accepted 8 August 2020
Available online 08 August 2020

ASD diagnoses. Studies that have included larger samples, longer durations have usually been directed toward populations other than ASD patients. For instance, Montelo et al. (1998) compared the behavioral effects of active vs. passive (listening-based) group music therapy over a for month period on preadolescents with emotional, learning, and behavioral disorders, showing that patients improved significantly in both conditions. A further review examining the overall efficacy of music therapy for children and adolescents with psychopathology revealed only one study focusing on ASD, but this study lacked a control group (Gold et al., 2004).

However, a recent comprehensive randomized clinical trial that has comprised both a substantial sample of patients (n = 384) specifically in the ASD spectrum and a duration that allowed for follow-up assessments has challenged the clinical relevance of music therapy in ASD (Bieleninik et al., 2017). In this study, participants received enhanced standard care (i.e. routine parental care plus three 60-minute sessions of parent counselling offered by professionals experienced with ASD) with or without improvisational music therapy. The authors found no significant difference between groups in terms of symptoms severity over 5 months concerning social affect (based on the Autism Diagnostic Observation Schedule, ADOS) and social responsiveness (based on the Social Responsiveness Scale, SRS). They nevertheless found a few significant measures for secondary outcomes as well as reporting that children with Intellectual Disability responded better. Moreover, the authors acknowledged some concerns about the implementation of the music therapy during the trial and raised issues about the relevance of focusing on specific autistic symptoms to apprehend the efficacy of music therapy. Finally, this study leaves unresolved a more basic question: is MT more effective than simply listening to music for children with ASD?

In order to try answering this question— while taking into account the two other methodological gaps we have just evoked —we have conducted a single-blind controlled study comparing 25 music therapy sessions to 25 music listening sessions. By doing so, we aimed at assessing the efficacy of music therapy on autistic symptoms of children with ASD. Our main hypothesis was that MT will lead to greater clinical improvement. We also hypothesized that this clinical improvement will be related to the improvement of specific autistic symptoms.

2. Methods

Participants: Thirty-seven participants aged 4 to 7 years old were recruited for this study in five psychiatric facilities located in Nantes, France, which accepted to take part in this a research. Sample size was determined taking into account the existing literature and the usual effect sizes in this domain (Geretsegger et al., 2014). All facilities are day-care centers with social workers, nurses, clinical psychologists and psychiatrists. These facilities provide different types of individual psychotherapeutic treatments as well as group therapy. In France, school is mandatory at age 6, but some children with ASD may remain in these facilities for a few years when the gravity of their symptom is incompatible with school attendance. All individuals were in these day-care facilities and returned home every night. We recruited more male than female participants in both groups, with a total of 31 boys and 5 girls. Demographics (age, gender) were collected for each participant from the psychiatric facilities. The inclusion criteria consisted of a previous validated diagnosis of ASD based on the Childhood Autism Rating Scale (CARS) (Rellini et al., 2004).

Of the 50 children, ages from for 4 to 7, who were in the five facilities, 13 individuals have been excluded: 6 children with auditory impairment (all individuals were tested at some point in their lives) or neurological diseases (with the exception of stabilized epilepsy); 4 children with psychotropic medication that was modified less than 6 months prior to the study; 3 children whose parents refused to take part in the study (see Fig. 1).

Ethical aspects: Informational material and meetings were organized with the staff of all psychiatric facilities to explain the purpose and study design. The families of the children gave informed consent based on a scripted consent form and verbal information describing in detail the nature of the study. Informed consent was obtained from both parents of all children participants included in the study. Ethical approval was granted by the Nantes CHU ethics committee. The study has been registered at the ISCRCTN registry with the identification number ID ISRCTN31179114 (http://www.isrctn.com/ISRCTN31179114)

Randomization: Amid the 37 children who were enrolled in the study, 19 completed all Music Therapy sessions (Group 1, 25 sessions of MT with a qualified music therapist) and 17 completed all Music Listening sessions (Group 2, 25 sessions of ML with no specific therapeutic intervention). Children were randomly assigned to one of these two groups using a generated randomization list for each group at t0 (inclusion and baseline evaluations). Each of the five facilities treated children from both groups (MT vs ML).

Outcome measures: The *Clinical Global Impression* (CGI) is a 7-point scale that allows clinicians to rate severity that is commonly used in ASD trials, with higher scores indicating greater severity. Scores range between: 1 = normal, not at all ill, 2 = borderline mentally ill, 3 = mildly ill, 4 = moderately ill, 5 = markedly ill, 6 = severely ill, and 7 = among the most extremely ill patients; The *Childhood Autism Rating Scale* (CARS) is a 15-item scale that is widely used to assess autistic symptoms. It results in a global score and allows for the assessment of a severity score based on specific autistic symptoms. The *Abberant Behavior Checklist* (ABC) is a 57-item checklist of non-adaptive behavior categories, such as sensory relation, body and object use, and language (Aman and Singh, 1994)

Procedure of evaluation: At t0 (start of the study) and t1 (after 25 sessions), all participants were assessed by two trained clinical psychologists working at the Nantes Hospital. They were totally blind, both at t0 and t1, to the groups the children belonged to (ML or MT condition). During the study, they had no interaction with the children, the children's parents, the therapists included in the study or the staff of the five mentioned facilities. Each clinical psychologist assessed half of the children (19 for one psychologist; 18 for the other) and they assessed the same children at t0 and t1. This assessment was conducted in a consulting room at each psychiatric facility. It took an average time of one hour and used the already validated French translation of these three tests in the following order: (1) the Childhood Autism Rating Scale (CARS, secondary outcome) (Schopler et al., 1980), (2) the Aberrant Behavior Checklist (ABC, secondary outcome) (Wadden et al., 1991) and (3) the Clinical Global Impression – Improvement (CGI-I, henceforth CGI, primary outcome) (Williams et al., 2013),

Interventions: All 25 sessions were held between October 2014 and June 2015 (8 months), lasted 30 minutes and took place in the same room in each facility on the same schedule every week, except during holiday periods. MT sessions were performed by a music therapist and a co-therapist (a nurse or educator from the psychiatry facility). All music therapists were educated and had graduated from the University of Nantes with a 3-year course period with a psychodynamic component to ensure homogenous training (Brun et al., 2019; Rabeyron, 2017).

Music Therapy sessions: All sessions were 30 minutes in duration. The room for MT was quiet and large enough (approximately 30m^2) with a sound system and an instrumentarium[1]. The session was attended by a music therapist (internal or external to the institution), a co-therapist (internal to the institution) and a clinical psychologist trainee. Children could choose freely to stay on the ground or sit on chairs. A playlist was available (commercial music including French and foreign songs) for

[1] An instrumentarium consists of a set of large percussion instruments (a djembe, a drum, a derbouka), two tambourines, a sanza, a slide flute, one balafon, two sound eggs, two maracas, 1 guiro, a set of BoomWhakers. Two wind instruments (a kazoo and a harmonica) were added in January at the music therapists' request.

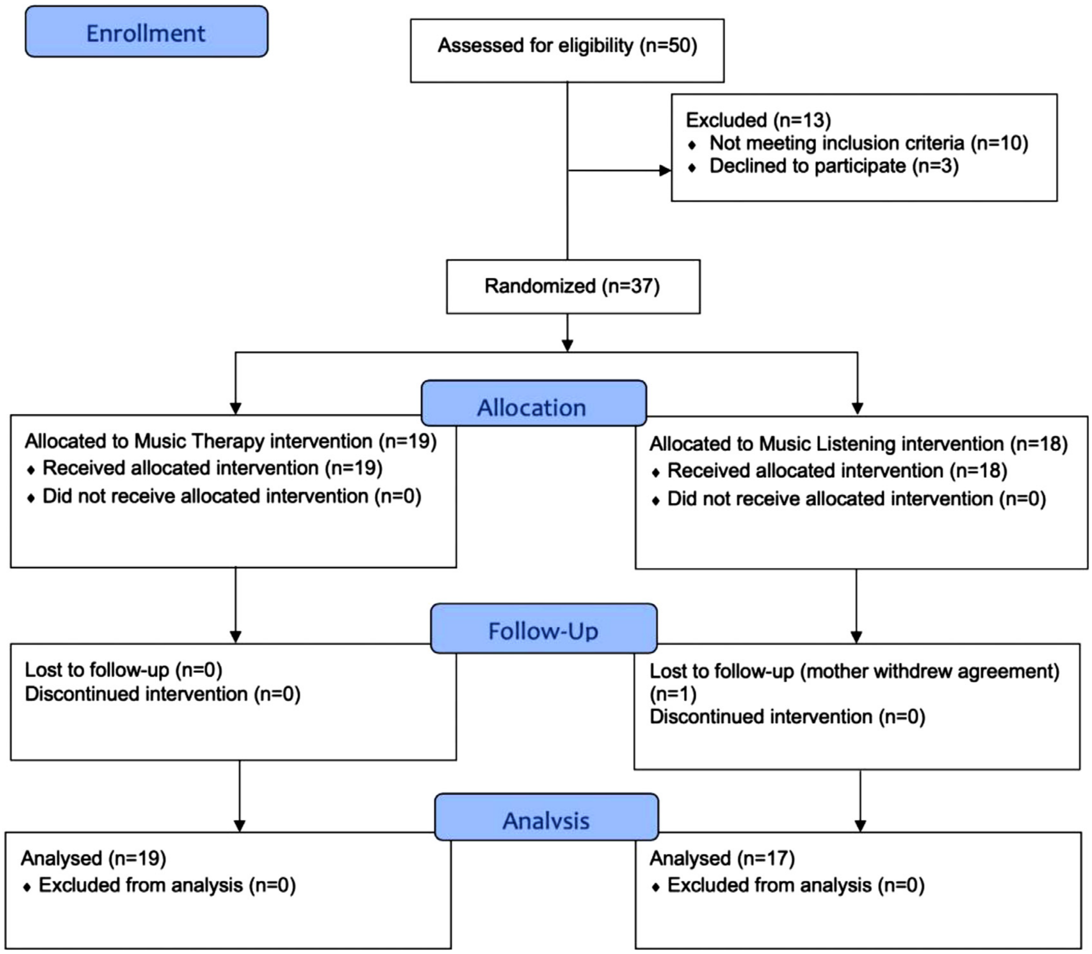

Fig. 1. Consort 2010 Flow Diagram.

opening and closing part of the sessions. These songs were chosen before the study after consensus was reached among music therapists. The sound level of the music and instruments was managed in order to avoid inducing any unpleasant sensation in the participants. The therapists had to follow a rigorous schedule during the sessions: (1) an 'opening ritual' (5-minute instrumental music listening) where children and therapist listened to the first part of the playlist; (2) an instrumental and vocal improvisation (20 minutes) where children had free access to the instruments and interacted with the other children and the therapists. (3) a 'closing ritual' (5-minute vocal music listening) where children and therapists listened to the last part of the playlist before the end of the session.

Music Listening sessions: The ML sessions had the same length (30 minutes) and were conducted in the same room as the MT sessions. ML sessions were performed by two non-qualified music therapists (a nurse or an educator) who were asked not to interact with the children unless necessary for security reasons. If the latter engaged in interaction or dialogue with the nurse/educator, the music therapists were asked to be polite and responsive while not pursuing further interaction or dialogue. Like for the MT group, a playlist was available (commercial music including French and foreign songs) whose songs were chosen before the study after consensus was reached among therapists. The children and the non-qualified music therapists had to listen to this playlist during 30 minutes. Children could choose freely to stay on the ground or sit on chairs while listening to the music.

Statistical analyses: Statistical analyses were performed using Bayesian methods as it is more and more recommended not only in science in general but also in psychiatry and experimental psychology in particular (Vanderkerckhove et al., 2018). In addition, in order to dispose of the highly criticized p-values, Bayesian statistics provide a distribution of the probability that (in our case) the performance of patients is lower than that of controls (which is not to be confounded with the null hypothesis testing of classical statistics using p-values). Moreover, Bayesian statistics offer the advantage of including prior knowledge in the analyses and of testing the robustness of the results by using sensitivity analyses. The latter test whether non-informative priors (a theoretical condition in which the researcher has no prior knowledge on a phenomenon to be tested) or even pessimistic priors (priors that challenge the authors' hypotheses) influence the results of the abovementioned probability.

Sociodemographic variables were compared between groups using univariate linear (for normally distributed data) or logistic (for binary data) regression analyses, and clinical variables with linear or Beta regressions (for non-normally distributed data) analyses according to (Hunger et al., 2012). The CGI score at t0 and t1 was treated as repeated measures. Therefore, a mixed model was used which took the intra-subject variance into account. The same method was used to analyze the CARS and ABC scores and subscores. Predictor variables included group (MT vs. ML) and time (t0 vs. t1). The influence of each predictor was first examined in separate univariate analyses. Interactions between predictors were secondarily examined in multivariate analyses including all predictors.

For all analyses, the probability for the score of each measure being higher in the MT group than in the ML group [indicated as $Pr(MT>ML)$] was calculated. A probability higher than 95% (i.e. $Pr(MT>ML) > 0.95$) for scores of MT being *higher* than those of ML was considered meaningful. Conversely, $Pr(MT>ML) > 0.95$ is equivalent to a probability lower than 5% for scores of MT being *lower* than those of ML (i.e. $Pr(\text{MT}<\text{ML}) < 0.05$). Therefore, $Pr(\text{MT}<\text{ML}) < 0.05$ was also considered meaningful. Interactions between factors were written $Pr(\text{OR}>1)$, OR corresponding here to *exp* of the interaction coefficient. While several studies with small sample sizes provided data in favor of MT, bigger-sample RCTs rather challenged such results (Bieleninik et al., 2017; Porter et al., 2017). Therefore, we used non-informative priors for the group factor as well as for the interaction between group and time factors. Since the decrease in autistic scores after intervention did not reach more than 5% of the initial score in the two mentioned studies, we also used non-informative priors for time. A burn-in of 5000 iterations, followed by 100 000 iterations was used for each of the 3 chains, yielding a final 300 000 iterations-sample for retrieving posterior distribution characteristics. Convergence of the McMC sample chains was checked graphically and was observed in each case. All computations were done using the R software (R Development Core Team, 2018) with all the required additional packages (Rjags, Plummer et al., 2018).

3. Results

Sample Characteristics, sociodemographic and clinical data: Out of the 37 children of the study, all but one completed the trial (one child stopped after the first ML session because her mother withdrew her previously signed agreement). Both groups did not differ significantly in terms of age ($M = 4.9$ years old, $SD = 0.79$ for MT vs. $M = 5.39$ years old, $SD = 0{,}725$ for ML, $Pr(\text{MT}>\text{ML}) = .39$). Despite subjects randomization, the MT group comprised a significantly higher proportion of female participants than the ML group (1/18 vs. 4/13; $Pr(\text{MT}>\text{ML}) = .962$). Therefore, later analyses were adjusted for gender. IQ, as assessed with the WPPSI[2], was comprised between 49 and 62 ($M = 52.2$; $SD = 3.2$) and did not differ between groups.

Hypothesis 1, CGI scores: A decrease in CGI scores was observed between t0 and t1 in both groups as reflected by a meaningful effect of time ($\Pr(\text{t0}>\text{t1}) < .001$) (Fig. 2). This effect was more pronounced in the MT than in the ML group as reflected by a meaningful interaction between the group and time factors (OR = 0.44, CI95% [0.20 – 0.93], $Pr(\text{OR}>1) = .017$). CGI scores did not differ between groups at t0 ($Pr(\text{MT}>\text{ML}) = .72$) but were lower in the MT than in the ML group at t1 ($Pr(\text{MT}>\text{ML}) = .005$) (See Fig. 2). Complementary analyses showed that the proportion of patients showing an improvement of at least 1 point at the CGI was higher in the MT group (17/19, 89.5%) than in the ML group (14/17, 82.4%) (OR = 5.56, CI95% [0.68-33.76], $Pr(\text{MT}>\text{ML}) = .94$). Sensitivity analyses revealed that this difference was even more meaningful when an improvement of at least 2 points was expected: (12/19 (63.2%) for MT vs. 5/17 (29.4%) for ML, OR = 7.36, CI95% [1.76-36.60], $Pr(\text{MT}>\text{ML}) > .99$). Results remained unchanged when gender was introduced in the analyzes.

Hypothesis 2, CARS and ABC scores (Table 1): CARS scores decreased between t0 and t1 (Mean difference (M_diff) = -3.96, CI95% [-6.34 to -1.51], $Pr(\text{t0}>\text{t1}) = .001$). Such a decrease was similar in both groups, as confirmed by the lack of a significant interaction between group and time ($Pr(\text{OR}>1) = .65$). CARS scores at t0 and t1 were higher in the MT than in the ML group. However, such differences were not meaningful ($Pr(\text{MT}>\text{ML}) = .92$). No meaningful interactions between group and time were found for the CARS subscores. Once again, gender had no influence on the results.

The ABC total scores showed marginal changes between t0 and t1 (M_diff = 7.23, CI95% [-3.22 to 17.43], $Pr(\text{t0}>\text{t1}) = .46$). However, ABC scores at t0 were higher in the MT than in the ML group (M_diff = 12.48, CI95% [0.35 to 24.68], $Pr(\text{MT}>\text{ML}) = .98$). This group difference was no longer observed at t1 (M_diff = 6.83, CI95% [-5.81 to 19.86], $Pr(\text{MT}>\text{ML}) = .85$) and the interaction between time and group was meaningful ($Pr(\text{OR}>1) = .03$). Regarding the ABC subscores, two meaningful interactions between time and group were observed for lethargy ($Pr(\text{OR}>1) = .01$) and stereotypy ($Pr(\text{OR}>1) = .03$). Following interventions, lethargy improved in the MT group but remained unchanged in the ML counterpart; MT stereotypy also improved while slightly worsening in the ML group. When including gender in the analyses, the difference between groups at t0 was no longer meaningful ($Pr(\text{MT}>\text{ML}) = .91$), implying that such a difference at t0 was explained both by the fact that female participants had higher ABC scores than their male counterparts, and that the MT group had more female participants than the ML group. All remaining results were unchanged by the introduction of gender.

Correlation analyses: Correlation analyses performed on the whole sample revealed that changes in CGI between t0 and t1 correlated with changes in both the CARS and ABC total scores ($r = 0.36$, CI95% [0.06 to 0.61], $Pr(r>0) > .99$, and $r = 0.50$, CI95% [0.24 to .71], $Pr(r>0) > .99$, respectively), as well as with changes in the lethargy and stereotypy subscores of the ABC ($r = 0.55$, CI95% [0.30 to 0.74], $Pr(r>0) > .99$, and $r = 0.43$, CI95% [0.14 to 0.66], $Pr(r>0) > .99$, respectively). Moreover, changes in CGI scores were more pronounced for participants with higher CARS and ABC scores at t0 ($r = 0.77$, CI95% [0.61 to 0.88] $Pr(r>0) > .99$, and $r = 0.36$, CI95% [0.06 to 0.61], $Pr(r>0) > .99$, respectively).

Sensitivity analyses: Additional analyses using optimistic priors (i.e. increasing the belief in a better efficacy of ML over MT) or pessimistic priors (i.e. challenging our data with belief in better efficacy of MT over ML) did not evidence any influence of the different facilities (where children were treated) on the main and secondary outcomes.

4. Discussion

The aim of this study was to determine whether music therapy during an 8 months period is more efficient than music listening in improving the mental health of children (4-7 years-old) with ASD. Concerning our main hypothesis, CGI scores improved to a better extent in the MT (d = 2.16) than in the ML condition (d = 1.33) with a large effect size at t1 (d = 0.80). This improvement was clinically meaningful as 63.2% of the children exhibited a decrease of at least 2 points at the CGI in the music therapy group compared to 29.4% in the music listening group. Moreover, clinical improvement was correlated with a significant improvement of autistic symptoms as assessed with ABC subscales concerning lethargy and stereotypy symptoms.

Our results thus bring original data in support of the clinical relevance of MT and confirm those of previous RCTs that have included samples of a similar size (Reschke-Hernandez, 2011). Moreover, this study is one of the first to include an active control group comprising music listening, which allowed us to study more specifically the impact of musicotherapy compared to that of music listening alone. The improvement in both groups suggests that joining a music group have a therapeutic effect but also that the special use of music within a structured musicotherapy program catalyzes such an effect. The clinical relevance of our results is further warranted by the fact that the evaluation was carried out after an 8-months period (25 sessions), which contrasts with most of the previous, comparatively shorter studies on music therapy in ASD— except for Bieleninink et al. (2017) and Porter et al. (2017). In this regard, given the chronic course of this clinical condition, having sufficient time to both notice clinical improvements and ensure their stability in time is crucial. Moreover, longer studies are more similar to naturalistic practice, in which MT training sessions usually last for approximately 6-9 months (Brun et al. 2019; Rabeyron et al. 2019).

Meaningful changes between t0 and t1 were not evidenced with the CARS. This scale addresses the relational aspect of the autistic symptoms and this result seems to confirm Bielinink et al.'s (2017) study in which this kind of symptoms was the main outcome measure and did not improved after MT. However, specific improvement of autistic

[2] This evaluation has been carried out before the study and was provided by the psychiatric facilities concerning every child taking part in the study.

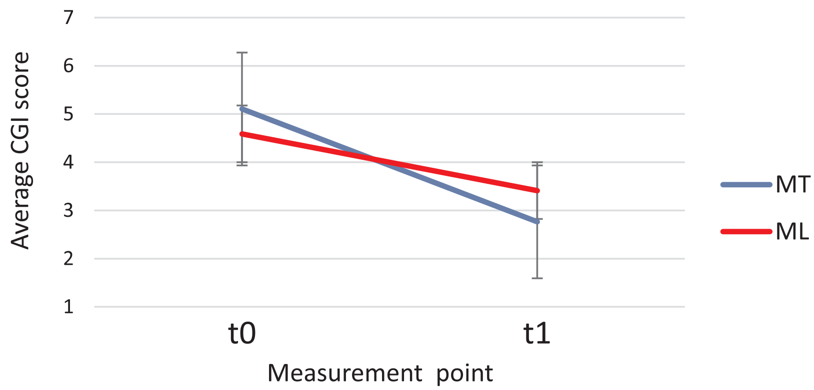

Fig. 2. Average CGI score per condition (MT or ML) and assessment point (t0 or t1). Square brackets indicate standard errors.

Table 1
Baseline and Endpoint scores for CGI, CARS and ABC (MT and ML).

	BASELINE (t0) Music Therapy (n = 19)		Music Listening (n = 17)		ENDPOINT (t1) Music Therapy (n = 19)		Music Listening (n = 17)			
	M	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*d (t0)*	*d (t1)*
CGI	**5.1**	**(1.1)**	**4.6**	**(1.2)**	**2.8**	**(0.9)**	**3.4**	**(0.6)**	**0.43**	**-0.80**
CARS TOTAL SCORE	**39.3**	**(7.7)**	**36.4**	**(8.6)**	**35.9**	**(8.2)**	**33.8**	**(10.8)**	**0.36**	**0.22**
Relationship to people	3.1	(0.8)	2.7	(0.5)	2.7	(0.8)	2.3	(0.8)	0.62	0.50
Imitation	2.7	(0.8)	2.5	(0.9)	2.4	(0.7)	2.1	(1.0)	0.24	0.35
Emotional response	2.7	(0.7)	2.9	(0.7)	2.5	(0.8)	2.9	(0.5)	-0.29	-0.62
Body use	2.4	(0.7)	2.4	(0.9)	2.2	(0.8)	2.1	(0.9)	0.00	0.12
Object use	2.7	(1.0)	2.1	(1.1)	2.5	0.8)	1.9	(1.0)	0.57	0.67
Adaptation to change	2.7	(0.8)	2.4	(1.0)	2.1	(0.6)	1.9	(1.0)	0.33	0.25
Visual response	2.6	(0.8)	2.5	(0.6)	2.3	(0.8)	2.1	(0.8)	0.14	0.25
Listening response	2.3	(0.7)	2.3	(0.7)	2.1	(0.7)	2.1	(0.7)	0.00	0.00
Taste-smell-touch response and use	2.3	(0.7))	2.1	(0.9)	2.1	(0.8)	1.9	(0.8)	0.25	0.25
Fear and nervousness	2.4	(0.6)	2.8	(0.8)	2.4	(0.8)	2.6	(0.8)	-0.57	-0.25
Verbal communication	3.0	(0.9)	2.6	(0.8)	2.9	(0.9)	2.3	(0.9)	0.47	0.67
Non-verbal communication	2.3	(0.9)	2.1	(1.0)	2.1	(0.8)	1.9	(0.9)	0.21	0.24
Activity level	2.4	(0.9)	2.6	(0.7)	2.2	(0.9)	2.2	(0.9)	-0.25	0.00
Level and consistency of intellectual response	2.7	(0.7)	2.4	(0.8)	2.6	(0.7)	2.3	(0.9)	0.40	0.38
General impressions	3.1	(0.6)	2.6	(0.7)	2.9	(0.7)	2.5	(0.7)	0.77	0.57
ABC TOTAL SCORE	**58.7**	**(24.4)**	**53.1**	**(23.3)**	**48.6**	**(27.2)**	**49.8**	**(28.6)**	**0.23**	**-0.04**
ABC irritability	10.4	(6.9)	10.7	(7.0)	8.6	(7.0)	9.7	(8.1)	-0.04	-0.15
ABC lethargy	**21.9**	**(11.4)**	**16.2**	**(11.2)**	**16.0**	**(11.8)**	**15.8**	**(13.4)**	**0.50**	**0.02**
ABC stereotypy	**8.4**	**(6.6)**	**7.0**	**(5.7)**	**6.0**	**(5.1)**	**7.7**	**(5.9)**	**0.23**	**-0.31**
ABC inappropriate speech	15.4	(11.4)	16.8	(11.3)	13.6	(10.6)	14.3	(11.0)	-0.12	-0.06
ABC hyperactivity	2.6	(2.4)	2.5	(2.4)	4.4	(11.9)	2.4	(1.7)	0.04	0.29

Note: CGI = Clinical Global Impression. CARS = Childhood Autism Rating Scale. ABC = Aberrant Behavior Checklist. *d* = effect size (Cohen's *d*)

symptoms has been hereby objectified with two ABC subscales (lethargy and stereotypy symptoms). From this point of view, it is worth taking onto consideration that one of the natural responses to music in general— and to beat, in particular —are vestibular movements such as rocking, swaying, or head movements (Zentner and Eerola, 2010). Such bodily responses to beat induction have been proved to be already functional in newborns and young infants (Winkler et al., 2009). One may therefore hypothesize that the use of music within a therapy program (but not simply listening to music) could induce such a motor facilitation into the participants, thus affecting other behaviors, and explaining the improvement on the lethargy scale. This hypothesis would need to be confirmed by direct observation in future studies of children's bodily responses during the sessions. A significant improvement was also found for stereotypies showing that they slightly improve in the MT group while slightly worsening in the ML group. A possible explanation for this result is that music listening, when not coupled with a specific therapy program, might improve rhythmic movements like stereotypies, the latter decreasing in the context of a music-based therapy. However, since stereotypies are highly sensitive to children's anxiety and environment (Goldman et al., 2009), such a decrease of stereotypies might just be the consequence of the overall improvement of the children as reflected by the CGI.

Finally, the discrepant results obtained with the CARS and the ABC may also be explained by the different properties of the scale and their distinct sensitivity to change. The fact that strong clinical improvements were observed on the CGI but weakly on the CARS and ABC scales also suggests that improvements were manifested in behaviors not contemplated by the ABC or CARS but that nevertheless somehow systematically impressed the clinicians. Recording the sessions should thus be considered as a necessary condition for a similar, future study in order to better understand this aspect.

Limitations of the study and future research directions: We would like to stress in particular four limitations to this study:

- It is worth considering that the present study did not rely on several of the 'gold standard' instruments for the evaluation of ASD - such as the Autism Diagnostic Observation Schedule (ADOS) - which might have shed further light on the matters above discussed.
- Another limitation concerns the number and temporal span of the longitudinal assessments. In this study participants were assessed at baseline (t0, start of the study) and endpoint (t1, after 25 sessions), but no follow-up evaluation was conducted after such an endpoint. Therefore, it is not possible to test the durability of the therapeutic intervention's effects. Future studies could include such a follow-up

assessment after several months and even years.

- During the meeting of all investigators after 3 months and at the end of the research protocol, we noted that in the ML group, professionals had concerns about remaining completely neutral. We had a discussion regarding this issue and concluded that complete neutrality was not fully achieved. This bias - a limitation often inherent to psychotherapeutic research - was, for example, found in comparable MT trials in schizophrenia (Morgan et al., 2011). It might have reduced the impact of MT when comparing it with ML, because ML group might have been more therapeutic than there were supposed to be.
- A similar limitation was the impossibility of a fully replicable psychotherapeutic session (in this case music-therapy) given that free play is part of musicotherapy itself and cannot be manualized step by step when interacting with the children (Rabeyron, 2017, 2020). We addressed these concerns by choosing music therapists who had been in the same training, with the same teachers and by defining a rigorous and similar session duration/schedule as well as the same music playlist. A larger study might try to compare the effect coming from the music therapist in order to understand more precisely what comes from the music therapy or the therapist itself, which is a classical questioning in the field of psychotherapy (Cuijpers et al., 2019).

5. Conclusion

Our findings address a fundamental matter, so far not directly questioned, showing that music therapy and its psychotherapeutic dimension is overall more efficient than simply listening to music for children with ASD. The present study thus favors the consideration of MT as a rightful add-on to the ASD healthcare program. Importantly, the latter was achieved with a sample size and intervention length that raised the standards of most existing literature. There is an important need for further studies in music therapy and all other psychotherapeutic treatments to include larger samples and more control groups. The latter would allow a more precise understanding of the present findings in general, as well as how they detach from music listening and standard care in terms of therapeutic benefits.

Compliance with ethical standards

This research received a grant from Entreprendre pour Aider (www.entreprendrepouraider.fr), a philanthropic foundation, and Academie Francaise (www.academie-francaise.fr). This research could not have been possible without the institutional support of CHU de Nantes, CHS Georges Daumezon, Institut des Hauts Thebaudieres and CHS de Blain.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the CHU de Nantes research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

CRediT authorship contribution statement

Thomas Rabeyron: Conceptualization, Methodology, Formal analysis, Writing - review & editing. **Juan-Pablo Robledo del Canto:** Formal analysis, Writing - review & editing. **Emmanuelle Carasco:** Conceptualization, Project administration, Investigation. **Vanessa Bisson:** Conceptualization, Investigation. **Nicolas Bodeau:** Formal analysis. **François-Xavier Vrait:** Conceptualization, Project administration, Investigation. **Fabrice Berna:** Formal analysis, Writing - review & editing. **Olivier Bonnot:** Conceptualization, Methodology, Formal analysis, Writing - review & editing.

Declaration of Competing Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2020.113377.

References

Aman, M., Singh, N., 1994. The Aberrant Behavior Checklist-Community:Manual. Slosson Educational publications, New-York.
Bieleninik, Ł., Geretsegger, M., Mössler, K., Assmus, J., Thompson, G., Gattino, G., Elefant, C., Gottfried, T., Igliozzi, R., Muratori, F., Suvini, F., Kim, J., Crawford, M.J., Odell-Miller, H., Oldfield, A., Casey, Ó., Finnemann, J., Carpente, J., Park, A.-L., Grossi, E., Gold, C., 2017. Effects of improvisational music therapy vs enhanced standard care on symptom severity among children with autism spectrum disorder: the TIME-A randomized clinical trial improvisational music therapy in autism spectrum disorder improvisational music therapy in autism spectrum disorder. JAMA 318 (6), 525–535.
Brondino, N., Fusar-Poli, L., Rocchetti, M., Provenzani, U., Barale, F., Politi, P., 2015. Complementary and alternative therapies for autism spectrum disorder. Evid.-Based Complement. Alternat. Med. eCAM, 258589.
Brownell, M.D., 2002. Musically adapted social stories to modify behaviors in students with autism: four case studies. J. Music Ther. 39 (2), 117–144.
Brun, A., Chouvier, B., Roussillon, R., 2019. Manuel des médiations thérapeutiques. Dunod, Paris.
Buday, E.M., 1995. The effects of signed and spoken words taught with music on sign and speech imitation by children with autism. J. Music Ther. 32 (3), 189–202.
Cuijpers, P., Reijnders, M., Huibers, M., 2019. The role of common factors in psychotherapy outcomes. Annual Rev. Clin. Psychol. 15 (1), 207–230.
Farmer, K., 2002. The Effect of Music vs. Nonmusic Paired with Gestures on Spontaneous Verbal and Nonverbal Communication Skills of Children with Autism between the Ages. Thesis, (PhD). Florida State University.
Gattino, G.S., dos Santos Riesgo, R., Longo, D., Loguercio Leite, J.C., Schüler Faccini, L., 2011. Effects of relational music therapy on communication of children with autism: a randomized controlled study. Nordic J. Music Therapy 20 (2), 142–154.
Geretsegger, M., Elefant, C., Mössler, K.A., Gold, C., 2014. Music therapy for people with autism spectrum disorder. Cochrane. Database. Syst. Rev. 6 (6), CD004381.
Gold, C., Voracek, M., Wigram, T., 2004. Effects of music therapy for children and adolescents with psychopathology: a meta-analysis. J. Child Psychol. Psychiatry 45 (6), 1054–1063.
Goldman, S., Wang, C., Salgado, M.W., Greene, P.E., Kim, M., Rapin, I., 2009. Motor stereotypies in children with autism and other developmental disorders. Dev. Med. Child Neurol. 51, 30–38.
Hunger, M., Döring, A., Holle, R., 2012. Longitudinal beta regression models for analyzing health-related quality of life scores over time. BMC Med. Res. Methodol. 12, 144.
Kern, P., Rivera, N.R., Chandler, A., Humpal, M., 2013. Music therapy services for individuals with autism spectrum disorder: a survey of clinical practices and training needs. J. Music Ther. 50 (4), 274–303.
Kim, J., Wigram, T., Gold, C., 2008. The effects of improvisational music therapy on joint attention behaviors in autistic children: a randomized controlled study. J. Autism Dev. Disord. 38 (9), 1758–1766.
Lim, H.A., 2010. Effect of "developmental speech and language training through music" on speech production in children with autism spectrum disorders. J. Music Ther. 47 (1), 2–26.
Lim, H.A., Draper, E., 2011. The effects of music therapy incorporated with applied behavior analysis verbal behavior approach for children with autism spectrum disorders. J. Music Ther. 48 (4), 532–550.
Montello, L., Coons, E.E., 1998. Effects of active versus passive group music therapy on preadolescents with emotional, learning, and behavioral disorders. J. Music Ther. 35 (1), 49–67.
Morgan, K., Bartrop, R., Telfer, J., Tennant, C., 2011. A controlled trial investigating the effect of music therapy during an acute psychotic episode. Acta Psychiatr. Scand. 124 (5), 363–371.
Plummer, M., Stukalov, A., Denwood, M. rjags: Bayesian Graphical Models using MCMC. 2018. https://CRAN.R-project.org/package=rjags.
Porter, S., McConnell, T., McLaughlin, K., Lynn, F., Cardwell, C., Braiden, H.j., Rogan, S., 2017. Music therapy for children and adolescents with behavioural and emotional problems: a randomized controlled trial. J. Child Psychol. Psychiatry 58 (5), 586–594.
R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
Rabeyron, T., 2017. Médiations thérapeutiques et processus de symbolisation : de l'expérience sensible à la modélisation. L'Evol. Psychiatrique 82 (2), 351–364.
Rabeyron, T., 2020. Le paradigme clinique du jeu: son évaluation et sa modélisation dans le champ des psychothérapies psychanalytiques. Bulletin de psychologie 566 (2),

83–98.
Rabeyron, T., Saumon, O., Dozsa, N., Carasco, E., Bonnot, O., 2019. De la médiation musicothérapique dans la prise en charge des troubles du spectre autistique chez l'enfant : évaluation, processus et modélisation. La Psychiatrie de l'enfant 62 (1), 147–171.
Rellini, E., Tortolani, D., Trillo, S., Carbone, S., Montecchi, F., 2004. Childhood Autism Rating Scale (CARS) and autism behavior checklist (ABC) correspondence and conflicts with DSM-IV criteria in diagnosis of autism. J. Autism Dev. Disord. 34 (6), 703–708.
Reschke-Hernández, A.E., 2011. History of music therapy treatment interventions for children with autism. J. Music Ther. 48 (2), 169–207.
Schopler, E., Reichler, R.J., DeVellis, R.F., Daly, K., 1980. Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). J. Autism Dev. Disord. 10 (1), 91–103.
Simpson, K., Keen, D., 2011. Music interventions for children with autism: narrative review of the literature. J. Autism Dev. Disord. 41 (11), 1507–1514.
Thompson, G., 2012. Making a Connection: Randomised Controlled Trial of Family Centred Music Therapy for Young Children with Autism Spectrum Disorder. Thesis (PhD). University of Melbourne.
Vandekerckhove, J., Rouder, J.N., Kruschke, JK., 2018. Bayesian methods for advancing psychological science. Psychon. Bull. Rev. 25, 1–4.
Wadden, N.P., Bryson, S.E., Rodger, R.S., 1991. A closer look at the autism behavior checklist: discriminant validity and factor structure. J. Autism Dev. Disord. 21 (4), 529–541.
Williams, K., Brignell, A., Randall, M., Silove, N., Hazell, P., 2013. Selective serotonin reuptake inhibitors (SSRIs) for autism spectrum disorders (ASD). Cochrane. Database. Syst. Rev. 8, CD004677.
Winkler, I., Haden, G., Ladinig, O., Sziller, I., Honing, H., 2009. Newborn infants detect the beat in music. In: Proceedings of the National Academy of Sciences. 106. pp. 2468–2471.
Zentner, M., Eerola, T., 2010. Rhythmic engagement with music in infancy. In: Proceedings of the National Academy of Sciences of the United States of America. 107. pp. 5768–5773.