Title: The Test–Retest Repeatability of a Rhythm Coordination Test Procedure in 4- to 6-Year-Old Children : A Pilot Study

Year: 2020

Version: Published version

Copyright: © The Author(s) 2020

Rights: CC BY 4.0

Rights url: https://creativecommons.org/licenses/by/4.0/

Please cite the original version:
Tuominen, Pipsa P. A.; Raitanen, Jani; Rinne, Marjo (2020). The Test–Retest Repeatability of a Rhythm Coordination Test Procedure in 4- to 6-Year-Old Children : A Pilot Study. Music and Science, 3. DOI: 10.1177/2059204320962281
The Test–Retest Repeatability of a Rhythm Coordination Test Procedure in 4- to 6-Year-Old Children: A Pilot Study

Pipsa P. A. Tuominen1,2, Jani Raitanen3,4 and Marjo Rinne3

Abstract
Moving to music combines the ability of rhythm and coordination. In relation to the musical and motor development of children, sensorimotor synchronization requires the ability to perceive and perform a steady beat. The present pilot study aimed to investigate the test–retest repeatability of a rhythm coordination test procedure in order to pilot the procedure for children. Test–retest repeatability reflects the variation in measurements taken by the rhythm coordination test on the same participant under the same conditions. Ten children (mean age 5.5 years, standard deviation (SD) 0.6) participated in the tests. The test performance was evaluated in points from 0 to 8, separately at a slow and fast pace, resulting in 16 points altogether. Test sessions were performed twice for each child within a four-day interval. Children reached, on average, 7.8 (SD 3.5) points in the first and 8.1 (SD 3.6) points out of 16 in the second test session. The test–retest mean intra class correlation was 0.96, indicating a high repeatability of the rhythm coordination test for children. Further, the children older than 5.5 years achieved a higher number of points than the younger ones, and they seemed to reach almost the same level as adults. The rhythm coordination test procedure for children provides a promising means of assessing children’s sensorimotor synchronization in the context of movement and music.

Keywords
Beat perception, child development, movement to music, performance, sensorimotor integration

Introduction
Synchronizing bodily movements to music combines both perceptual and motor tasks, such as detecting a beat and tempo and coordinating movements to the beat. This sensorimotor synchronization (SMS) requires the ability to perceive and perform a steady beat (Repp & Su, 2013). Rhythm coordination integrates both sensor and motor abilities needed for beat detection, timing, and motor coordination of different rhythms (e.g., in dancing, ball games) (Rinne, 2010).

In children, the development and ability to perceive and produce a steady beat, and the ability to synchronize movement with an external beat, have been primarily studied in tapping and walking studies (Repp & Su, 2013). In natural settings, children were found to perform movements such as hopping, circling, or swaying when listening to familiar music (Eerola et al., 2006). A metric complexity of music affects perceptual sensitivity to beat alignment (Einarson & Trainor, 2016). However, most 2- to 4-year-olds did not adjust the tempo of their movement according to the tempo of the music (Eerola et al., 2006). Synchronization has been found to improve as children develop, especially from ages...
3 to 7 (Repp & Su, 2013). Further, the 4-year-old children were more likely to adjust their movements with the tempo changes compared with the younger ones (Eerola et al., 2006).

In previous studies, steady beat synchronization was examined during tapping, stepping, or walking tasks, using slow, medium, and fast musical examples (Dalla Bella et al., 2017; Gill, 2015; Rose, 2016). Children have also been found to perform hand movements better than feet movements, with the highest accuracy being at a medium tempo (Rose, 2016). Similarly, children have been found to adapt to faster and slower tempos (Rose, 2016). Adults have been found to overcorrect the interval between the taps within the tempo changes (Dalla Bella et al., 2017). In general, children have more trouble matching the metronome beat at a slow pace compared with adults (Gill, 2015).

Multi-limb coordination tasks have been used to assess motor difficulties instead of beat detection and synchronization. Simultaneous clapping while walking has also been used to examine multi-limb coordination and stability, consistency, and frequency changes over childhood (Getchell, 2006). Older children have shown a higher accuracy in matching limb frequency to a metronome beat or other limb movements (Getchell, 2007; Ireland et al., 2018).

In the field of music research, rhythm coordination has been studied as a part of perception and synchronization. In physiotherapy, the focus has been on the development of movement and coordination skills. However, to our knowledge, there is no functional field test available for children that combines movements of upper and lower extremities and synchronization with the external beat. When developing a new test or transferring an old test to a new age group, carefully designed pretests of feasibility, reliability, and repeatability are essential. Test–retest reliability (i.e., repeatability) reflects the variation in measurements taken by a specific instrument (i.e., the rhythm coordination test) on the same participant under the same conditions (Koo & Li, 2016).

This pilot study evaluates the repeatability of the children’s rhythm coordination test procedure, which was used in the Step into music! intervention study. In the Step into music! intervention, the main aim was to assess children’s and their parents’ accelerometer-measured physical activity and sedentary time when using a music mat at home. The Step into music! study also assessed the coordination (including the rhythm coordination) of the children, and the association between children’s musical background and their physical exercise adherence.

Methods

Aims

Our test aimed to assess children’s ability to rhythmically synchronize and coordinate movements of their upper and lower extremities to the given metronome beat. The pilot test was done prior to and separate from the Step into music! study group in autumn 2017. The Step into music! study was approved by the Ethics Committee of University of Jyväskylä, Jyväskylä, Finland.

Participants

Ten children (seven girls, three boys) were recruited via Tursola daycare center in Kangasala, Finland. Children were eligible to participate if they were between 4 and 6 years old, had normal vision and hearing with or without glasses or a hearing aid, and could perform marching and clapping hands as instructed. All the children were Caucasian, and their primary language at home was Finnish. The mean age of the participating children was 5.5 (standard deviation (SD) 0.6) years. Parents gave written consent for participation on their child’s behalf. The study was conducted following prevailing ethics principles.

Rhythm Coordination Test Procedure and Test–Retest Protocol

The rhythm coordination test assessed beat perception ability with synchronization of rhythmic movement, as well as coordination skills of upper and lower extremities as demonstrated by a performance of marching and clapping hands. In the original test for adults, participants took steps in place for every single beat for the first 30 seconds and added hand-claps on every other beat for the next 30 seconds using slow and fast tempos, separately (Rinne et al., 2001). Figure 1 shows the marching and clapping procedure. The tempo was given by the beat of a metronome. The rhythm coordination test procedure was modified for children by adding a rehearsal for a test (i.e., tester showed marching and clapping and did it with a child) to ensure that the child understood what to do. The addition of the rehearsal was based on the earlier finding that young children synchronized their hand movements with higher accuracy in a social context (Kirschner & Tomasello, 2009).

The rhythm coordination test procedure for children consisted of three separate parts: 1) rehearsal of marching and clapping hands together with the tester using a moderate tempo (104 bpm); 2) the first section of the test with a slower, low-impact tempo (92 bpm); and 3) the second section of the test with a fast, high-impact tempo (138 bpm). The slow tempo section started immediately after the rehearsal, and the fast tempo section started immediately after the slow tempo section, once the tester had set the metronome to the correct tempo. The rehearsal was performed as many times as needed, while both the slow and fast rhythm sections were performed only once. The tempo for the rehearsal (104 bpm) was based on an earlier study by Rose (2016) that showed the highest overall accuracy among children at a medium tempo. The tempos for
the slow- (92 bpm) and fast-paced (138 bpm) sections were adapted from the original test by Rinne et al. (2001).

The test performance was evaluated using separate points for the slow- (92 bpm) and fast-paced (138 bpm) sections, resulting in a maximum of 16 points (8 points for each section) (Rinne et al., 2001). Each section of the performance was assessed, as follows:

- 0–10 s, degree of starting accuracy (0 = asynchronous marching, 1 = getting into the marching rhythm gradually, 2 = synchronous marching rhythm at the first go).
- 10–30 s, keeping the beat of exact rhythm (0 = asynchronous marching, 1 = some difficulties in keeping the rhythm, 2 = synchronous and exact marching rhythm).
- 30–40 s, degree of starting accuracy within the first 10 seconds when starting hand-claps (0 = asynchronous marching and hand-clapping, 1 = getting into the marching and hand-clapping rhythm gradually, 2 = synchronous marching and a hand-clapping rhythm at the first go).
- 40–60 s, keeping the beat of the exact rhythm (0 = asynchronous marching and clapping, 1 = some difficulties in keeping the rhythm, 2 = synchronous and exact marching and clapping rhythm).

Each child was tested during two separate measurement sessions by a physiotherapist who was trained to conduct the test. Instructions for children included a description and rehearsal of marching and hand-clapping. Instructions also outlined the possibility of the volume control of the metronome. Each child performed the test procedure, as previously described. During the test procedure, the children were encouraged to continue (e.g., “Keep on going”), and they were told when to start clapping (e.g., “Go on marching and start clapping hands”). The test and retest sessions were conducted at the same time of day within a four-day interval.

### Statistical methods

The means, SDs, and ranges of the measurements are presented as descriptive statistics. The intraclass correlation coefficient (ICC) with a two-way mixed-effects model, type of a single rater, and an absolute agreement as a definition was used to estimate the repeatability (i.e., an experiment capable of giving renewable results) of the rhythm coordination test procedure (Koo & Li, 2016). The ICC reports within-group change (i.e., consistency of children’s performance from test to retest) and systematic change in mean (i.e., in the average performance of children as a group over time) (Atkinson & Nevill, 1998; Lexell & Downham, 2005; Vaz et al., 2013). ICC values of lower than 0.70 represent poor reliability, values from 0.70 to 0.79 show fair reliability, 0.80 to 0.89 demonstrate good reliability, and values equal to or higher than 0.90 represent high reliability (Currier, 1984). A conservative interpretation by Currier (1984) was used due to the study having only a single rater that may cause a smaller variation between test and retest scores compared with the situation with two raters. In other words, when there is only little variation between scores, the ICC tends to be high.

Bland-Altman plot analysis was used to describe an agreement between test and retest by constructing limits of agreement. This method evaluates a bias between the mean differences and estimates 95% limits for the agreement interval (i.e., the total error between test and retest) (Atkinson & Nevill, 1998; Giavarina, 2015).

Differences between the age groups at test and retest were tested by Mann-Whitney U test. The mean age of the children was used as a cut-off point to separate the groups into younger or older.

All analyses were performed using IBM SPSS Statistics 24.0.

### Results

The descriptive data of the pilot study are presented in Table 1. Because of the small data set, the normality was not tested. The mean difference between the test and retest was -0.3 points (95% CI 1.3 to 0.7), which is statistically not significant ($p=0.50$). The test–retest mean ICC was 0.96, indicating high reliability, with a 95% confidence interval (CI) being 0.86 to 0.99, indicating at least good reliability.

The Bland and Altman plot (Figure 2) shows only small differences between the test–retest results. The mean difference of the tests ($\gamma=-0.3$) indicates the bias between the test and the retest, showing that the children achieved higher results in the retest than in the test. All of the data points lie within $\pm 1.96$ SD of the mean difference ($-2.9$ to 2.3), which indicates the limits of agreement (95% CI). Further, the magnitude of the systematic difference (95% CI of the mean difference being from 1.3 to 0.7) and the heteroscedasticity of the plots indicate that the retest does not under- or over-estimate the results compared with the first test.
We also found differences that were related to the age of children. A Mann-Whitney U test indicated that the points in the tests were higher among older children (Md=11.5, n=4) than among younger children (Md=5.5, n=6) with a large effect, U=0.0, p=0.010, r=0.82. The corresponding values at retest for older (Md=11.5) and younger (Md=5.0) children were U=0.0, p=0.009, r=0.82.

### Discussion

The present rhythm coordination test was piloted for the Step into music! study, using a small group of volitional 4- to 6-year-old children. In the present pilot study, the test-retest repeatability of the children’s rhythm coordination test procedure was studied and found to be high. In addition, we found that older children received higher scores than younger ones.

Concerning the consistency of test–retest sessions, ICC was used as the general measure of repeatability in the rhythm coordination test procedure. The ICC value in this population was higher than in earlier studies (Rinne et al., 2001; Vartiainen et al., 2006), which were, however, conducted among adults. The high ICC value might be related to the heterogeneity of the small data set: the ICC values tend to be smaller when the data are homogeneous (Atkinson & Nevill, 1998). It was also seen that children tended to reach slightly higher average values during the retest session compared with the test session. We assume that this may be related to the fact that the test situation and the physiotherapist were more familiar the second time, making it easier for a child to concentrate on the actual task. The stability in the rhythm coordination task, as observed in the Bland and Altman plot, showed that the individual differences in test–retest sessions were at most 2 points, which is in line with Rinne et al. (2001). The slight improvement from test to retest may also reflect a learning effect, since the test was done twice within a short period. A similar improvement between test and retest sessions was found by Béigel et al. (2018), who studied perceptual and sensorimotor timing skills among healthy adults. They concluded that when the same musical excerpts were used at both testing times, improvement between the sessions reflected a learning effect (Béigel et al., 2018).

In addition, rehearsal of marching and clapping hands was added for the children, which was not done in the original test procedure for adults. The rehearsal was added to the test procedure in order to ensure that the children understood what they were expected to do. In addition, Kirschner and Tomasello (2009) found that young children synchronized their hand movements with higher accuracy in the social context; therefore, we believed the rehearsal of

| Gender | Test (points, range 0–16) | Retest (points, range 0–16) | Mean of test and retest (Test + Retest)/2 (points) | Test–retest difference (points) | (Test–retest)/Mean (%) |
|--------|--------------------------|-----------------------------|-----------------------------------------------|-------------------------------|-----------------------|
| male   | 5.29                     | 3                           | 4.0                                           | -2                            | -50.0                 |
| female | 4.78                     | 4                           | 5.0                                           | -2                            | -40.0                 |
| male   | 5.21                     | 5                           | 4.5                                           | 1                             | 22.2                  |
| female | 5.47                     | 6                           | 5.5                                           | 1                             | 18.2                  |
| female | 5.02                     | 6                           | 5.5                                           | 1                             | 18.2                  |
| female | 5.07                     | 8                           | 8.0                                           | 0                             | 0.0                   |
| male   | 5.65                     | 11                          | 11.0                                          | 0                             | 0.0                   |
| female | 6.40                     | 11                          | 11.5                                          | -1                            | -8.7                  |
| female | 6.60                     | 12                          | 11.5                                          | 1                             | 8.7                   |
| female | 5.62                     | 12                          | 13.0                                          | -2                            | -15.4                 |
| Mean   | 5.51                     | 7.8                         | 8.1                                           | -0.3                          |                       |
| Standard deviation | 0.59 | 3.5 | 3.6 | 1.3 |       |

**Table 1.** Data showing agreement between test and retest. Values are presented for each child, ranking them from the lowest to the highest.

**Figure 2.** Bland-Altman plot for data from Table 1. The dashed black line shows the mean difference of the tests. The darker area illustrates 95% CI of the mean difference (-1.3 to 0.7), indicating the magnitude of the systematic difference. The dotted lines (±1.96*SD) as the random error indicate 95% CI (-2.9 to 2.3) as the representation of the limits of agreement.

---

### Additional Table

| Gender | Test (points, range 0–16) | Retest (points, range 0–16) | Mean of test and retest (Test + Retest)/2 (points) | Test–retest difference (points) | (Test–retest)/Mean (%) |
|--------|--------------------------|-----------------------------|-----------------------------------------------|-------------------------------|-----------------------|
| male   | 5.29                     | 3                           | 4.0                                           | -2                            | -50.0                 |
| female | 4.78                     | 4                           | 5.0                                           | -2                            | -40.0                 |
| male   | 5.21                     | 5                           | 4.5                                           | 1                             | 22.2                  |
| female | 5.47                     | 6                           | 5.5                                           | 1                             | 18.2                  |
| female | 5.02                     | 6                           | 5.5                                           | 1                             | 18.2                  |
| female | 5.07                     | 8                           | 8.0                                           | 0                             | 0.0                   |
| male   | 5.65                     | 11                          | 11.0                                          | 0                             | 0.0                   |
| female | 6.40                     | 11                          | 11.5                                          | -1                            | -8.7                  |
| female | 6.60                     | 12                          | 11.5                                          | 1                             | 8.7                   |
| female | 5.62                     | 12                          | 13.0                                          | -2                            | -15.4                 |
| Mean   | 5.51                     | 7.8                         | 8.1                                           | -0.3                          |                       |
| Standard deviation | 0.59 | 3.5 | 3.6 | 1.3 |       |
the test would improve the results of the test and retest in some children. In this study, four children achieved better scores and four children achieved lower scores on the retest. Thus, we assume that rehearsal did not affect the results in the meaning of learning.

The children older than 5.5 years achieved almost the same scores in the rhythm coordination tests as the adults did (Rinne et al., 2001). This is in line with Gill (2015), who found that 5- to 7-year-olds showed the ability to modify their walking, even if there were differences with meeting the exact timing of metronome paces. We found that the youngest children demonstrated a higher deviation in results between the test and retest than the older ones. This result also reinforces Ireland et al.’s (2018), Gill’s (2015), and Getchell’s (2006) findings that task difficulty (slow or fast pace, hand-clapping simultaneously while walking) might influence children’s performance. We considered that the differences between younger and older children might be related to the development of their motor and musical abilities: the motor and musical development of older children is likely to be more advanced than that of younger children. Specifically, among children with musical training, years of music lessons predicted performance on rhythm synchronization tasks over and above the effect of age (Ireland et al., 2018). However, children in their study were older than the children in this study. Further, we did not ask whether the children had music-based hobbies.

Task-related differences in action-dependent and age-related tasks have been found to stay constant across age groups (Getchell, 2006). Children with developmental coordination disorder (DCD) demonstrate more within-individual and between-participant variability than children without DCD or adults in timing the claps with the footfalls, in timing each limb separately, and in using the pairs of limbs (Mackenzie et al., 2008; Volman et al., 2006). The availability of visual or auditory cues has not found to have any role in the coordination of clapping/marching (Mackenzie et al., 2008). Since hand–hand coordination patterns have been found easier than contralateral (different body side) hand–foot coordination patterns among healthy children (Rose, 2016) and children with developmental coordination disorder (Volman et al., 2006), we assume that simultaneous clapping while marching would be a good task to examine dual- and multi-limb rhythm coordination among children.

The main limitations of the study are the small sample size and the use of only one tester. Owing to the small sample size, possible deviation from the normality of the distribution cannot be detected and therefore, for instance, the ICC can be biased. Thus, for future research, a larger study population is needed to examine the differences between younger and older children. In addition, the inter-rater reliability of the rhythm coordination test should be assessed. It would also be important to examine testers’ subjective assessment concerning measured accuracy of the child’s marching and clapping, as well as possible differences in performance when the child has an audio cue versus an audiovisual cue. As for the strength of the study, there has been a lack of rhythm coordination tests available for children. To our knowledge, this kind of rhythm coordination test has not yet been performed among children. Thus, this study provides valuable information for further studies and clinical practice.

Conclusion
Concerning the consistency of pilot test sessions, the preliminary results of test–retest mean ICC indicated a high repeatability of the procedure within this population. Thus, simultaneous clapping while marching would be a good task for examining dual- and multi-limb rhythm coordination among children. In addition, the rhythm coordination test has clinical relevance, for example, when testing children in music education or inspecting children’s coordination.

Acknowledgements
We thank Professor Jukka I. Louhivuori and Professor Urho M. Kujala from the University of Jyväskylä, and Pauliina Husu from the UKK Institute for Health Promotion Research for their support.

Action Editor
Graham Welch, University College London, Institute of Education.

Peer Review
Valentin Bégel, McGill University, Department of Psychology. Nina Politimou, University College London, Department of Psychology and Human Development.

Author Contribution
PPAT and MR were responsible for the initial design of the study. PPAT and JR designed and implemented the statistical analysis. PPAT prepared the first version of the manuscript. All authors were involved in revising it for content and have given final approval of the version to be published. All authors read and approved the final manuscript.

Declaration of conflicting interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval
The Step into music! study was approved by the Ethics Committee of Jyväskylä University, Jyväskylä, Finland (Code 6/2017_Louhivuori).

Funding
The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a grant from the Ellen and Artturi Nyysönen Foundation (Kalevi Heinilä’s fund), the Jenny and Antti
Ireland, K., Parker, A., Foster, N., & Penhune, V. (2018). Rhythm and melody tasks for school-aged children with and without musical training: Age-equivalent scores and reliability. *Frontiers in Psychology*, 9, 426.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, 102(3), 299–314. http://doi.org/10.1016/j.jecp.2008.07.005

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. http://doi.org/10.1016/j.jcm.2016.02.012

Lexell, J. E., & Downham, D. Y. (2005). How to assess the reliability of measurements in rehabilitation. *American Journal of Physical Medicine & Rehabilitation*, 84(9), 719–723. http://doi.org/10.1002002060-200509000-00013

Mackenzie, S. J., Getchell, N., Deutsch, K., Wilms-Floet, A., Clark, J. E., & Whitall, J. (2008). Multi-limb coordination and rhythmic variability under varying sensory availability conditions in children with DCD. *Human Movement Science*, 27, 256–269. http://doi.org/10.1016/j.humov.2008.02.010

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, 20(3), 403–452. http://doi.org/10.3758/s13423-012-0371-2

Rinne, M. (2010). Effect of physical activity, specific exercise and traumatic brain injury on motor abilities. Theoretical and pragmatic assessment (Doctoral thesis). Studies in sport, physical education and health. University of Jyväskylä.

Rinne, M. B., Pasanen, M. E., Miilunpalo, S. I., & Oja, P. (2001). Test-retest reproducibility and inter-rater reliability of a motor skill test battery for adults. *International Journal of Sports Medicine*, 22, 192–200.

Rose, P. (2016). Effects of movement, tempo, and gender on steady beat performance of kindergarten children. *International Journal of Music Education*, 34(1), 104–115. http://doi.org/10.1177/0255761415533309

Vartiainen, M. V., Rinne, M. B., Lehto, T. M., Pasanen, M. E., Sarajuuri, J. M., & Alaranta, H. T. (2006). The test–retest reliability of motor performance measures after traumatic brain injury. *Advances in Physiotherapy*, 8(2), 50–59. http://doi.org/10.1080/14038190600700195

Vaz, S., Falkmer, T., Passmore, A. E., Parsons, R., & Andreou, P. (2013). The case for using the repeatability coefficient when calculating test-retest reliability. *PloS One*, 8(9), e73990. http://doi.org/10.1371/journal.pone.0073990

Volman, M. J., Laroy, M. E., & Jongmans, M. J. (2006). Rhythmic coordination of hand and foot in children with developmental coordination disorder. *Child: Care, Health and Development*, 32(6), 693–702. http://doi.org/CCH679