Annals of Physical and Rehabilitation Medicine 61 (2018) 395–400

Available online at ScienceDirect www.sciencedirect.com

Elsevier Masson France EM|consulte www.em-consulte.com

Original article

Test-retest reliability of the Battery for the Assessment of Auditory Sensorimotor and Timing Abilities (BAASTA)

Valentin Bégel [a,b,1], Laura Verga [c,d,1], Charles-Etienne Benoit [a,e,h], Sonja A. Kotz [a,c,d,*], Simone Dalla Bella [a,f,g,h,*]

[a] EuroMov, Montpellier University, 34090 Montpellier, France
[b] NaturalPad, SAS, 34090 Montpellier, France
[c] Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany
[d] Department of Neuropsychology & Psychopharmacology, Faculty of Psychology & Neuroscience, Maastricht University, 6229 ER Maastricht, The Netherlands
[e] Cognition and Actions Laboratory, Institute of Neuroscience, Université Catholique de Louvain, 1200 Brussels, Belgium
[f] International Laboratory for Brain, Music, and Sound Research (BRAMS), H3C 3J7 Montreal, Canada
[g] Department of Psychology, University of Montreal, H3C 3J7 Montreal, Canada
[h] Department of Cognitive Psychology, WSFiZ, 01-030 Warsaw, Poland

ARTICLE INFO

Article history:
Received 1st March 2017
Accepted 7 April 2018

Keywords:
Rhythm perception
Rhythm performance
Sensorimotor synchronization
Reliability
Test-retest
Timing
Rhythm

ABSTRACT

Perceptual and sensorimotor timing skills can be thoroughly assessed with the Battery for the Assessment of Auditory Sensorimotor and Timing Abilities (BAASTA). The battery has been used for testing rhythmic skills in healthy adults and patient populations (e.g., with Parkinson disease), showing sensitivity to timing and rhythm deficits. Here we assessed the test-retest reliability of the BAASTA in 20 healthy adults. Participants were tested twice with the BAASTA, implemented on a tablet interface, with a 2-week interval. They completed 4 perceptual tasks, namely, duration discrimination, anisochrony detection with tones and music, and the Beat Alignment Test (BAT). Moreover, they completed motor tasks via finger tapping, including unpaced and paced tapping with tones and music, synchronization-continuation, and adaptive tapping to a sequence with a tempo change. Despite high variability among individuals, the results showed good test-retest reliability in most tasks. A slight but significant improvement from test to retest was found in tapping with music, which may reflect a learning effect. In general, the BAASTA was found a reliable tool for evaluating timing and rhythm skills.

1. Introduction

Humans are well equipped to process the temporal information of events in the environment. Most of us have a good grasp of the duration of events ranging from a few milliseconds to several years [1]. This highly developed skill is particularly evident in the ability to perceive and move along to a temporally regular stimulus such as the beat of music or the ticking of a clock. These abilities, shared by musicians and non-musicians alike [2,3], are sustained by a complex neuronal network involving both subcortical (e.g., basal ganglia and cerebellum) and cortical brain structures (e.g., premotor cortex and supplementary motor area) [4–7].

In the last 2 decades, timing abilities have been assessed by a variety of tests, from perceptual tasks such as comparison of durations (duration discrimination [8,9]), detection of shifts in regular sequences (anisochrony detection [3,10,11]), to sensorimotor tasks (e.g., finger tapping with regular sequences [12,13]). Data collected with these tasks have fueled the development of a few influential models and theories such as the Scalar Expectancy Theory (e.g., [14]), neural-integration based models [15–17], and the Dynamic Attending Theory [18]. A surge of interest in timing and rhythmic skills in recent years has been motivated by growing evidence that these skills are linked to important cognitive abilities such as working memory and reading skills [19,20] and may play a role in literacy and language acquisition [7,21,22]. In addition,

* Corresponding authors.
E-mail addresses: sonja.kotz@maastrichtuniversity.nl (S.A. Kotz), simone.dalla.bella@umontreal.ca (S. Dalla Bella).
[1] Shared first authorship.

https://doi.org/10.1016/j.rehab.2018.04.001
1877-0657/

malfunctioning of rhythm processing co-exists with motor or cognitive deficits in several disorders (for a review, see [23]), such as Parkinson disease [24–26], attention deficit hyperactivity disorder [27], or autism [28]. In these cases, rhythmic stimuli can be useful tools for rehabilitation. For example, auditory rhythmic stimulation, whereby rhythmic stimuli (metronome or music) are presented to a patient, has beneficial effects on motor behaviour (e.g., gait in Parkinson disease [24,29,30]) and speech perception and production [22,31–36]. In the context of rehabilitation, fine-grained evaluation of rhythm and timing skills is likely useful to develop individualized rhythmic training for patients with movement disorders [47].

In the last few years, tools for the systematic assessment of timing abilities have been devised. For example, Fuji and Schlaug [37] proposed the Harvard Beat Alignment Test (H-BAT), based on the original Beat Alignment Test (BAT [38]). These batteries include perceptual and sensorimotor beat-based tasks, aimed at testing the ability to extract a beat from a complex auditory sequence. These tasks consist in judging whether a sequence of isochronous tones is aligned to the beat of music or not, to compare duple or triple meters, and to perceive changes in sequences of tones (perception tasks). Moreover, the battery includes tapping tasks to the beat of music or metrical sequences (production tasks). Merchant and collaborators [39] developed a battery of duration-based tests (categorization, discrimination, and reproduction of durations) and beat-based tests with non-musical material (synchronization-continuation, i.e., continue tapping at the same rate right after an isochronous sequence is presented, and synchronization by circle drawing). The battery allowed for discriminating between different profiles of timing impairments in Parkinson disease.

Recently, we proposed an alternative tool for testing timing and rhythmic skills, namely, the Battery for the Assessment of Auditory Sensorimotor and Timing Abilities (BAASTA) [24,40]. The battery includes 9 tasks (4 perceptual and 5 motor tests) for testing beat-based and duration-based timing skills. Perceptual tasks include duration discrimination, anisochrony detection (with tones and music), and a version of the BAT. Production tasks are unpaced tapping, paced tapping (with tones and music), synchronization-continuation and adaptive tapping.

In recent studies [24,40–42], we demonstrated that the BAASTA is sensitive to rhythm deficits. The battery can detect different profiles of beat deafness [3,10,41], a condition in which healthy individuals have difficulties with the perception and/or production of rhythm [43–46]. The BAASTA was also used to uncover rhythm impairments in patient populations such as adults and children with developmental stuttering [42] and patients with Parkinson disease [24,47] and to assess the change in perceptual and sensorimotor rhythmic abilities following a rhythm-based training in Parkinson disease [24,47]. These studies were particularly successful in revealing that perceptual and sensorimotor timing skills improved in a longitudinal test-retest design. However, a learning effect may contribute to explaining these findings because the tasks from the BAASTA were repeated. Hence, the test-retest reliability of the BAASTA needs to be assessed so that it can be used more systematically for assessing the effect of training programs on rhythmic abilities in patient populations.

The goal of this study was to conduct a test-retest reliability study [48] of the BAASTA in a sample of 20 older adults. This target group was selected because it was comparable to the age group of participants in the aforementioned training study of patients with Parkinson disease [24,47]. Participants repeated all tests of BAASTA twice at a 2-week interval. Test-retest reliability was expected to provide a good index of the consistency of measures of individuals' scores [49]. For comparison purposes, the performance obtained with the BAASTA for a group of patients with Parkinson disease from a previous study is also reported [24].

2. Material and methods

2.1. Participants

Twenty older adults recruited from the database of the Max Planck Institute for Human Cognitive and Brain Sciences (Leipzig, Germany) took part in the study. The mean (SD) age was 62.95 years (5.92) (range 50 to 76 years) and they were mostly non-musicians (mean [SD] number of years of formal musical training = 2.25 [3.21]). Two participants received 7 and 12 years of formal musical training.

2.2. Tests and procedure

The BAASTA [40] consists of 4 perceptual tasks (duration discrimination, anisochrony detection with tones, anisochrony detection with music, and the BAT) and 5 production tasks (unpaced tapping, paced tapping to an isochronous sequence, paced tapping to music, synchronization-continuation, and adaptive tapping), implemented on a tablet mobile device (LG G Pad 8.0 model). The stimuli were delivered over headphones (Sennheiser HD201). The mean (SD) interval between the test and retest was approximately 2 weeks (16.25 days [3.92]). Ten participants received the perceptual tasks before the production tasks; the other participants performed the tasks in the reverse order. The order within each group of tasks was fixed as indicated above.

2.3. Perceptual tasks

For the first 3 tasks (duration discrimination, anisochrony detection with tones and with music), a 2 down/1 up staircase procedure [40,50] was used to obtain the perceptual threshold. Each task included 2 trials preceded by 4 examples and 4 practice trials with feedback.

2.3.1. Duration discrimination

This task tested the participants' ability to compare the duration of 2 tones (frequency = 1 kHz) presented one after the other (interval between tones = 600 ms). The first tone (standard duration) lasted 600 ms and the second one (comparison duration) between 600 and 1000 ms. The participant judged whether the second tone lasted longer than the first one.

2.3.2. Anisochrony detection with tones

This task assessed the participants' ability to detect a time shift in a sequence of 5 isochronously presented tones (1047 Hz, tone duration = 150 ms, mean inter-onset interval [IOI] = 600 ms). The 4th tone was presented earlier than expected by up to 30% of the IOI. The task was to judge whether the sequence was regular or not.

2.3.3. Anisochrony detection with music

Participants judged whether a musical sequence (i.e., 2 bars from Bach's "Badinerie" orchestral suite for flute, BWV 1067, played with a piano timbre, inter-beat interval [IBI] = 600 ms) was regular or not. The time shift occurred on the 4t^{h} beat (by up to 30% of the IBI).

2.3.4. Beat Alignment Test

This task assessed the participants' ability to perceive the beat in 72 musical fragments from Bach's "Badinerie" and 2 from Rossini's "William Tell Overture" (duration = 20 beats or quarter notes, IBI = 450, 600, and 750 ms). From the 7th musical beat, a metronome with a percussion timbre was superimposed onto the music and aligned or not with the beat (i.e., out of phase by 33% of the music IBI, or with a different period by 10% of the IBI). The task was to judge whether the metronome was aligned or not with the musical beat.

2.4. Production tasks

In the second and third tasks (paced tapping with an isochronous sequence and paced tapping with music), the task was repeated twice for each excerpt and was preceded by one practice trial.

2.4.1. Unpaced tapping

The task assessed the mean inter-tap interval (ITI) and tapping variability in the absence of a pacing stimulus. The participants tapped with their index finger at their preferred rate for 60 sec.

2.4.2. Paced tapping with an isochronous sequence

This task assessed the participants' ability to synchronize their finger tapping with a metronome (60 piano tones, with a tone frequency of 1319 Hz, presented at 600-, 450-, and 750-ms IOI).

2.4.3. Paced tapping with music

This task was similar to the previous tapping task, but musical excerpts taken from Bach's ''Badinerie'' and Rossini's ''William Tell Overture'' (duration = 20 beats or quarter notes, IBI = 600 ms) were used as the pacing stimuli.

2.4.4. Synchronization-continuation

This task tested the ability to continue tapping at the rate provided by a metronome. A metronome (10 isochronous tones) was first presented at one of 3 tempos (600-, 450-, and 750-ms IOI). The participants synchronized with the metronome and continued tapping at the same rate after the sequence stopped for a duration corresponding to 30 IOIs of the pacing stimulus.

2.4.5. Adaptive tapping

This task assessed the participants' ability to adapt their tapping rate to a tempo change in a synchronization-continuation task. As in the previous task, an isochronous sequence (10 tones) was presented, but in 40% of the trials the IOIs between the last 4 tones increased or decreased by 30 or 75 ms, or remained constant. Participants were instructed to tap to the tones, to adapt to the tempo change, and to keep tapping at the new tempo after the stimulus stopped (continuation phase, corresponding to 10 IOIs). Then, they judged whether they perceived a change in stimulus tempo (acceleration, deceleration, or no change) after each trial. Trials were presented in random order. A training block preceded the first experimental trial.

3. Analyses

3.1. Perception tasks

In the duration discrimination and anisochrony detection tasks, the task stopped after 8 turnaround points (i.e., the values corresponding to the direction changes from up to down or from down to up) in the staircase paradigm. The threshold was calculated by averaging the last 4 turnaround points. This threshold is expressed as the percentage of the standard duration in the three tasks (Weber fraction). Trials including more than 30% false alarms (FAs, scored when the participant reported a difference when there was none) were rejected. The mean thresholds across the 2 trials in each task were presented. In the BAT, the d' was calculated by dividing the number of Hits (when a misaligned metronome was correctly detected) by the number of FAs (when a misalignment was erroneously reported).

3.2. Production tasks

Pre-processing of the data included the following steps: the first 10 taps were discarded in paced and unpaced tapping. Trials containing fewer than 10 continuous taps and more than 30 taps in the continuation phase of the synchronization-continuation task were discarded. To be valid, trials in the adaptive tapping task needed to contain at least 8 taps in the continuation phase, and in the synchronization phase, at least 4 taps had to be in the vicinity of the metronome sounds. Taps leading to ITIs smaller than 100 ms were considered artifacts. Outliers, defined as taps for which the ITI between the actual tap and the preceding tap was smaller than Q1 – 3*Interquartile range (IQR) or greater than Q3 + 3*IQR (Q1 = first quartile, Q3 = third quartile) were rejected.

Mean ITI and motor variability (coefficient of variation, CV, of the ITI) for unpaced tapping and motor variability for the continuation phase of synchronization-continuation and adaptive tapping were calculated. Circular statistics [51] were used for the analysis of synchronization performance as in previous studies [3,10,20,42,52,53]. In circular statistics, a circle of 360° degrees represents the inter-stimulus interval (ISI) where zero degrees on the circle indicate the sound or beat. Each tap is represented by an angle on the circleand refers to a unit vector. The resultant vector R is calculated on the basis of the unit vectors for all taps in a trial. The length of R, between 0 and 1, is an index of consistency of the performance and reflects the variability of the relative phase between taps and pacing stimuli. A value of 1 means that all the taps and pacing stimuli occur exactly at the same time, and a value of 0 means that the distribution is totally random (i.e., lack of synchronization). The direction of R is expressed in degrees and represents the accuracy of synchronization, namely, whether taps occurred before or after the pacing stimulus. A positive direction means that the taps occurred after the tones or beats, on average, and a negative one means that the taps anticipated the tones or beats (180° is the antiphase). If the participants' synchronization was below chance level, as assessed by the Rayleigh test for circular uniformity [51,54], accuracy was not calculated. A logit transformation was applied to the vector length before further statistical analyses [40–42]. The mean performance of the 2 trials is presented for paced tapping and synchronization-continuation.

Finally, an adaptation index, defined by the value of the slope of a regression line fitted to the ITIs of the final sequence tempo and a sensitivity index (d') for detecting tempo changes were calculated for adaptive tapping.

3.3. Test-retest analysis

We compared the first and the second testing sessions by standard t test when the distributions were normal in both sessions and by Wilcoxon-Mann-Whitney U test when at least one of the distributions was not normal (Shapiro-Wilks test). Rejection of the null hypothesis implies a significant difference between the 2 sessions and reveals a systematic error that may result from an experimental bias such as the effect of weariness or learning. To control for sources of error due to chance (random error; e.g., biological variability), we calculated the intraclass correlation coefficient (ICC3.1). The ICC3.1 is considered a good index of test-retest reliability with the same experimenter because it does not include the variance associated with systematic error (bias) [48,49,55]. The ICC is commonly used as a measure of relative reliability, namely, the consistency of the relative position of one individual in a group [56]. Following Shrout and Fleiss [57], ICC values were interpreted as > 0.75, excellent; 0.40–0.75, good; and < 0.40, poor. The other form of reliability is absolute reliability

Table 1
Measures of reliability for the Battery for the Assessment of Auditory Sensorimotor and Timing Abilities (BAASTA) in healthy participants compared with patients with Parkinson disease (PD) from Benoit et al. (2014).

		Healthy participants (n = 20)								PD patients (n = 15)
Task	Variable	Session 1 Mean (SD)	Session 2 Mean (SD)	Relative reliability indices	Absolute reliability indices					Mean (SD)
				P	ICC	SEM	SEM%	CR	CR%	
Duration discrimination	Threshold (Weber fraction)	16.49 (5.45)	18.08 (3.47)	1	0.39	3.29	19.06	9.13	52.82	34.1 (11.00)**
Anisochrony detection with tones	Threshold (Weber fraction)	13.67 (3.27)	12.92 (2.86)	1	0.68	1.74	13.08	4.82	36.25	10.73 (3.26)
Anisochrony detection with music	Threshold (Weber fraction)	11.06 (3.54)	13.20 (3.86)	1	0.45	2.75	22.65	7.61	62.75	14.96 (7.12)
Beat Alignment Test	d'	1.67 (1.36)	1.82 (1.49)	1	0.94	0.36	20.69	1.00	57.22	2.49 (1.02)
Paced tapping	Metronome (consistency)	0.91 (0.12)	0.91 (0.11)	1	0.75	0.46	16.45	1.29	45.71	0.94 (0.06)
	Metronome (accuracy)	−3.84 (4.21)	−3.63 (4.01)	1	0.80	0.46	−12.43	1.29	−34.54	11.55 (5.51)**
	Music (consistency)	0.73 (0.20)	0.77 (0.19)	< 0.05	0.96	0.24	16.62	0.67	45.80	0.83 (0.18)
	Music (accuracy)	0.21 (6.68)	1.88 (6.03)	0.62	0.77	0.24	23.28	0.67	64.16	19.63 (14.49)**
Adaptive tapping	Adaptation index (acceleration)	1.76 (1.00)	1.47 (0.55)	1	0.83	0.34	20.95	0.94	58.16	1.78 (0.77)
	Adaptation index (deceleration)	1.05 (0.76)	1.06 (0.93)	1	0.01	0.84	79.65	2.32	220.18	1.16 (0.59)
	d' (acceleration)	2.46 (0.54)	2.7 (0.65)	1	0.20	0.53	20.95	1.46	55.97	2.62 (0.57)
	d' (deceleration)	2.46 (0.61)	2.70 (0.56)	1	0.16	0.54	21.11	1.51	58.54	2.51 (0.57)
Synchronization-continuation	Motor variability (CV ITIs)	0.08 (0.13)	0.11 (0.12)	1	0.31	0.10	107.66	0.28	293.26	0.04 (0.01)*
Unpaced tapping	Inter-tap interval (ms)	493.20 (167.66)	537.73 (264.89)	1	0.32	182.74	35.35	506.18	97.91	597.4 (334.99)
	Motor variability (CV ITIs)	0.12 (0.15)	0.16 (0.19)	1	0.17	0.15	109.94	0.43	305.18	0.05 (0.03)

ICC: intraclass correlation coefficient; CR: coefficient of reliability. Note. P-values are obtained with standard t tests or Wilcoxon-Mann-Whitney U tests depending on the normality of the distributions, and were submitted to Bonferroni correction based on the number of comparisons. PD patients' performance was compared to Session 1 from healthy controls: $^{*}P < 0.05$, $^{**}P < 0.01$.

[48,49], which deals with the degree to which measurements vary and provides an index of the expected trial-to-trial noise in the data. SEM, SEM%, and coefficient of reliability (CR, also referred to as the smallest real difference, or minimum difference) were used to assess absolute reliability. The SEM is expressed in the same unit as the measurement of interest and quantifies the variability between the 2 sessions. It is calculated as the square root of the within-subjects error variance. CR represents the 95% confidence interval, a value for which any retest score outside that interval would be interpreted as having 95% chances to reflect a real difference. It is calculated by multiplying the SEM by 2.77 ($\sqrt{2}$ times 1.96). For comparison between tasks in different units, the SEM% and CR% were calculated by dividing the SEM or the CR by the mean of all measurements from both sessions and multiplied by 100.

4. Results

Results are presented in Table 1. Systematic error tests showed a significant change from test to retest for synchronization consistency ($t(19) = -3.91$, $P < 0.05$ with Bonferroni correction) only in paced tapping with music. All the other tasks showed no significant differences between the 2 times of testing. In the perceptual tasks, the ICC just failed to reach the threshold for good reliability (= 0.40) in duration discrimination (0.39) and was good to excellent (0.45 to 0.94) in the other tasks. In the production tasks, the ICC was excellent for accuracy and consistency in all the paced tapping tasks (0.75 to 0.96). However, it was poor in unpaced tapping, synchronization-continuation, and adaptive tapping (0.01 to 0.32) except for the adaptation index (acceleration) in adaptive tapping (0.83).

The SEM% was between 13.08 and 22.65 in the perception tasks and between 12.43 and 109.94 in the production tasks. The CR% ranged from 34.54 to 305.18. Note that the changes observed between the first and the second sessions were not related to the age of participants. Indeed, correlations between age and the difference between the second and the first sessions did not reach significance in any of the tasks (average $r < 0.14$).

Finally, we compared the current results obtained with the BAASTA in a sample of healthy participants to those of a patient population that can benefit from the assessment of rhythmic abilities. To this aim, we report in Table 1 the results obtained in a previous study with a group of patients with Parkinson disease (n = 15, Benoit et al., 2014 [24])[2]. Patients' performance was compared to the results for healthy participants at pre-test. Patients were impaired in a few measures, such as the threshold in duration discrimination and accuracy in the paced tapping tasks. Because these measures showed good reliability, they are appropriate targets for testing the effect of rhythmic training on the performance of patients with Parkinson disease.

5. Discussion

We conducted a test-retest study of the BAASTA, a tool for the systematic assessment of perceptual and sensorimotor timing skills [40]. Twenty adults completed the battery on 2 occasions, 2 weeks apart. The results of the 2 measurements were compared by using reliability indices. First, a standard comparison of the mean differences with repeated measures tests was used to assess systematic error [49]. The performance remained stable in most of the tasks at both times of testing, which indicates that the performance in these tasks is robust and not affected by a learning

[2] The version of BAASTA from Benoit et al. (2014) [24] was slightly different from the one used in this study (the adaptive algorithm was different in perceptual tasks, and a triangle timbre was used instead of a woodblock for the superimposed metronome in the BAT). Moreover, in perceptual tasks and in paced tapping, the best performance across trials was considered, whereas the mean performance was computed in this study.

effect or annoyance resulting from task repetition. However, we found a difference between the test and retest in paced tapping to music. Participants slightly but significantly improved their performance, showing lower variability at retest than at test (i.e., an improvement of 4.05% in synchronization consistency) when tapping to the beat of the music. This improvement likely reflects a learning effect because the same music excerpts were used to assess synchronization with music at both testing times.

The ICC was used to assess the consistency of the performance at an individual level. It is encouraging to observe that ICC values ranged from satisfactory to excellent, especially in the perceptual tasks and in paced tapping with a metronome and music, regardless of the age of participants. However, test-retest reliability was poor for the mean ITI in unpaced tapping and for variability in unpaced tapping and synchronization-continuation. In addition, 3 of the 4 variables of adaptive tapping yielded poor ICCs, in perception (d' for acceleration), and in production (adaptation index for deceleration). Despite poor test-retest reliability for some measures in the adaptive tapping task, some of them (e.g., the adaptation index, d') have been found useful to discriminate individuals with rhythmic deficits [24,41]. Moreover, they are likely to be successful predictors of the response of patients with Parkinson disease to a rhythm-based protocol [47]. Thus, it is still valuable to include these measures in a systematic evaluation of rhythmic skills using the BAASTA, with the caveat that the variability of these measures at different times may be high for an individual.

One critical aspect of the BAASTA is that the whole battery takes 2 hours, on average, to complete [40]. This is an important limitation especially for assessing timing and rhythmic abilities in a patient population in a clinical setting. Thus, the most relevant tasks sufficient to provide a screening of patients' rhythmic abilities must be selected. The current study suggests that motor variability in unpaced tapping or in a synchronization-continuation test, as well as the interval between the taps in unpaced tapping, yielded poor test-retest reliability. Exclusion of these tasks may be advisable to reduce the testing time while increasing the reliability of the BAASTA. In adaptive tapping, poor test-retest reliability was limited to a certain variable, namely the adaptation index for deceleration and the sensitivity indices (d'). Therefore, this task may not be excluded as a whole.

Finally, indices of absolute reliability were provided. The SEM was used to define the boundaries around which a participant's true score lies, and the CR is useful to estimate whether a performance in a test-retest experiment is likely to reflect a real difference (i.e., an improvement or a deterioration of performance). These indices will be valuable in future studies to consider the performance of individuals, in particular older adult populations, in light of test-retest reliability. The comparison of the results obtained in healthy controls with previous results from a group of patients with Parkinson disease shows that some measures from the BAASTA with good reliability may be valuable indicators of the beneficial effects of a rhythmic training in this population. More generally, further experiments are needed to test the reliability of the battery in a test-retest protocol in patient populations, such as patients with Parkinson disease.

Acknowledgements

The study was supported by a grant from the European Community (EBRAMUS, 7th Framework Programme, grant agreement N^{o} 238157) to SDB and SK, and by a junior grant from the Institut Universitaire de France to SDB.

Disclosure of interest

The authors declare that they have no competing interest.

References

[1] Grondin S. Psychology of time. Bingley: Emerald Group Publishing Ltd; 2008.
[2] Repp BH. Sensorimotor synchronization and perception of timing: effects of music training and task experience. Hum Movement Sci 2010;29:200–13.
[3] Sowiński J, Dalla Bella S. Poor synchronization to the beat may result from deficient auditory-motor mapping. Neuropsychologia 2013;51:1952–63.
[4] Coull JT, Cheng RK, Meck WH. Neuroanatomical and neurochemical substrates of timing. Neuropsychopharmacol 2011;36:3–25.
[5] Merchant H, Harrington DL, Meck WH. Neural basis of the perception and estimation of time. Annu Rev Neurosci 2013;36:313–36.
[6] Schwartze M, Rothermich K, Kotz SA. Functional dissociation of pre-SMA and SMA-proper in temporal processing. Neuroimage 2012;60:290–8.
[7] Schwartze M, Kotz SA. A dual-pathway neural architecture for specific temporal prediction. Neurosci & Biobehav R 2013;37:2587–96.
[8] Grondin S. Duration discrimination of empty and filled intervals marked by auditory and visual signals. Percep Psychophys 1993;54:383–94.
[9] Merchant H, de Lafuente V. Neurobiology of interval timing. New York: Springer; 2014.
[10] Dalla Bella S, Sowiński J. Uncovering beat deafness: detecting rhythm disorders with synchronized finger tapping and perceptual timing tasks. J Vis Exp 2015;97:e51761.
[11] Ehrlé N, Samson S. Auditory discrimination of anisochrony: influence of the tempo and musical backgrounds of listeners. Brain Cognition 2005;58: 133–47.
[12] Repp BH. Sensorimotor synchronization: a review of the tapping literature. Psychon B Rev 2005;12:969–92.
[13] Repp BH, Su YH. Sensorimotor synchronization: a review of recent research (2006–2012). Psychon B Rev 2013;20:403–52.
[14] Gibbon J, Church RM, Meck WH. Scalar timing in memory. Ann NY Acad Sci 1984;423:52–77.
[15] Taatgen NA, van Rijn H, Anderson J. An integrated theory of prospective time interval estimation: the role of cognition, attention, and learning. Psychol Rev 2007;114:577–98.
[16] Wing AM, Kristofferson AB. Response delays and the timing of discrete motor responses. Percept & Psychophys 1973;14:5–12.
[17] Wing AM, Kristofferson AB. The timing of interresponse intervals. Percept & Psychophys 1973;13:455–60.
[18] Large EW, Jones MR. The dynamics of attending: how people track time-varying events. Psychol Rev 1999;106:119–59.
[19] Tierney AT, Kraus N. The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. Brain Lang 2013;124:225–31.
[20] Woodruff Carr K, White-Schwoch T, Tierney AT, Strait DL, Kraus N. Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. P Natl Acad Sci 2014;111:14559–64.
[21] Gordon RL, Shivers CM, Wieland EA, Kotz SA, Yoder PJ, McAuley JD. Musical rhythm discrimination explains individual differences in grammar skills in children. Developmental Sci 2014;18:635–44.
[22] Kotz SA, Gunter TC. Can rhythmic auditory cuing remediate language-related deficits in Parkinson's disease? Ann NY Acad Sci 2015;1337:62–8.
[23] Allman MJ, Meck WH. Pathophysiological distortions in time perception and timed performance. Brain 2011;135:656–77.
[24] Benoit CE, Dalla Bella S, Farrugia N, Obrig H, Mainka S, Kotz SA. Musically cued gait-training improves both perceptual and motor timing in Parkinson's disease. Front Hum Neurosci 2014;8:494.
[25] Jones CR, Jahanshahi M. Contributions of the basal ganglia to temporal processing: Evidence from Parkinson's disease. Timing & Time Perception 2014;2: 87–127.
[26] Kotz SA, Schwartze M. Differential input of the supplementary motor area to a dedicated temporal processing network: Functional and clinical implications. Front Integr Neurosci 2011;5:86.
[27] Noreika V, Falter CM, Rubia K. Timing deficits in attention-deficit/hyperactivity disorder (ADHD): evidence from neurocognitive and neuroimaging studies. Neuropsychologia 2013;51:235–66.
[28] Allman MJ. Deficits in temporal processing associated with autistic disorder. Front Integr Neurosci 2011;5:2.
[29] Thaut MH, McIntosh GC, Rice RR, Miller RA, Rathbun J, Brault JM. Rhythmic auditory stimulation in gait training for Parkinson's disease patients. Mov Disord 1996;11:193–200.
[30] Thaut MH, Abiru M. Rhythmic auditory stimulation in rehabilitation of movement disorders: a review of current research. Mus Percept 2010;27:263–9.
[31] Kotz SA, Schmidt-Kassow M. Basal ganglia contribution to rule expectancy and temporal predictability in speech. Cortex 2015;68:48–60.
[32] Thaut MH, McIntosh KW, McIntosh GC, Hoemberg V. Auditory rhythmicity enhances movement and speech motor control in patients with Parkinson's disease. Funct Neurol 2001;16:163–72.
[33] Mainka S, Mallien G. Rhythmic speech cueing (RSC). In: Thaut MH, Hoemberg V, editors. Handbook of neurologic music therapy. Oxford: Oxford University Press; 2014. p. 150–60.
[34] Przybylski L, Bedoin N, Krifi-Papoz S, Herbillon V, Roch D, Leculier L, et al. Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. Neuropsychol 2013;27:121–31.
[35] Schön D, Tillmann B. Short- and long-term rhythmic interventions: perspectives for language rehabilitation. Ann NY Acad Sci 2015;1337:32–9.
[36] Stahl B, Kotz SA, Henseler I, Turner R, Geyer S. Rhythm in disguise: why singing may not hold the key to recovery from aphasia. Brain 2011;134:3083–93.

[37] Fujii S, Schlaug G. The Harvard Beat Assessment Test (H-BAT): a battery for assessing beat perception and production and their dissociation. Front Hum Neurosci 2013;7:771.

[38] Iversen JR, Patel AD. The Beat Alignment Test (BAT): surveying beat processing abilities in the general population. In: Miyazaki KI, Hiraga Y, Adachi M, Nakajima Y, Tsuzaki M, editors. Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC 10). Adelaide: Causal Productions; 2008. p. 465–8.

[39] Merchant H, Zarco W, Bartolo R, Prado L. The context of temporal processing is represented in the multidimensional relationships between timing tasks. PLoS One 2008;3:e3169.

[40] Dalla Bella S, Farrugia N, Benoit CE, et al. BAASTA: Battery for the assessment of auditory sensorimotor and timing abilities. Behav Res Meth 2017;49:1128–45.

[41] Bégel V, Benoit CE, Correa A, Cutanda D, Kotz SA, Dalla Bella S. "Lost in time" but still moving to the beat. Neuropsychologia 2017;94:129–38.

[42] Falk S, Müller T, Dalla Bella S. Non-verbal sensorimotor timing deficits in children and adolescents who stutter. Front Psychol 2015;6:847.

[43] Launay J, Grube M, Stewart L. Dysrhythmia: a specific congenital rhythm perception deficit. Front Psychol 2014;5:18.

[44] Mathias B, Lidji P, Honing H, Palmer C, Peretz I. Electrical brain responses to beat irregularities in two cases of beat deafness. Front Neurosci 2016;10:40.

[45] Palmer C, Lidji P, Peretz I. Losing the beat: deficits in temporal coordination. Philos T Roy Soc B 2014;369:20130405.

[46] Phillips-Silver J, Toiviainen P, Gosselin N, Piche O, Nozaradan S, Palmer C. Born to dance but beat deaf: a new form of congenital amusia. Neuropsychologia 2011;49:961–9.

[47] Dalla Bella S, Benoit CE, Farrugia N, Keller P, Obrig H, Mainka S, et al. Gait improvement via rhythmic stimulation in Parkinson's disease is linked to rhythmic skills. Sci Rep 2017;7:42005.

[48] Fleiss JL. The design and analysis of clinical experiments. New York: John Wiley and Sons; 1986.

[49] Weir JP. Quantifying test-retest reliability using the intraclass correlation, coefficient and the SEM. J Strength Cond Res 2005;19:231–40.

[50] Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am 1971;49:467.

[51] Fisher NI. Statistical analysis of circular data. Cambridge: Cambridge University Press; 1995.

[52] Kirschner S, Tomasello M. Joint drumming: social context facilitates synchronization in preschool children. J Exp Child Psychol 2009;102:299–314.

[53] Pecenka N, Keller PE. The role of temporal prediction abilities in interpersonal sensorimotor synchronization. Exp Brain Res 2011;211:505–15.

[54] Wilkie D. Rayleigh test for randomness of circular data. Appl Stat 1993;32:311–2.

[55] Brozek J, Alexander H. Components of variation and the consistency of repeated measurements. Research Quarterly. American Association for Health. Phys Educ Recreation 1947;18:152–66.

[56] Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The case for using the repeatability coefficient when calculating test-retest reliability. PLoS One 2013;8:e73990.

[57] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420–8.