Poor synchronization yet adequate tempo-keeping in adults with autism

Keren Kasten | Nori Jacoby | Merav Ahissar

1Department of Cognitive Science, The Hebrew University of Jerusalem, Jerusalem, Israel
2Computational Auditory Perception Group, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany
3The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel
4Department of Psychology, The Hebrew University of Jerusalem, Jerusalem, Israel

Correspondence
Merav Ahissar, The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Edmond J Safra Campus-Givat Ram, Jerusalem 9190401, Israel and Department of Psychology, The Hebrew University of Jerusalem, Mt Scopus, Jerusalem 9190501, Israel.
Email: msmerava@gmail.com

Funding information
European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program, Grant/Award Number: 833694; Israel Science Foundation, Grant/Award Number: 1650/17

Abstract
Sensorimotor synchronization to external events is fundamental to social interactions. Adults with autism spectrum condition (ASC) have difficulty with synchronization, manifested in both social and non-social situations, such as paced finger-tapping tasks, where participants synchronize their taps to metronome beats. What limits ASC’s synchronization is a matter of debate, especially whether it stems from reduced online correction of synchronization error (the “slow update” account) or from noisy internal representations (the “elevated internal noise” account). To test these opposing theories, we administered a synchronization-continuation tapping task, with and without tempo changes. Participants were asked to synchronize with the metronome and continue the tempo when it stopped. Since continuation is based only on internal representations, the slow update hypothesis predicts no difficulty, whereas the elevated noise hypothesis predicts similar or enhanced difficulties. Additionally, tempo changes were introduced, to assess whether adequate updating of internal representations to external changes is possible when given a longer temporal window for updating. We found that the ability to keep the metronome’s tempo after it stopped did not differ between ASC and typically developing (TD) individuals. Importantly, when given a longer period to adapt to external changes, keeping a modified tempo was also similar in ASC. These results suggest that synchronization difficulties in ASC stem from slow update rather than elevated internal noise.

KEYWORDS
adults; auditory; learning; motor (control, system); sensory integration; sequencing

INTRODUCTION
Sensorimotor synchronization to an external event is a fundamental aspect of our social interactions. Synchronization plays a bonding role in daily situations such as dancing, playing music ensemble, and even catching a ball (Van Der Steen & Keller, 2013). Sensorimotor synchronization also supports language learning (Gordon et al., 2015) and the development of social interaction (Feldman et al., 2011), and frequently takes place in a social context, such as interpersonal synchrony (Van Der Steen & Keller, 2013). While there is a growing awareness of reduced sensorimotor skills in autism spectrum condition (ASC), even in the absence of social interactions (Bhat et al., 2011), most studies of synchronization in autism have focused on social synchronization. These studies have found difficulties (Fitzpatrick et al., 2016; Marsh et al., 2013), for example in clapping with others and synchronization of participants’ speech and gestures to those of their partners (Fitzpatrick et al., 2013; Fitzpatrick et al., 2017; Kaur et al., 2018). Moreover, the degree of difficulty in social motor synchronization is associated with ASC severity (Fitzpatrick et al., 2017). However, social environments might pose additional challenges to individuals with autism. Hence, to understand the general, rather than socially related,
mechanisms underlying ASCs’ difficulties in synchronization, we administered a simple paced finger synchronization task, with and without external changes of tempo (illustrated in Figure 1). In this task, the participant is instructed to synchronize (temporally align) her taps to the beats of a metronome. Synchronization is measured by two parameters. The first is mean asynchrony, namely, the mean temporal interval between the participant’s tap and the corresponding (nearest) metronome beat. This mean reflects participants’ perceptual accuracy, and it is typically negative. On average, participants tap slightly before the metronome, and perceive this interval as aligned with the metronome beat (Aschersleben, 2002; Aschersleben & Prinz, 1995; Repp, 2005; Vishne et al., 2021). The second parameter is the variability around this mean (often denoted by standard deviation (SD) of the asynchrony). Smaller variability indicates greater consistency (precision) around this mean (Jacoby et al., 2015).

Several studies have characterized tapping in autism (Morimoto et al., 2018; Sheridan & McAuley, 1997; Tryfon et al., 2017; Vishne et al., 2021), with mixed results. Two child studies found no synchronization difficulties. Thus, Sheridan and McAuley (1997) asked children with ASC and age-matched typically developing (TD) subjects to synchronize their taps to tones presented with fixed intervals of 600 ms and to continue tapping at the same tempo after the metronome stops. In the synchronization phase, when the metronome was on, no consistent difference in either the mean asynchrony or the SD of the asynchrony was found. In the continuation phase, in which participants kept tapping without the metronome at the same tempo, a difference between the groups was found only in younger (age 7–8 years) children. In a more complex synchronization task, children were instructed to listen to rhythms of varying metrical complexity and then tap in synchrony with them. No difference was found between children with ASC and age-matched TDs (Tryfon et al., 2017).

By contrast, two large studies found synchronization impairment in ASC. In one, which included both children and adolescents, participants were asked to tap their index and thumb fingers in synchrony with auditory stimuli presented at a fixed interval of 500 ms (for 15 s). Mean asynchrony and ITI (inter-tap interval, schematically illustrated in Figure 1) did not differ between the TD and ASC groups. However, the ASC group (children and adolescents) had larger variability in their asynchrony and ITI compared to age-matched TDs (Morimoto et al., 2018). Similar results were found in a recent study of paced finger tapping, which was administered to three adult groups, matched for age and cognitive skills (Vishne et al., 2021): TD (n = 56), ASC (n = 38), and individuals with dyslexia (n = 39). Two tapping conditions were administered: fixed intervals (isochronous) and switching tempo. In the fixed condition, the metronome was set to 500 ms intervals (for 50 s). There was no group difference in mean asynchrony, but the ASC group had greater variability around this mean, compared to the two other groups, indicating that the greater variability is not a general characteristic of developmental disabilities.

To decipher the specific mechanisms underlying ASCs’ difficulties in tapping synchronization, Vishne et al. (2021) used a computational model, which allowed the researchers to separate between mean noise levels and serial tapping effects. The model assumes that tapping is the outcome of three underlying processes, two of which are internal representations. The first is timekeeping—the internal representation of the tempo of the external metronome; the mean of this representation is equal to that of the metronome, but it has an inherent (normally distributed) noise. The second is the timing of the motor response—the interval between the specific timing of the current and previous tapping response (finger tap), which is also somewhat noisy. The third process is error correction—the fraction of the perceived synchronization error (in the previous tap) which is corrected in the current tap. This component is crucial for the synchronization process. Typically, participants correct a fraction of this error in the following tap, while the remaining fraction is “carried” to the next tap. The partial correction yields the serial dependence (positive correlation in asynchronies) between consecutive asynchronies. When a
small fraction is corrected, a large fraction is carried to the next tap, yielding a larger serial dependence. Modeling the tapping process revealed that individuals with autism do not have an elevated internal noise, in either their keeping of internal intervals or in the timing of their motor tapping. However, they have a reduced rate of error correction.

The analysis that indicated reduced online error correction is in line with the “slow update” hypothesis, which proposes that a core difficulty of individuals with ASC is slower updating of internal representations (Lieder et al., 2019). Thus, people with ASC adequately learn internal statistics and use them, and hence adequately acquire repeated routines, but it takes them longer than it takes TDs to do so. In a dynamically changing environment, whether social or sensorimotor, where they have to update their predictions online (within sub-seconds to a few seconds), they will have difficulties. This means that online error correction, needed for improving synchronization, is less efficient in autism.

An opposing account—the “elevated internal noise” hypothesis—suggests that increased internal noise plays a role in the symptoms of ASC. This account is supported by various studies using different methods and measures (Baron-Cohen & Belmonte, 2005; Dakin et al., 2005; Rubenstein et al., 2003; Simmons et al., 2009). For example, fMRI and EEG studies have found increased trial-by-trial variability in the neural responses to sensory stimuli in individuals with ASC (Dinstein et al., 2012; Haigh et al., 2016; Milne, 2011). This increased variability may lead to unreliable and less predictable internal representations (Dinstein et al., 2015), which may be especially problematic in social situations where many different cues must be perceived through multiple senses. In addition, elevated internal noise may also contribute to other symptoms of ASC such as motor clumsiness (Whyyatt & Craig, 2013), differences in visual perception (Dakin & Frith, 2005; Simmons et al., 2009) and abnormalities in behavioral variability (Karalunas et al., 2014).

Interestingly, neither of these two recent tapping studies assessed the reliability of paced tapping without the actual metronome beats, namely, the ability to reliably retain the metronome tempo without external information. The expected results differ between the two hypotheses. The elevated internal noise hypothesis predicts that with no external feedback, continuation based on the reliability of internal representations will be noisier in individuals with ASC than in TDs. By contrast, the slow update hypothesis predicts that individuals with ASC will perform the task similarly to TDs, since no external error needs to be integrated and updated. The current study aimed to test these hypotheses.

Importantly, the slow update hypothesis further predicts that individuals with autism will keep metronome’s pace adequately without external feedback, even when this tempo changes, as long as participants are given sufficiently long (several seconds) update time. By contrast, the elevated internal noise hypothesis does not predict that additional taps will reduce difficulties. To test this prediction, we examined whether an extended updating period enables participants with autism to perform similarly to TDs.

We designed a protocol in which participants were asked to synchronize their taps with the metronome at the start of each tapping sequence, and keep tapping at the tempo of the last metronome beats when the metronome stopped. We administered both isochronous (fixed) sequences (500 ms intervals), which are considered automatic, that is, do not require explicit attention and rely on basic sensorimotor skills (Repp & Keller, 2004), and sequences with tempo changes, which are more cognitively demanding and rely on executive functions (Repp & Keller, 2004). To assess the rate of updating to the new beat, we administered two types of sequences with tempo changes. In one, the metronome stopped 12 beats (~6 s) after the tempo change. We assumed that this interval is sufficient for tempo-updating in ASC. In the other, the metronome stopped shortly (2 beats) after the tempo change. We assumed that if individuals with ASC are slow updaters, this condition will be particularly challenging for them.

METHODS

Participants

Participants with autism (and no language difficulty, as assessed by self-report and current reading scores) were recruited through clinicians, support centers, and designated facilities. All had community-based diagnoses, conducted by multi-disciplinary teams of medical specialists who had used various standard diagnostic tools, including ADI-R (Lord et al., 1994), ADOS-G (Lord et al., 2000), and CARS (Schopler et al., 1980) to ensure that all participants met DSM-IV criteria (American Psychiatric Association, 2000). Since we assumed that tapping skills may be enhanced by previous musical experience (e.g., Repp, 2010), which is difficult to quantify and match (Zentner & Strauss, 2017), we ensured that all recruited participants had no more than minimal (less than 2 years) formal musical education.

Data were collected from 64 participants (31 TD and 33 ASC), but 4 were excluded: 3 participants (1 TD individual and 2 ASC) had too many errors (as described below, 5% or more of their sequences were excluded, while the average fraction of missing sequences was 0.29%). One individual with autism was excluded due to very poor reading, suggesting language difficulties. The final groups consisted of 60 participants: 30 TDs and 30 individuals with ASC. All experiments were approved.
by the Hebrew University Committee for the Use of Human Subjects in Research. All participants provided written informed consent before their participation, after which they completed the cognitive assessment and AQ questionnaire (described below). They were then administered the finger-tapping task. All participants were financially compensated for their participation.

Cognitive assessments and AQ questionnaire

All participants completed a cognitive assessment, which evaluated non-verbal reasoning skills with the Block Design visuospatial reasoning task (Wechsler, 2008) and basic linguistic skills with single-word and paragraph reading. Single-word reading was assessed using a list of 24 single words in Hebrew (Deutsch & Bentin, 1996). Participants were instructed to read the words aloud, as quickly and accurately as possible. The paragraph reading was measured by a four-paragraph academic-level text (Ben-Yehudah et al., 2001). Participants were instructed to read the text aloud, as quickly and accurately as possible but slowly enough to be able to answer a simple content question at the end. Both accuracy and rate were scored. We chose these measurements based on our extensive experience in assessing the verbal skills of individuals with dyslexia, some of whom have a history of delayed verbal skills. We estimate that these measures most accurately characterize adult skills and are less sensitive to a tendency for verbal elaboration than vocabulary scores. All participants also completed the AQ (Autism Spectrum Quotient) questionnaire, a self-report measure aimed to quantify participants’ estimation of their autistic traits.

As shown in Table 1, the groups were matched for age, non-verbal reasoning skills, and reading accuracy. Reading rate of a paragraph was slower in the ASC group, as expected where syntactic and semantic information facilitate reading rate (Brock & Caruana, 2014; O’Connor & Klein, 2004). By contrast, AQ scores were substantially higher than TDs’, as expected (Baron-Cohen et al., 2001). All participants were native Hebrew speakers with no language disabilities.

Protocol of the tapping task

As illustrated in Figure 2, each tapping sequence consisted of either 18 or 28 metronome beats. All sequences began with a baseline Inter Stimulus Interval (ISI) of 500 ms (2 Hz) for 16 beats. Then the tempo could either remain constant or change (to 410, 450, 550, and 590 ms ISI). When the tempo changed, it lasted either 2 or 12 beats. Altogether, 10 types of sequences (5 tempos × 2 metronome durations) were administered eight times. The sequences were divided into two blocks separated by a 15-min break.

Participants listened to these sequences through headphones at a comfortable presentation level and tapped with their right index finger on a custom-made wooden box. A microphone, installed in the box, recorded the participant’s taps on the box. We used Focusrite Saffire 6 USB, which simultaneously recorded the output from the microphone and the headphone signals. In this way, the overall latency and jitter in measuring tapping onset is small (~2 ms; Anglada-Tort et al., 2022). Participants were instructed to start tapping in synchrony with the third beat of each sequence, and to continue tapping after

<table>
<thead>
<tr>
<th>TABLE 1 Cognitive and AQ50 scores of the two groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>TDs (n = 30) female = 11</strong></td>
</tr>
<tr>
<td><strong>ASDs (n = 30) female = 6</strong></td>
</tr>
<tr>
<td><strong>Group difference Mann–Whitney (effect size-Cliff’s delta)</strong></td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
<tr>
<td>General cognitive test (scaled WAIS score)</td>
</tr>
<tr>
<td>Block design (mean = 10, SD = 3)</td>
</tr>
<tr>
<td>Reading accuracy (% correct)</td>
</tr>
<tr>
<td>Word accuracy</td>
</tr>
<tr>
<td>Paragraph accuracy</td>
</tr>
<tr>
<td>Reading rate (words/minute)</td>
</tr>
<tr>
<td>Paragraph rate</td>
</tr>
<tr>
<td>Autistic traits</td>
</tr>
<tr>
<td>AQ50 Range: 0–200</td>
</tr>
</tbody>
</table>

Note: The two groups were matched for spatial reasoning scores (both groups’ scores were above average, scaled score = 10) and reading accuracy. The ASC participants were somewhat slower paragraph readers, and, as expected, had significantly higher AQ50 scores. All the assessments were calculated with 30 individuals with ASC and 30 TDs, except for the AQ50, which was calculated with 28 TDs (the scores of 2 participants were lost due to technical issues). Bold indicates a significant difference (p-value < 0.05).
the metronome stopped with the last tempo that they heard, until they heard a low-pitched sound, which signaled them to stop tapping.

Tapping analyses

Tapping onsets were extracted from the stereo audio signal using a MATLAB script. In the synchronization phase of the experiment, we excluded taps that were outside a window of ±200 ms from the nearest metronome beat, as is common practice (Repp, 2005; Vishne et al., 2021). Additionally, we excluded taps that deviated from the required ISI by more than ±280 milliseconds from our analysis in order to eliminate outliers and ensure consistency in the percentage of the outliers between the continuation and synchronization phases of the experiment. None of our results depend on this specific (somewhat arbitrary) choice, that is, choosing 250 or 300 ms yields similar statistics. Sequences where more than 50% of the taps were excluded from further analysis, and participants were excluded if 5% or more of their sequences (more than 4 sequences) were excluded based on the above criteria. Overall, there was a small number of excluded sequences (less than 1%) and taps (less than 5%) (mean fraction ± SEM: TD: 0.08 ± 0.03 ASC: 0.37 ± 0.10, Mann–Whitney: p = 0.28). All statistical analyses were performed using MATLAB (version 2021a). The non-parametric, Mann–Whitney U Test was used due to the high variability in the ASC group. In addition, the sensitivity index (d’) from signal detection theory was used to examine the change in tracking rate after environmental changes. Specifically, we used the delay interval (Figure 1), and compared mean delays before and after the tempo change, divided by the mean of the standard deviations. To examine patterns of error correction in two consecutive asynchronies, we fitted a 2-d Gaussian model, after which we observed the two main axis of the corresponding ellipsoid (the two eigen-vectors of the covariance matrix).

RESULTS

Synchronization versus internal temporal reliability

At the beginning of each sequence, participants tapped in synchrony with the metronome, with an ISI of 500 ms. The two groups had a similar mean asynchrony during this phase (Median [interquartile range] (ms): TD −37.1 [16.7], ASC: −28.3 [34]; Mann–Whitney U test: p = 0.2, though, as shown in Figure 3a, cross-participant variability was larger in the ASC group). Yet, the within-participant variability (measured as SD) around this mean was significantly larger in the ASC group (Median [interquartile range] (ms): TD:26.4 [11.2], ASC: 39.2 [20.8], Mann–Whitney U test: p = 0.002), in line with Vishne et al. (2021). As shown in Figure 3b, the variability of 80% (24/30) of the ASC participants was larger than the TDs’ median variability.

According to the computational model of Vishne et al. (2021), ASCs’ larger variability around the mean asynchrony stems from reduced online error correction (slow update). This means that we expect higher correlations between consecutive asynchronies (errors). If errors are fully corrected, we expect 0 correlation between consecutive asynchronies. By contrast, if an error is kept across taps and is not corrected at all, the expected correlation is 1 (assuming no noise). In general, higher correlations indicate reduced error correction. As shown in Figure 4a, the correlation between consecutive asynchronies was significantly higher in the ASC group (Median [interquartile range] (ms):TD: 0.4 [0.2], ASC: 0.5 [0.2], Mann–Whitney U test: p = 0.04). Similar results are found when we fitted a Gaussian distribution to the

FIGURE 2 A schematic illustration of the temporal structure of long and short sequences with a tempo change. All sequences began with a fixed Inter Stimulus Interval (ISI) of 500 ms for 16 beats. Then the tempo accelerated (shortened intervals by 50/90 ms), decelerated (elongated intervals by 50/90 ms) or remained the same, for 12 (top) or 2 (bottom) additional beats. Then, the metronome stopped (right vertical dashed bars). Participants were instructed (in advance) to continue tapping with the last metronome’s tempo. The duration of this continuation was 20 taps (8.2–11.8 s, depending on the condition), and its end was denoted by a clear low-pitched sound. Each black bar represents a metronome beat.
tap and its subsequent tap and observed the two main axis (eigenvectors) of this distribution (Figure 4b shows the cluster of all consecutive pairs). Note that the first main axis component is a measure of the main direction of variation in a dataset, while the second axis captures additional patterns of variation in the error. This ratio is significantly larger in the ASC group (TD: 3.1, ASC: 4.5, Mann–Whitney U test: $p = 0.02$), indicating that an error in a given tap (t) is more similar to the error in the next tap (t + 1) in the ASC group. This means that individuals with ASC have greater reliance on the previous tap and are less able to adjust and correct their tapping in response to errors. These two analyses indicate ASCs' reduced online error correction.

As shown in Figure 3, the ASC group shows greater variability around their mean asynchrony. Figure 4 shows that their online error correction is, on average, smaller than TDs'. Together these analyses suggest that ASCs' errors are corrected to a lesser extent, and hence tend to increase before they are substantially corrected. Overall, most mistakes, which are defined as the absolute difference between the asynchrony of a given tap and the mean asynchrony of the participant, are small in both groups (<30 ms). However, there are large mistakes (>70 ms) as well. The range of small mistakes was defined as <30 ms since the distribution of mistake magnitudes was not bell-shaped, that is, extending the range from 30 to 40 ms resulted in only a 10% increase in mistakes. To maintain consistency, large mistakes were defined as those between 70 and 100 ms. Together, small and large mistakes make up approximately 80% of all mistakes, with the remaining ~20% being intermediate mistakes (between 30 and 70 ms). As shown in Figure 5—represented as log of the number of mistakes for each participant—ASCs have fewer small-mistakes (Figure 5a, Median [interquartile range] (ms): TD: 6.9 [0.2], ASC: 6.8 [0.3], Mann–Whitney U test: $p = 0.02$), and more large-mistakes (Figure 5b, Median [interquartile range] (ms): TD: 2.8 [1.6], ASC: 3.7 [1.3], Mann–Whitney U test: $p = 0.006$), and the ratio between them is significantly higher than TDs' (Figure 5c, Median [interquartile range] (ms))

**Figure 4**  The correlation between consecutive asynchronies (synchronization errors) is larger in the ASC group, indicating a smaller fraction of error correction between consecutive taps. (a) Single participant correlations are larger in the ASC group. The median of each group is denoted as a black horizontal bar; error bars around this median denote the interquartile range

**Figure 3**  Tapping during the synchronization phase: mean asynchrony (a) is similar in the two groups, but the standard deviation around this mean (b) is significantly higher in the ASC than in the TD group. Each dot represents the performance of one participant. In general, the performance of ASC participants is more broadly distributed than TDs'. The median of each group is denoted as a black horizontal bar; error bars around this median denote the interquartile range

**Figure 5**  The correlation between the reliability of internal beat and of motor responses in the two groups

When the metronome stopped, participants were asked to keep its pace. In the isochronous sequences its tempo was 2 Hz–500 ms between consecutive beats. Since there was no external feedback, the participants had to rely on their internal tempo. In order to measure the ability to
keep a fixed tempo without a metronome, we used an additional measure—ITI (Inter Tap Interval), which reflects the participant’s consistency in tapping intervals.

As shown in Figure 6, during the synchronization phase (Figure 6a), like the asynchrony, ITI had a similar mean for the two groups—500 ms (Median [interquartile range] (ms): TD: 498.7 [2.5], ASC: 499.1 [2.6], Mann–Whitney U test: p = 0.3). Yet the ASC group had a significantly larger SD around this mean (Median [interquartile range] (ms): TD: 25.9 [9], ASC: 29.6 [11.5], Mann–Whitney U test: p = 0.02). In contrast, during the continuation phase (Figure 6b), there was no group difference, neither in the mean ITI nor in the SD around this mean. Mean ITI was reliably kept by both groups (Median [interquartile range] (ms): TD: 500.7 [12.9], ASC: 499.6 [13.8], Mann–Whitney U test: p = 0.4). SD around this mean was also similar in both groups (Median [interquartile range] (ms): TD: 23.9 [4.7], ASC: 26.1 [10.1], Mann–Whitney U test: p = 0.3), indicating the adequacy of ASCs’ ability to tap based on internal representations of intervals.

To assess the interaction between group and synchrony state, namely, whether turning off the metronome had a significantly larger effect on variability in the ASC group than in the TD group, we calculated, for each participant, the difference between the SD in the synchronization phase (Figure 6a) and the continuation phase (Figure 6b). The difference in SD between with and without a metronome is significantly larger in the ASC group (Figure 6c; Median [interquartile range] (ms): TD: –2.5 SD, ASC: 2.4 SD, Mann–Whitney U test: p = 0.048).

**FIGURE 5** The ASC group has a larger number of large mistakes and fewer small mistakes compared to the TD group. The number of mistakes is shown in log scale. (a) The number of the small mistakes (<30 ms) is smaller in the ASC group. (b) The number of the large mistakes (>70 ms) is higher in the ASC group. (c) The ratio between the number of large mistakes and the number of small mistakes is larger in the ASC group. The median of each group is denoted as a black bar; error bars around this median denote the interquartile range.

**FIGURE 6** ITIs with the metronome (synchronization) and without the metronome (continuation)—stopping the metronome has a larger effect on the ASCs’ standard deviation. Mean and SD of ITI of each group, (a) During the synchronization phase. (b) During the continuation phase. Participants of both groups manage to reliably keep the mean metronome tempo, and mean ITI is similar in the two groups in both phases. However, SD is significantly larger in the ASC group during the synchronization phase, whereas during the continuation phase the variability is similar in the two groups. (c) The difference in participants’ SD between the synchronization and continuation phases. In both groups SD is larger in the synchronization phase, where errors are corrected both for mean ITI and for synchronization with the metronome. The reduction of SD is significantly larger in the ASC group, which corrects less during synchronization. Each dot represents the performance of one participant. The median of each group is denoted as a black horizontal bar; error bars around this median denote the interquartile range.
[5.4], ASC: −6.2 [6.9], Mann–Whitney U test; p = 0.048). Interestingly, in both groups, SD is smaller in the continuation phase, indicating a reliable rhythmic continuation, and an additional variability in the synchronization phase due to the need to keep online synchrony with the external metronome.

The ASC group needed more metronome beats to update a change in tempo

To assess the participants’ rate of tracking changes in the external rhythms, we introduced tempo changes, acceleration or deceleration, each with one of two step sizes (Figure 2). After these tempo changes, the metronome continued for either a very brief period (2 beats) or for a longer one (12 beats), allowing us to measure whether these periods were sufficient to internalize the new metronome rhythm, and hence keep it reliably even after the metronome had stopped. For this analysis, we used the delay interval, which is the time between the last beat of the metronome and the next tap (as shown in Figure 1). The delay is the complement of the asynchrony and, as such, will yield the same results as the asynchrony except in cases where there is an environmental change (e.g., a change in tempo or a stop in the metronome). In these cases, the delay has an advantage because it is aligned with the last metronome beat, which is the last external stimuli that the participant can respond to in the current response tap. Thus, the delay is aligned with the subject’s perspective. As shown in Figure 7a, updating of the tempo was not immediate in either group. Although the

![Graphs showing delay intervals and distributions](https://onlinelibrary.wiley.com/doi/10.1002/aur.2926)

**Figure 7** The rate of tracking metronome acceleration is slower in ASC. (a) Mean group delay following tempo changes (baselined—i.e. after subtracting mean asynchrony) Top: the two acceleration conditions. Bottom: the two deceleration conditions. Left: large tempo changes (± 90). Right: small tempo changes (±50). The y-axis measures the delay interval in each beat, aligned to the last beat. The x-axis values are the tap number: Beat 0 is the first metronome beat with the new tempo. Participants’ delay still has not changed, since they did not predict the tempo change when planning the tap. Beat 1 is the first that shows (partial) update. Black bars indicate the metronome intervals. Error bars denote cross participants’ standard errors. (b) Distributions of delay intervals of all participants before (beats 4–12) and after the tempo change (beats 4–12 after the change). The mean of each distribution is denoted by a black vertical line. The ASC group did not adapt fully to changes in acceleration tempo.
tempo change always occurred at beat 0 (after 16 beats), the participants did not predict it and they did not track the number of beats. The result that participants did not predict the change is expected given that the tempo change only occurred in 80% of the sequences, and when it did, it could involve lengthening or shortening of the ISI. Hence, at beat 0, their delay did not yet change. At beat 1, participants already received information about the change, since they could detect the pattern in the experiment where only one tempo change in a sequence is present, yet they updated only partially at this beat. The TDs fully updated within two additional taps. The ASCs updated at a similar rate in deceleration, but updated only partially in acceleration, as shown at the top of Figure 7a. For clarity of presentation, we subtracted the pre-change delay to 500 ms—the tempo before the tempo change, plotting only the change in delay.

To quantify the rate of update to the change in tempo, we defined a sensitivity index $d'$ for each participant for each tempo change, namely, the difference between the mean delays before and after the tempo change, divided by the mean of the standard deviations (Figure 7b). Higher values indicate better separation between the participants' delays before and after the tempo change, indicating better updating to the new tempo. The $d'$ was calculated for long trials only, for taps 4–12 before and after the tempo change (the first three taps at the start of the sequence and immediately after the change were excluded, since tempo was not yet stable, as shown in Figure 7a). Moreover, this index allowed us to average performance across the two acceleration steps, which showed similar trends, and across the two deceleration conditions, which also showed similar trends (Figure 7a). We found that the groups' reliability of updates differed only in the acceleration (Median $[\text{interquartile range}]$ (ms): TD: 2.4 $[0.8]$, ASC: 1.9 $[1.5]$, Mann–Whitney $U$ test: $p = 0.04$). No difference between the groups was found in tempo deceleration (Median $[\text{interquartile range}]$ (ms): TD: 1.9 $[1.2]$, ASC: 1.2 $[1.6]$, Mann–Whitney $U$ test: $p = 0.2$).

Next, we characterized the continuation phase. To characterize both the immediate update and subsequent updating we analyzed the ITI and the delay separately—for the first taps and for the following (steady state) taps. For the initial responses we calculated the SD of delay 1—the first delay that participants responded to with the metronome stop (since the metronome stop was unexpected in delay 0, the participants did not yet respond to the change, Figure 8). For the rest of the continuation taps, we calculated the absolute error (distance) of the ITI in each tap, as detailed below (Figure 9). In both the initial and steady state tapping, we found updating difficulties in the ASC group in the short, but not in the long, metronome sequences following the tempo change. Namely, given an additional adapting window (of ~5 s), participants with ASC successfully integrated the new tempo.

Analyzing the delay of tap 1 (Figure 8), we found that the SD was larger in the ASC than in the TD group, in short acceleration sequences (Median $[\text{interquartile range}]$ (ms): TD: 53.2 $[26.6]$, ASC: 66.7 $[42.3]$, Mann–Whitney $U$ test: $p < 0.01$). No difference between the groups was found in long sequences (Median $[\text{interquartile range}]$ (ms): TD: 50.2 $[35.8]$, ASC: 56.1 $[39]$, Mann–Whitney $U$ test: $p = 0.3$). The interaction between group and sequence length was significant. Namely, the difference in SD between short and long sequences significantly differed between groups (Median $[\text{interquartile range}]$ (ms): TD: 5.7 $[39.8]$, ASC: 12 $[23.8]$, Mann–Whitney $U$ test: $p = 0.04$). No group difference was found in deceleration sequences, where we found similar performances in both the short (Median $[\text{interquartile range}]$ (ms): TD: 82.2 $[39]$, ASC: 73.8 $[34]$, Mann–Whitney $U$ test: $p = 1$), and the long sequences (Median $[\text{interquartile range}]$ (ms): TD: 50.4 $[28.2]$, ASC: 60.1 $[34.8]$, Mann–Whitney $U$ test: $p = 0.2$), and there was no interaction between group and sequence length (Median $[\text{interquartile range}]$ (ms): TD: 25.6 $[49.2]$, ASC: 20.4 $[39.4]$, Mann–Whitney $U$ test: $p = 0.3$).

To characterize the steady state taps (Figure 9), namely, whether participants manage to reliably

![Figure 8](https://example.com/image8.png)

**Figure 8** Standard deviation in the first tap after the metronome stops (delay 1). Metronome stopped either shortly (2 beats), or few seconds (12 beats) after its tempo acceleration. (a) Short metronome continuation—the ASC group had a larger SD compared to the TD group. (b) Metronome continued for several seconds (12 beats)—the two groups had similar SD. (c) The interaction of group × metronome duration following tempo switch is significant—the ASC group reduced their SD significantly more than the TDs between the short and long conditions. The median of each group is denoted as a black bar; error bars around this median denote the interquartile range.
integrate the modified tempo within a few taps after the metronome stopped, we assessed the ITIs. We quantified the distance between the attained ITIs and the required one (the last metronome tempo). The ASC group showed a large variability in their tapping intervals, particularly when the metronome stopped shortly after the tempo change. Some participants had a high SD around the required ITI, whereas others had a small SD around a mean that was different from the required ITI. To capture broad variability, we defined a distance index—the tap’s (absolute) distance from the required ITI—and averaged it for each participant, for short and long sequences separately. There was a marginally larger variability in the ASC group in the short sequences (Median [interquartile range] (ms): TD: 8.4 [9], ASC: 11.6 [12.2]). Mann–Whitney U test: \( p = 0.061 \). Namely, here too the ASC group gained more from elongating the sequences before the metronome stopped.

DISCUSSION

In this study, we aimed to test two opposing theories, the elevated internal noise account and the slow update account, by using a synchronization-continuation tapping task in both fixed and changing tempo conditions. This task required the retention of internal representations of metronome beats. Additionally, we investigated whether elongating the tempo-update period allows participants with ASD to keep the updated tempo reliably. We tested this by comparing two continuation conditions, with short and long update periods, respectively.

Our results supported the slow update hypothesis (Lieder et al., 2019; Vishne et al., 2021), which predicts that difficulties in ASC will be specific to the required rate of error correction when interacting with the external world. Hence, when there is no requirement for synchronization with an external stimulus, people with ASC will not have difficulties. We found that individuals with ASC had difficulties in correcting errors and tracking external tempo changes, particularly tempo accelerations. Furthermore, our results showed that, although maintaining an internal pace requires enhanced cognitive memory resources (Perbal et al., 2002), individuals with ASC performed reliably when no online updating was required and when they were given a few seconds to update their internal tempo and retain it. These findings suggest that a slow update account may explain the main difficulties observed in both the automatic isochronous sequences and in the more demanding sequences with tempo changes, which require cognitive processing.

Our results are consistent with previous studies that found synchronization difficulties of individuals with ASC, manifested in larger SD (Morimoto et al., 2018; Vishne et al., 2021). Our observation that memory load does not increase the difficulties of ASC is in line with a recent study (Tryfon et al., 2017) that at first appears to contradict the current study. In this study, 31 children with ASC and 23 age-matched TDs listened several times to rhythms of varying metrical complexity and then tapped in synchrony with them. This study found no group difference. It is counter-intuitive that online synchronization to simple fixed beats is challenging for ASC individuals while the production of complex rhythms is adequate. However, it is consistent with our current findings that the online requirement is the main impediment in ASC. Tapping a complex tune requires some practice and relies largely on learned internal rhythms (since the next interval is not equal to the previous one), allowing
longer update durations and more predictability in the sequence that can be exploited.

Similarly, in a recent study that administered a reproduction task (Edey et al., 2019), the performance of adult participants with autism was not worse than TDs’ (25 adults with ASC and 24 matched age TDs). Participants were presented with a series of four tones with fixed inter-tone intervals (either 600 or 900 ms). They were instructed to listen to the first two tones and press in synchrony with the third and fourth tones. There was no group difference in response accuracy (to auditory stimuli), again indicating adequate representation and reproduction of temporal intervals. The adequacy of representation of intervals was also tested directly (Poole et al., 2022) in a battery of auditory and visual timing tasks (at the sub-second range), administered to adults with ASC and TDs. These tasks measured interval-timing sensitivity by temporal discrimination threshold tasks (participants made relative judgments about intervals presented using the adaptive method), time estimation (participants made direct judgments about different durations in milliseconds range), event-timing sensitivity (participants determined the order of two stimuli), and interval memory (participants were presented with a standard duration which they were asked to memorize and subsequently judge whether comparison durations were the same or different from the standard duration). In addition, participants completed a self-reported questionnaire about timing behaviors in daily life. Although participants with ASC tended to report more difficulties with timing in daily life, no significant group difference was found in any of the measurements, which focused on seconds and sub-second intervals.

Taken together, our observation that people with autism have an adequate representation of intervals at the sub-second to second range, while their ability to synchronize with an external signal is hampered due to slow error-correction, is in line with the previous literature. One deviation is a tapping study of children with autism, where tapping was administered to 18 children (ages 7–15 years) with ASC and 11 age-matched TDs. The paradigm of this task was synchronization-continuation. No consistent difference in the mean or SD of the asynchrony was found in the synchronization phase. However, in the continuation phase, differences between the groups were found, though only in the younger children (Sheridan & McAuley, 1997). In our study, we tested only adults. Yet, a previous study with children (Morimoto et al., 2018) found synchronization difficulties in the ASC group. Hence, given the small sample size and large age range during which tapping skills substantially improve, (Monier & Droit-Volet, 2018; Monier & Droit-Volet, 2019; Thompson et al., 2015), replications are required for better understanding. Additionally, since synchronization skills might relate to cognitive skills (Bailey & Penhune, 2010), the lack of cognitive matching may have also affected group differences.

The current study investigates the performance of individuals with autism in the synchronization-continuation tapping task, which requires repetitive taps at a fixed interval, and as such might seem similar to repetitive restricted behaviors (RRB), a common trait of autism. Yet, these are very different behaviors. First, RRB can be classified into two categories: “lower-order” behaviors, including repetitive motor movements and sensory stimuli responses, and “higher-order” behaviors, involving a need for sameness and limited interests (Leekam et al., 2011; Turner, 1999). Studies have shown that lower-order RRB is more frequent in young children with autism and is often associated with intellectual disability (Esbensen et al., 2009). Accordingly, these behaviors tend to decrease with age (Barrett et al., 2018). However, the participants in the current study were high-functioning adults without intellectual disabilities. Second, RRB is an involuntary behavior (Freeman et al., 2010), and individuals with ASC often report a lack of awareness of the initiation of these behaviors (Kapp et al., 2019), while the synchronization-continuation task requires intentional and accurate performance. Interestingly, repetitive behaviors have also been reported in disorders such as Parkinson’s disease, yet people with Parkinson’s have difficulties with the continuation part of the synchronization-continuation tapping, unlike those with ASC (Benoit et al., 2014; Harrington et al., 1998). Moreover, the slow update of both perceptual and motor plans in autism leads to difficulties adapting to rapidly changing environmental cues (including, but not specific to human gestures). As a result, individuals with autism may resort to implicit strategies, such as engaging in repetitive behaviors and limiting attention tounchanging fields, in order to increase their exposure to predictable signals.

This study has a number of limitations. One is the exclusive inclusion of individuals with high-functioning autism, which may limit the generalizability of the findings to the autism spectrum in general. A second limitation is the use of a specific tempo. We chose 500 ms intervals and its temporal vicinity, since this is the most natural tempo for humans. However, the use of other intervals could potentially reveal continuation difficulties.

To conclude, we found that while individuals with ASC have difficulties in synchronizing to an external stimulus, they have adequate tempo-keeping when the metronome stops, and adequate keeping of a modified tempo after the metronome stops when given several seconds of updating with the metronome. These results suggest that temporal internal representations in autism are adequate and do not have elevated noise levels, while synchronizing with external beats is impaired due to reduced online error correction.

ACKNOWLEDGMENTS

We thank Gal Vishne for her help in the initial stages of the project.
FUNDING INFORMATION
The European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Program Grant Agreement 833694 and the Israel Science Foundation Grant 1650/17. Both awarded to Merav Ahissar.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT
The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All experiments were approved by the Hebrew University Committee for the Use of Human Subjects in Research. All participants provided written informed consent before their participation, after the instructions about the experiment were provided.

ORCID
Keren Kasten https://orcid.org/0000-0002-1504-807X

REFERENCES


