Perception of timing is more context sensitive than sensorimotor synchronization

BRUNO H. REPP
Haskins Laboratories, New Haven, Connecticut

In three experiments, the effects of contextual temporal variation on the perception of timing and on sensorimotor synchronization were examined. Experiment 1 showed that exposure to a variably timed auditory precursor sequence reduces the detectability of deviations from isochrony in a musical test sequence. By contrast, in Experiment 2 there was only a small and transient effect of identical precursor sequences on the variability of finger taps that were synchronized with a similar test sequence. Moreover, the precursor did not impede phase error correction following deviations from isochrony in the test sequence. Experiment 3 employed a within-subjects design that required simultaneous detection of irregularities in and synchronization with nonmusical auditory sequences. Precursor variability impaired only detection, not synchronization performance. These results suggest that perception of deviations from regularity engages context-sensitive timing processes, probably related to conscious awareness, that are not involved in sensorimotor synchronization.

Perception of Timing as Internal Synchronization

The human ability to synchronize a repetitive action (such as finger tapping) with a regular sequence of events (such as tones) seems to have much in common with the ability to perceive temporal regularity in such a sequence. Synchronized movement requires an internal timekeeper or oscillator that times actions so as to make them coincide with the expected times of future events (see, e.g., Mates, 1994; Pressing, 1999; Vorberg & Wing, 1996).[1] Perception of temporal regularity in an event sequence requires a similar internal process that generates expectations of when future events should occur (Barnes & Jones, 2000; Jones, 1976; Jones & Boltz, 1989; Large & Jones, 1999). Asynchronies between taps and sequence events in synchronization are analogous to discrepancies between temporal expectations and sequence events in perceptual monitoring. Maintenance of synchronization with an isochronous sequence requires phase error correction (Mates, 1994; Vorberg & Wing, 1996). Likewise, it has been proposed that an attentional oscillator entrained to a sequence can adapt its phase (Large & Jones, 1999). Perception of timing thus may be regarded as internal synchronization, without any overt motor activity. Conversely, sensorimotor synchronization might be regarded as a form of externalized time perception (Michon, 1967).

One important theoretical issue is whether perception and action are governed by a single general timekeeping process or whether separate, perhaps even task-specific, timers are involved. Some studies have found significant correlations between participants' relative accuracy in time perception and time production tasks, which may reflect individual differences in the variability of a general underlying process (Keele, Pokorny, Corcos, & Ivry, 1985; Smith, 1957). Cerebellar damage has been found to impair performance on both types of task (Ivry, 1997; Ivry & Keele, 1989). Ivry and Hazeltine (1995) observed similar increases in variability as a function of interval duration in interval perception and production, which to them suggested a common timing process. Internal oscillators with similar properties have been claimed to underlie time perception and the timing of movement (Treisman, Faulkner, & Naish, 1992). Meegan, Aslin, and Jacobs (2000) found that training on an interval discrimination task significantly increased the accuracy of interval production for the specific interval duration that had been used as the standard during training. A recent brain imaging study (Schubotz, Friederici, & von Cramon, 2000) obtained evidence that similar brain structures are engaged in time perception and motor timing.

However, there are also arguments against a general timing process. For example, Kolers and Brewster (1985) found that sensorimotor synchronization is more accurate in the auditory than in the visual modality and concluded that different timekeeping processes are involved. Their views are close to a dynamic systems perspective, according to which timing is an emergent property of action and thus is specific to different motor or perceptual tasks (Turvey, 1977; Wallace, 1996). Evidence in support of that position has recently been presented by Robertson et al. (1999) and Zelaznik, Spencer, and Doffin (2000), who found that there was little relation between timing accuracy in finger-tapping and periodic drawing tasks. Ivry and Richardson (2002) have proposed a multiple-timer

model that assumes functionally similar but nonidentical timers for different limbs carrying out similar actions. Nevertheless, there may be a close relationship between perception and action within the same task situation (see, e.g., Viviani & Stucchi, 1989, 1992a, 1992b). Robertson et al. attributed earlier findings of a correlation between finger-tapping accuracy and the relative accuracy of temporal judgment to the relatively discrete nature of finger tapping, which may involve time estimation more than does continuous motor control. Thus, it may still be true that perception of timing and movement timing in similar situations rely on the same underlying timekeeping processes.

Even though timing processes underlying perception and action may be identical in certain situations, additional task-specific processes are involved in each mode of behavior. In particular, perceptual tasks generally require judgments about temporal differences, which means that these differences must reach conscious awareness. Such awareness is not required in sensorimotor synchronization, where subliminal temporal changes can effectively guide action (Repp, 2000a; Thaut, Tian, & Azimi-Sadjadi, 1998). The very existence of perceptual thresholds for temporal differences suggests that temporal information undergoes additional transmission or processing before it reaches awareness, whereas action systems evidently have access to this information at an earlier stage. This additional processing may make perception more vulnerable to contextual influences than sensorimotor synchronization, even if some basic timing processes are shared.

The present research addressed this issue by investigating the relative sensitivities of perception of timing and of action timing to contextual timing variability. The contextual manipulation is described in more detail below. The perceptual task (Experiment 1) required detection of deviations from temporal regularity in auditory sequences. The action timing task (Experiment 2) was the synchronization of finger taps with the same sequences. The two tasks were carried out simultaneously in Experiment 3. The dependent variable in the perceptual task was detection accuracy (percentage correct). The dependent variable in the synchronization task was the variability of the finger taps. Both dependent variables were assumed to reflect mainly central variability (although there are peripheral sources of variability as well): Detectability of deviations from regularity necessarily depends on central variability, and a considerable portion of the variance of finger taps has also been attributed to central variability (Ivry & Hazeltine, 1995; Vorberg & Wing, 1996; Wing & Kristofferson, 1973). Part of that central variability is due to timekeeper variability. Therefore, if there is a common timekeeper underlying both tasks and if contextual variation increases the variability of that timekeeper, this should be reflected in both lower detection accuracy and higher tapping variability.

Effects of Contextual Timing Variability

The contextual influence investigated here is that of a precursor sequence exhibiting large timing variability on the behavioral response to a subsequent test sequence. Several previous studies have shown that exposure to such a variable context impedes detection of local deviations from regularity in a subsequent test sequence. Large and Jones (1999; Experiment 1) found that a reduction in temporal sensitivity was caused not only by variation in the immediately preceding context, but also by variation in other trials in the same session. Related long-term effects of contextual temporal variability on time perception were observed by Jones and Yee (1997) and Yee, Holleran, and Jones (1994). Large and Jones attributed these effects to a widening of the temporal expectancy region of an attentional oscillator and to a slow rate of adaptation of this parameter.

In a musical version of the same paradigm, Repp (1998c: Experiment 1) found that exposure to an expressively timed performance of a Chopin excerpt impeded detection of local deviations from regularity in a subsequent, mechanically timed performance of the same excerpt. Moreover, that effect did not seem to decline in the course of the test excerpt, which lasted nearly 20 sec. However, there was no control condition in Repp's (1998c) experiment; the context effect was gauged by comparing the results with those of an earlier experiment that had a somewhat different design and different participants, and in which there had not been any precursors. One motivation for the present Experiment 1 was to replicate the context effect in a within-subjects design that included both variably timed and isochronous precursors.

In contrast to these perceptual findings, one study found no effect of passive exposure to an expressively timed music precursor on the variability of finger tapping in synchrony with a subsequent, mechanically timed musical excerpt (Repp, 2000a: Experiments 3 and 4). This result suggested that perception of timing is more vulnerable to interference from contextual timing variability than is synchronized action. This could mean one of two things: Either the context effect arose at a basic timekeeping level, in which case perception and motor control must have engaged different timekeepers, or the context effect arose at a level specific to perception, in which case there could have been a common timekeeper underlying perception and action. Experiments 2 and 3 attempted to distinguish between these two possibilities by requiring participants to tap in synchrony with both the precursors and the test excerpts, so that the precursors engaged the motor timekeeper in the same irregular way as they engaged the perceptual timekeeper in Experiment 1. If synchronized tapping showed a context effect under these conditions, this would suggest that the effect derives from increased variability in a timekeeper specific to each activity. However, if synchronized tapping still failed to show any context effect, this would suggest that the context effect in perception does not arise at the level of basic timekeeping, but at an additional level of temporal processing that is specific to the perceptual task.

The test excerpts in all three experiments contained small deviations from regularity. In the perceptual task, these had to be detected, whereas in the synchronization task

they had to be compensated for in the timing of the taps (i.e., through phase correction) to maintain synchrony. Previous research (Repp, 1999b, 2000a, 2001) has demonstrated that phase correction is independent of conscious detection of irregularities or asynchronies and operates effectively below the detection threshold. Therefore, it was predicted that the speed of phase correction would be independent of any effect that contextual timing variation might have in the perceptual task. However, a context effect on tapping variability in the synchronization task might slow phase correction as well.

EXPERIMENT 1
Perception

Although Experiment 1 used musical test sequences for reasons of continuity with the author's previous research, the predicted effect of temporally varying context on time perception was not believed to be specific to music. To demonstrate its generality, Experiment 1 employed four kinds of precursor: (1) isochronous music, (2) expressively timed music, (3) isochronous clicks, and (4) expressively timed clicks (i.e., with the same timing pattern as the expressively timed music). It was expected that small hesitations in the musical test excerpts would be more difficult to detect after expressively timed than after isochronous precursors, and that expressively timed music and click

precursors would have comparable effects. On the basis of earlier results (Repp, 1998c), it was also expected that the context effect would persist throughout the test excerpt.

Method
Materials. The musical excerpt was the opening of Chopin's Etude in E major, op. 10, No. 3, a computer-generated score of which is shown at the top of Figure 1. (The final chord represents a modification of the original notation, intended to give closure to the excerpt.) The initial eighth-note upbeat was always present in the music but was excluded from all analyses and graphs because of its longer duration. All subsequent interonset intervals (IOIs) in the excerpt are nominally sixteenth-note intervals, as long as distinctions between voices are ignored. In the sound pattern, IOIs were defined as the intervals between the onsets of the tones corresponding to the highest notes in successive score positions (the "top line"). In the computer-controlled isochronous version of the excerpt that served as the starting point for the experimental materials, all the IOIs were 500 msec in duration, nominally simultaneous tones started at very nearly the same time, all the tones had durations corresponding to the note values in the score, and all the key depression velocities were set to the same arbitrary MIDI value of 60.[2]

From this isochronous version of the excerpt, test sequences were generated that were not entirely isochronous but contained four lengthened IOIs (i.e., hesitations). These four IOIs were lengthened by the same amount Δt and occurred at unpredictable locations, separated by at least four unchanged IOIs. All of the tones that were sounding during a lengthened IOI were lengthened by Δt as well, so that legato articulation was maintained. In the course of a block of nine trials, each of the 36 IOIs in the excerpt was lengthened once
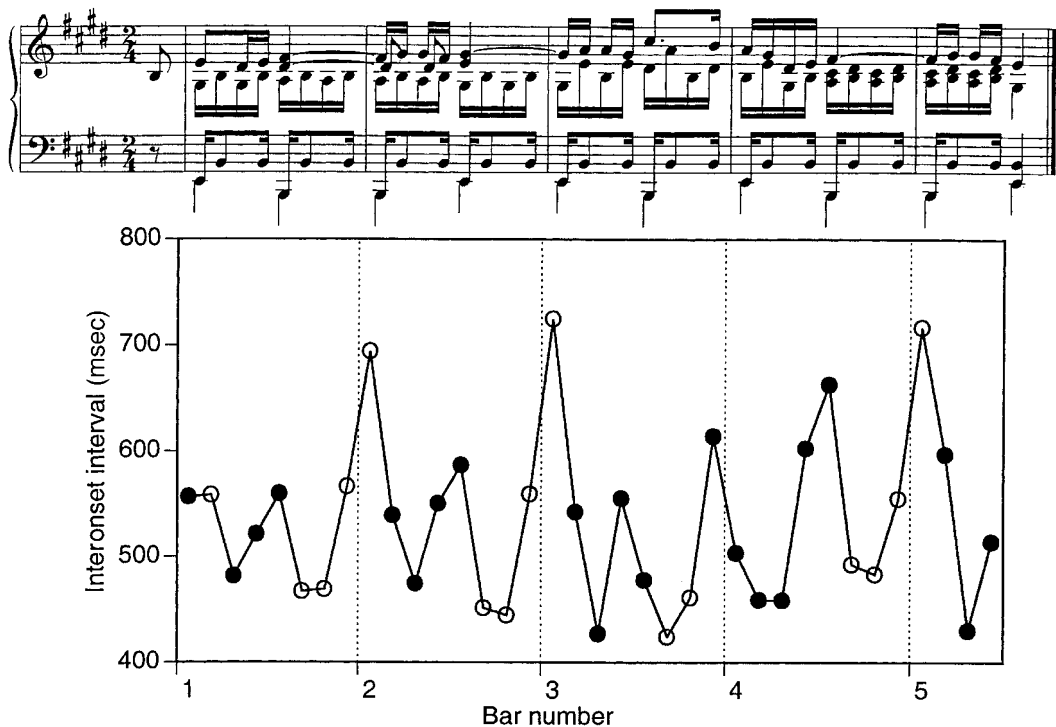

Figure 1. Computer-generated score of the opening of Etude in E major, op. 10, No. 3, by Frédéric Chopin, and the expressive timing profile used for the precursors in Experiment 1. Filled circles indicate interonset intervals (IOIs) initiated by melody tones; open circles indicate IOIs initiated by accompaniment tones.

by Δt. The Δt values ranged from 80 msec (16%) to 20 msec (4%) in steps of 10 msec. Three different trial blocks with different randomizations of the positions of the lengthened IOIs were created for each Δt value.

The expressively timed music precursor was a computer-controlled performance of the same Chopin excerpt with the pattern (or profile) of expressive timing shown in Figure 1. This timing pattern was derived from a principal component analysis of a large sample of expert performances (Repp, 1998b) and hence was musically appropriate, although it was not very typical. (Pattern typicality was assumed to be irrelevant.) It represented the fourth principal component of the performance analysis, whereas Repp's (1998c) previous study had used the second principal component pattern. The expressive timing pattern had a mean IOI duration of 533 msec and a standard deviation of 80 msec.[3] The precursor differed from the test excerpts in several other ways, which helped prevent their confusion: It contained typical expressive intensity variation, as well as small tone onset asynchronies and pedaling, which enhanced its naturalness and made it more appealing aesthetically than the "dead-pan" test excerpts.

The isochronous music precursor was derived from the expressively timed one by setting all IOIs between top-line tones to 500 msec, leaving all other temporal details (asynchronies, pedaling) relatively invariant. Details about how this was done may be found in Repp (2000b).

The expressively timed click precursor consisted of a sequence of 38 very high pitched (C8, 4168 Hz), rapidly decaying digital piano tones of equal intensity, with no nominal duration (i.e., MIDI "note offset") specified. They are called *clicks* here for convenience and because this term has been used in other studies by the author, although *pings* would be more accurate. Their onsets were timed in exactly the same way as those of the top-line tones of the expressively timed music. The isochronous click precursor had constant IOIs of 500 msec, except for the initial ("upbeat") IOI, which was twice as long.

All materials were generated from preassembled MIDI instructions on a Roland RD-250s digital piano under control of a Macintosh Quadra 660AV computer via a MIDI interface. The experiment was run using a program written in MAX. The participants listened binaurally over Sennheiser HD540 II earphones.

Participants. The participants were 12 paid volunteers, 19–33 years old, mostly summer students at Yale University, who had responded to a campus advertisement. They represented a wide range of musical training, from none at all to as much as 26 years of instruction (on several instruments combined).

Procedure. The participants came for two individual sessions on different days. On the 1st day, they were given written instructions and listened to a completely isochronous version of the Chopin excerpt a few times for familiarization. Then a pretest consisting of three blocks, each containing nine test trials without precursors, was presented. The participants were told that each trial contained four hesitations in the regular rhythm of the music. The task was to press the space bar on the computer keyboard whenever a hesitation was

perceived. The computer monitor provided immediate visual feedback for each correct response and showed counters in which hits and false alarms were added up within each block. The participants were told to respond quickly (any response slower than 1 sec or faster than 100 msec was counted as a false alarm) and to make as few false alarms as possible (i.e., not to guess randomly). The first block used the largest Δt value (80 msec). On the basis of the total score for that block, the experimenter chose the Δt value of the second block, and similarly for the third block, the aim being to find a difficulty level at which the participant detected about 60% of the hesitations. This was done in order to avoid ceiling or floor effects for different positions in the music; the participants' absolute accuracy was not of interest. The data of this pretest were not analyzed further.

The first half of the main part of the experiment followed. It consisted of six blocks in which each of the nine trials was preceded by a precursor, either music (for half the participants) or clicks (for the other half). The participants were asked merely to listen to the precursors; no response of any kind was required. A nominal silence of 2 sec intervened between a precursor and a test sequence, and one of 3 sec between a test sequence and the next precursor.[4] The Δt value was the same in all six blocks, and the precursor remained constant within each block. Blocks with isochronous and expressively timed precursors alternated, according to a counterbalanced design. Blocks containing different randomizations of the positions of hesitations were likewise balanced across the two precursor conditions. The fixed Δt values assigned to the 12 participants on the basis of their pretest performance were 60 msec (2), 50 msec (2), 40 msec (2), 30 msec (5), and 20 msec (1). This wide range is representative of the range of detection thresholds (50% correct) obtained in this task for unpracticed listeners with varying musical backgrounds (Repp, 1999b).

The second session began with one practice block without precursors, using a Δt value 10 msec longer than that in the test blocks of the preceding session. Then six test blocks followed, using the same Δt value as the test blocks in the first session. Those participants who had received music precursors in the first session now received click precursors, and vice versa.

Results and Discussion

Each participant had three opportunities to detect a hesitation in each of 36 positions in the music in each precursor condition. The correct detection scores (hits) were submitted to a repeated measures analysis of variance (ANOVA), with the fixed variables of precursor type (music vs. clicks), precursor timing (isochronous vs. expressive), and position (36 levels). The main results are summarized in Table 1 in terms of hit percentages averaged across positions and participants. As was predicted, hit percentages were lower in the expressively timed than in the isochronous precursor conditions [$F(1,11) = 27.7$,

Table 1
Average Percentages of Hits and False Alarms (FAs) in the Four Precursor Conditions of Experiment 1

Precursor Type	Precursor Timing					
	Isochronous		Expressive		Difference	
	Hits	FAs	Hits	FAs	(Iso − Exp)	SE
Clicks	69.4	2.3	62.7	2.1	6.7	1.6
Music	64.5	2.1	52.9	2.2	11.6	2.5

Note—Differences between precursor conditions are shown in the last column, with standard errors.

$p < .0004$]. Unexpectedly, hit percentages were also lower after music precursors than after click precursors [$F(1,11) = 10.2, p < .009$]. The effect of precursor timing was larger for music than for click precursors, but the type $\times$ timing interaction fell short of significance [$F(1,11) = 4.1, p < .07$].

It may be asked whether the context effects obtained were due to differences in sensitivity or in simple response bias.[5] The detection paradigm used does not lend itself easily to a signal detection theory analysis, because it is unclear how false alarm percentages should be calculated. However, if a simple response bias had been operating, there should have been significant differences in the false alarm rates across conditions, paralleling the differences in hit rates. A two-way repeated measures ANOVA on the numbers of false alarms, with the variables of precursor type and precursor timing, revealed no significant effects, which suggests that the differences among precursor conditions were not due to simple response bias. False alarm percentages were calculated on the simple assumption that each trial contained 36 observation periods (i.e., positions in the music), 4 of which contained hesitations and 32 of which did not. Thus, the average number of false alarms per trial was divided by 32 and multiplied by 100. The resulting percentages, shown in Table 1, are perhaps unrealistically low because task constraints and response contingencies have not been taken into account in their calculation, but they do convey that false alarms were rather insensitive to the contextual manipulation.

As was expected, there was a significant main effect of position in the ANOVA on hits [$F(35,385) = 14.3, p < .0001$], indicating that the detectability of hesitations varied across positions in the music.[6] However, none of the interactions of position with precursor type and precursor timing approached significance, despite the large number of degrees of freedom. This suggests that neither of the two precursor main effects (of type and of timing) declined across positions in the test excerpt.

EXPERIMENT 2
Synchronization

Experiment 2 employed a subset of the materials of Experiment 1, but instead of reporting perceived hesitations, the participants tapped in synchrony with the music, and the variability of the taps was the primary measure of interest. The test excerpts still contained small timing perturbations, so that the effectiveness of phase correction could be assessed as well. Only two of the four precursors employed in Experiment 1 were used—namely, the musical ones, which had produced a larger context effect than the click precursors.

The participants were required to tap in synchrony with the precursors, as well as with the following test excerpts. It is known that attempts to synchronize with a sequence exhibiting variable and unpredictable timing result in tracking behavior; that is, the timing of the taps echoes the timing of the sequence at a lag of 1 (Michon, 1967; Repp,

2002). If a fixed timing pattern is presented repeatedly, it may gradually be remembered, and tracking may change to prediction (anticipation). For the present purposes, it was sufficient that the taps accompanying the expressively timed precursor be highly variable in their timing, which would be true regardless of whether tracking or prediction occurred. By contrast, taps synchronized with the isochronous precursor were expected to exhibit much lower variability. The question of interest was whether temporal variability in an immediately preceding motor activity (as well as in the sequence that elicited that variability) would increase the variability of taps synchronized with the test excerpts and whether it would also retard phase correction in response to small perturbations in the timing of the test sequences.

Method

Materials. The two musical precursors, one being isochronous and the other expressively timed, and the test excerpts were taken from Experiment 1. The latter was the set with $\Delta t = 20$ msec; that is, each excerpt contained four small hesitations that were quite difficult to hear.

Participants. There were 12 participants, none of whom had been in Experiment 1. It was considered necessary that the participants be able to synchronize their taps reasonably well with the musical test excerpts. Initially, the participants were recruited through an advertisement, without requiring or assessing musical training. However, 5 of the first 11 participants (perhaps those with little or no musical training) provided poor data, owing to missing taps, phase drift, and excessive variability. To avoid further rejections, 5 highly trained undergraduate musicians (present or former members of the Yale Symphony Orchestra) were recruited; surprisingly, 1 of them also provided poor data. In addition to the 10 successful paid participants, a graduate research assistant (with little musical training) and a postdoctoral researcher (with moderate musical training) participated.

Procedure. The participants tapped with the index fingers of their preferred hands on a white key of a Fatar Studio 37 MIDI controller (a silent three-octave piano keyboard), which they held on their laps. The key depressions were recorded via a MIDI interface by the MAX program that also controlled presentation of the musical excerpts.

The participants were given a few practice trials, followed by three blocks of test trials without precursors (27 trials in all), to provide further practice. Tapping started with the first downbeat (the second tone) in each excerpt (cf. Figure 1, top) and continued in synchrony with the tones corresponding to sixteenth notes, for a total of 37 taps. The participants were not informed that there were small hesitations in the music, and none of them spontaneously mentioned hearing irregularities of timing. The main part of the experiment consisted of six blocks of test trials, with each trial being preceded by a precursor. The precursors were constant within each block but alternated between blocks, with some participants starting with the isochronous precursor and others with the expressively timed one. The participants were requested to tap in synchrony with each precursor, which was easy when the precursor was isochronous, but quite difficult when it was expressively timed. The silent intervals between precursors and test excerpts were the same as those in Experiment 1.

Results and Discussion

Timing variability. The within-trial variability of the taps was assessed in terms of the standard deviations of the tap-tone asynchronies and of the intertap intervals (ITIs). The initial three taps (during which participants "tuned in"

to the sequence tempo) and the final tap in each trial were disregarded. The standard deviations were averaged across trials and blocks. They included not only quasi-random variability owing to the participant's timekeeping and motor control processes, but also systematic variability caused by the musical structure (discussed below) and by the small hesitations in the test trials. Analogous standard deviations were also calculated for the asynchronies and ITIs produced in synchronizing with the precursors.[7]

Table 2 shows the results. Variability in tapping to the isochronous precursors (top) was comparable to that in tapping to the test excerpts (bottom), despite differences in the construction of the materials. By contrast, the variability of tapping to the expressively timed precursors was very high and similar to the variability of IOIs in the precursor itself ($SD = 80$ msec). Not shown in Table 2 is the finding of a clear decrease in this high variability across the three blocks, both for asynchronies [$F(2,22) = 44.3$, $p < .0001$] and for ITIs [$F(2,22) = 44.7$, $p < .0001$], which probably reflects temporal pattern learning. Despite this improvement, however, synchronization with expressively timed precursors was still quite poor in Block 3.

The bottom part of Table 2 shows the variability of the taps that were synchronized with the test excerpts following precursors. The average standard deviation of the asynchronies was slightly higher following expressively timed precursors than following isochronous precursors, but that difference was not significant. The average standard deviation of the ITIs, too, was increased slightly after exposure to expressively timed precursors, and that difference did reach significance in a repeated measures ANOVA with the variables of precursor (2) and block (3) [$F(1,11) = 6.5$, $p < .03$].

The interpretation of this effect depends on whether it reflects an increase in quasi-random variability, in systematic variability, or in both. Systematic deviations from regularity in the timing of taps synchronized with isochronous music have been observed in several previous studies, some of which used the same musical excerpt (Repp, 1999a, 1999b, 1999c, 2002). To determine this variation

in the present data, the time series of asynchronies and ITIs were averaged across all trials and then across participants. To avoid possible contamination by phase correction, the taps coinciding with hesitations in the test excerpts and the two following taps were removed from the data. The averages across the nine trials in a block thus were based on only six datapoints per position. Figure 2 shows the average asynchrony profiles and ITI profiles for test excerpts in the two precursor conditions.

As is commonly observed in synchronization tasks, the taps generally occurred before the tones they were intended to coincide with, hence the negative values of the asynchronies (Figure 2A).[8] However, the asynchronies were somewhat less negative following the expressively timed precursors than following the isochronous precursors. This difference was significant in a repeated measures ANOVA with the variables of precursor condition (2) and position [34 or 33; $F(1,11) = 9.1$, $p < .02$]. The initial three datapoints were omitted in the ANOVA, since they represent the "tuning in" to the sequence tempo.[9] There was a significant main effect of position for both asynchronies [Figure 2A; $F(33,363) = 8.3$, $p < .0001$] and ITIs [Figure 2B; $F(32,352) = 13.2$, $p < .0001$]. Thus, both profiles exhibited the expected systematic deviations from regularity. Asynchronies tended to be most negative in positions 4 and 8 of each bar, and ITIs consequently were longest following the tones in these positions, which preceded the metrically strong tones in positions 5 and 1, respectively. The patterns of these profiles thus may be seen as a reflection of the metrical structure (the level of quarter-note beats) of the music. What is important here is that there was a significant condition × position interaction for both asynchronies [$F(33,363) = 1.5$, $p < .04$] and ITIs [$F(32,352) = 1.8$, $p < .006$]. It can be seen in Figure 2 that both asynchronies and ITIs were more strongly modulated initially (bars 1–2) after expressively timed precursors than after isochronous precursors. The visual impression that these interactions derived from the initial portions of the profiles was confirmed by omitting the first 10 datapoints following the initial tuning-in and re-

Table 2
Average Within-Trial Standard Deviations (in Milliseconds)
of Asynchronies (Asynch) and Intertap Intervals (ITIs)
For Tapping in Synchrony With (Top) Precursors
and (Bottom) Test Excerpts

	Synchronization With Precursors			
	Isochronous	Expressive		
Asynch	23.3	74.3		
ITIs	22.2	91.7		
	Synchronization With Test Excerpts			
	Precursor Condition		Difference	
	No Precursor	Isochronous	Expressive	(Exp − Iso)	SE
Asynch	23.9	22.5	23.1	0.6	0.4
ITIs	21.7	22.6	23.7	1.1	0.5

Note—The last column shows the difference between the two precursor conditions, with standard errors.
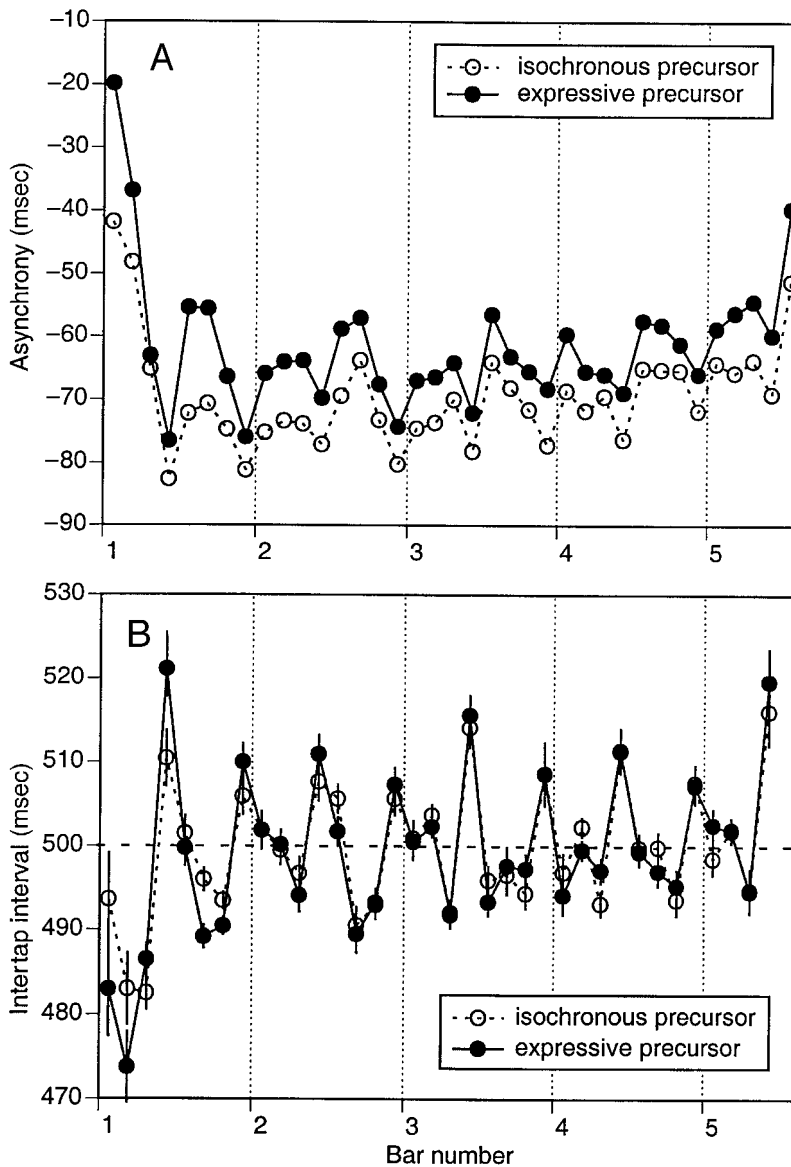
Figure 2. (A) Average asynchrony profiles for test excerpts in the two precursor conditions of Experiment 2. (B) Average intertap interval profiles, with between-subjects standard error bars.

peating the ANOVAs on the remaining data; these analyses yielded condition × position interactions that were clearly nonsignificant ($p > .5$).

Thus, it seems that the expressively timed precursor caused a transient increase in the systematic variability of the taps to test excerpts. To examine whether this effect accounted for the overall differences in ITI variability (Table 2), the standard deviations of the ITIs were recalculated with the same 10 datapoints omitted. The resulting values were 22.7 and 22.8 msec following isochronous and expressively timed precursors, respectively—obviously not a significant difference. Thus, the significant difference reported in Table 2 was indeed due to the initial portion of the ITI profile only, and moreover, it was largely or entirely

due to a difference in systematic variability, not in quasirandom variability (which tends to vanish when data are averaged across trials, as in Figure 2). The difference in systematic variability may have been caused by a heightened sensitivity to musical structure following exposure to expressively timed precursors.

Phase correction. To examine the speed of compensation for hesitations in the test excerpts, the datapoint triplets that had been extracted from the asynchrony time series before computing the average asynchrony profiles were analyzed. Each triplet represented three successive asynchronies, starting with the tap that coincided with a hesitation in the music. These asynchronies were expressed as deviations from each participant's average asynchrony
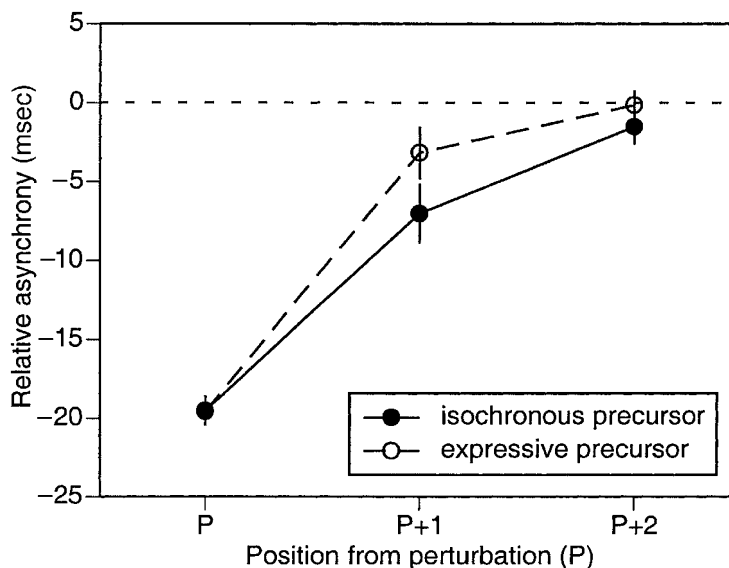
Figure 3. Phase correction functions in Experiment 2: Average relative asynchrony as a function of position for the two precursor conditions, with between-subjects standard error bars.

profile. The first of the resulting relative asynchronies had an expected value of -20 msec (i.e., $-\Delta t$). A rapid return of the subsequent relative asynchronies to zero would be indicative of phase correction. There were 36 datapoint triplets in each block of test trials, representing the 36 possible positions of the hesitation in the music. The relative asynchronies were averaged over 32 of these triplets (omitting the first 2 and the last 2) and then across participants.

The resulting phase correction functions for the two precursor conditions are shown in Figure 3. As was expected, the average relative asynchronies were close to -20 msec in position P (the perturbation point) and then returned toward the zero baseline. Unexpectedly, phase correction was more rapid following expressively timed precursors than following isochronous precursors. In a repeated measures ANOVA with the variables of condition (2) and relative position (3), the condition $\times$ position interaction was significant [$F(2,22) = 7.7, p < .003$]. Although this result is difficult to explain, it indicates that contextual timing variability did not slow down phase correction.

EXPERIMENT 3
Perception and Synchronization

Experiments 1 and 2 replicated earlier results that had suggested that time perception is sensitive to temporally varying context, whereas sensorimotor synchronization is not. However, there were differences between the two experiments that perhaps reduce the impact of the results. One difference was in the magnitude of the hesitations in the musical test excerpts: These perturbations were generally larger in Experiment 1, where they were adjusted to each participant's ability, than in Experiment 2, where

they were fixed at $\Delta t = 20$ msec. This difference could have interacted somehow with the context effect of interest. Another difference was in the musical training of the participants, which on average was greater in Experiment 2 than in Experiment 1. Furthermore, several general design features carried over from previous studies—the complex musical materials, the use of an expressive timing pattern to represent temporal variability, the repetition of the same timing pattern, and the presence of multiple hesitations in each test excerpt—may have introduced complications.

Experiment 3 addressed all these concerns. It combined the detection and synchronization tasks, which were carried out simultaneously on the same materials by the same participants. The materials were simple click sequences. The temporal variability in the precursors was random and changed from trial to trial. Each test sequence contained only a single hesitation, and a fixed range of hesitation magnitudes was used, intended to span the detection threshold. The only drawback of this design was that fewer detection responses were obtained. However, the context effect on perception was expected to be sufficiently robust to survive this reduction in the number of observations.

Method
Materials. All sequences were composed of high-pitched digital piano tones, as used in the click precursors of Experiment 1. Both precursor and test sequences consisted of 20 tones, with a baseline IOI of 500 msec. Precursor sequences were either isochronous or contained random deviations from isochrony. Twelve different random timing patterns were generated by adding to each 500-msec IOI a value drawn randomly from between -60 and $+60$ msec. The mean IOIs of these 12 patterns ranged from 482 to 513 msec (grand mean = 499 msec), and their standard deviations ranged from 47 to 69 msec (grand mean = 57 msec). Test sequences either were isochronous or contained a single lengthened IOI. The amount of

lengthening, Δt, was 10, 15, 20, 25, or 30 msec, and the hesitation occurred in one of 10 possible positions, ranging from the 6th to the 16th IOI, with the 11th IOI excluded. Thus, the hesitations could be divided roughly into early and late occurrences. The 50 test sequences (5 Δt values $\times$ 10 positions) were arranged into five blocks of 10, so that each Δt occurred twice in each block and each position occurred once. Two isochronous test sequences were included in each block. The order of sequences within each block was random.

Participants. The participants were 10 individuals who had participated in a number of earlier sensorimotor synchronization experiments in the author's laboratory, but not in Experiments 1 and 2. Their musical training varied widely, ranging from a few years to professional level, and their ages ranged from 19 to 56 years. Seven participants were paid and were unaware of the purpose of the study; the other 3 were a postdoctoral research associate, a research assistant, and the author.

Procedure. Each participant was first presented with one practice block without precursors. The task was to tap in synchrony with the tones and to report, after the end of each sequence, whether or not the sequence had contained a hesitation. The participants tapped on a quiet MIDI controller (as in Experiment 2) and made their perceptual judgments by pressing one of two keys on the computer keyboard. No feedback was provided. The registration of the computer keyboard response started the next trial after a delay of 3 sec. One participant, who had not heard any hesitations, was given a second practice block.

Subsequently, five test blocks were presented with each of the two types of precursors. Precursor type alternated from block to block. Some participants started with isochronous precursors (order of blocks: 1, 2, 3, 4, 5), others with randomly timed precursors (order of blocks: 4, 5, 1, 2, 3). The interval between precursors and test trials was 2 sec. In the randomly timed precursor condition, a different precursor preceded each test trial in a block, but the same 12 precursors occurred in different blocks. The participants were required to tap in synchrony with both precursor and test sequences and to report after the end of each test sequence whether or not they had heard a hesitation. Guessing was to be avoided.

Results and Discussion

Detection responses. Figure 4 shows the average percentages of positive detection responses. Those for an IOI increment size of 0 msec are false alarms; the others are hits. As was expected, hit percentages increased with the magnitude of the hesitation. The detection threshold (50% correct) in the isochronous precursor condition was about 18 msec (close to 4% of the IOI duration), which is consistent with previous studies (Friberg & Sundberg, 1995; Repp, 2000a). However, it was higher in the randomly timed precursor condition; that is, there were fewer correct detection responses. This was confirmed in a one-way repeated measures ANOVA[10] on the total number of correct responses [$F(1,9) = 13.9, p < .005$]. In theory, this difference could have been due to a response bias, especially since the purpose of the study was transparent to some participants. However, the false alarm rates were low and identical in the two precursor conditions, which suggests that no simple response bias was operating. To determine whether there was any decline of the precursor effect in the course of the test sequence, separate detection scores were calculated for hesitations in earlier and later positions, for all Δt values combined. A repeated measures ANOVA on these data, with the variables of precursor condition (2) and position (2), did not yield a significant interaction [$F(1,9) = 1.3, p > .25$]. Thus, there was no evidence that the precursor effect declined across positions.

Variability of finger taps. The average standard deviation of the asynchronies of the taps synchronized with isochronous precursors was 13.8 msec, and that of the corresponding ITIs was 14.7 msec. For the taps accompanying randomly timed precursors, only the ITIs were determined; their standard deviation was 57.1 msec. Thus, the
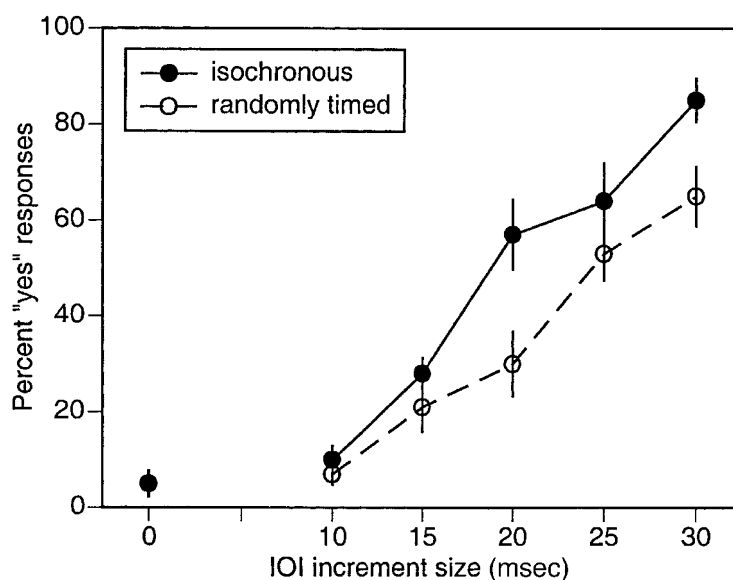

Figure 4. Average percentages of "yes" responses in the two precursor conditions of Experiment 3, with between-subjects standard error bars. IOI, interonset interval.
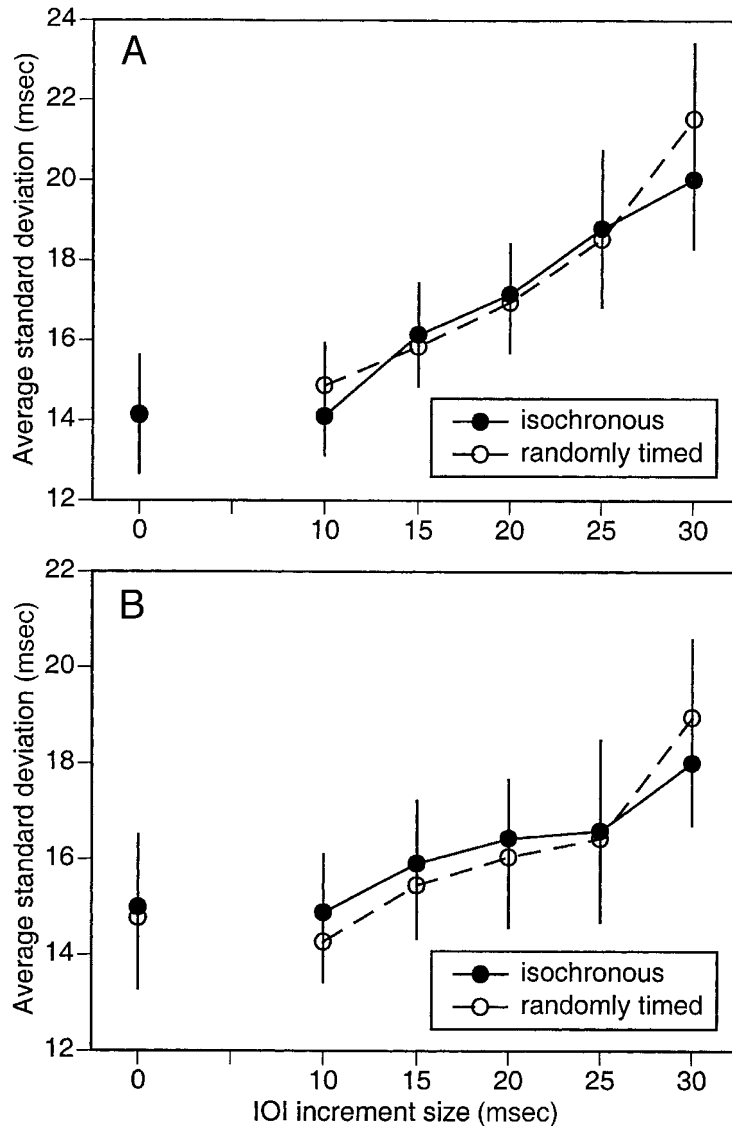
Figure 5. Average standard deviations of (A) tap shifts and (B) intertap intervals following isochronous and randomly timed precursors in Experiment 3, with between-subjects standard error bars. IOI, interonset interval.

randomly timed precursors clearly made the participants tap with variable timing.

The variability of the taps accompanying the test sequences were analyzed in terms of tap shifts, rather than tap–tone asynchronies.[11] The within-trial standard deviations of the shifts and ITIs were computed and averaged across all trials with the same Δt. The initial three taps and the final tap were disregarded. Since the phase correction was included in this computation, the standard deviations of the shifts and ITIs were expected to increase with Δt. Figure 5 shows the results in a format comparable to that of the detection scores in Figure 4.[12] The standard deviations increased with Δt, as was expected, but they were very similar in the two precursor conditions. Repeated measures ANOVAs on each set of standard deviations,

with the variables of precursor condition (2) and hesitation magnitude (5), yielded only significant main effects of hesitation magnitude for both shifts and ITIs [$F(1,9) > 9$, $p < .0001$], which is a trivial finding. The condition main effects [$F(1,9) < 1$] and interactions [$F(4,26) < 1.8, p > .14$] did not approach significance.

Phase correction. For each of the four taps following a hesitation, a measure of relative shift was calculated by subtracting the shift of the tap coinciding with the hesitation from that of the following taps. Then a linear regression of relative shift on Δt was performed. (Trials with $\Delta t = 0$ were not included.) By definition, the slope of the regression line was zero at the point of the hesitation. The slope of the regression line for each subsequent tap reflects the extent of phase correction and reaches 1 when

phase correction is complete, showing that the tap has shifted by Δt. These slopes were determined for each participant in each precursor condition. The average slopes for the two precursor conditions are shown in Figure 6.[13] It is clear that the randomly timed precursor did not impair phase correction. In fact, there was a nonsignificant tendency to overcorrect (slope > 1) in that condition. A repeated measures ANOVA on the slopes, with the variables of precursor condition (2) and position (4), yielded no significant effects. Thus, the finding of Experiment 2 that phase correction was more effective following a variably timed precursor was not replicated, although some individual participants seemed to show an effect of that kind.

Relation between detection scores and relative shift of taps. If a single timekeeper underlies perception and synchronized action, there should be correlated noise (i.e., timekeeper variability) in the two activities. In particular, there should be a negative correlation between the probability of detecting a hesitation of a given magnitude and the relative shift of the tap coinciding with it. The correlation should be negative, because a hesitation, by definition, creates a negative perceptual asynchrony (i.e., the temporal expectation for a tone precedes the actual time of occurrence of a tone); therefore, the more negative this asynchrony is, the more likely it is that the hesitation will be detected. This is obviously true across different values of Δt. However, it should hold even for a fixed value of Δt near the detection threshold. If the same timekeeper controls perception and action, the timing of the taps should be positively correlated with temporal expectations, and this implies that the tap coinciding with a hesitation should tend to occur relatively earlier on trials in which the hesitation is detected than on trials in which it is not detected (cf. Repp, 1999b).

The relevant measure was taken to be the relative shift between the tap immediately preceding a hesitation and the tap coinciding with the hesitation. (Similarly, it is presumably the local change in the unobservable perceptual asynchrony, rather than its absolute magnitude, that is most relevant to detection of a hesitation.) By considering relative shift rather than relative asynchrony (relative shift = relative asynchrony $+ \Delta t$), the effect of Δt was factored out, so that data for different Δt values could be combined. Thus, average relative shifts were computed separately for trials with positive and negative detection responses across Δt values from 15 to 30 msec, separately for each precursor condition. The resulting values were subjected to a repeated measures ANOVA with the variables of precursor condition (2) and detection response (2). The main effect of condition was significant [$F(1,9) = 8.4, p < .02$]: Taps coinciding with a hesitation occurred, on average, 3.3 msec earlier on trials in which the hesitation was detected than on trials in which the hesitation was not detected, which is the predicted negative correlation. The difference seemed to derive mainly from the randomly timed precursor condition, but the interaction did not reach significance [$F(1,9) = 3.7, p < .09$]. These results, although merely suggestive, are consistent with the hypothesis that the perceptual task and the motor task engaged the same internal timekeeper or oscillator.

GENERAL DISCUSSION

The present research examined the sensitivity of two ostensibly similar timing tasks—perceptual detection of hesitations and synchronized finger tapping—to contextual temporal variability. Experiments 1 and 3 both demonstrated that exposure to a variably timed auditory precur-
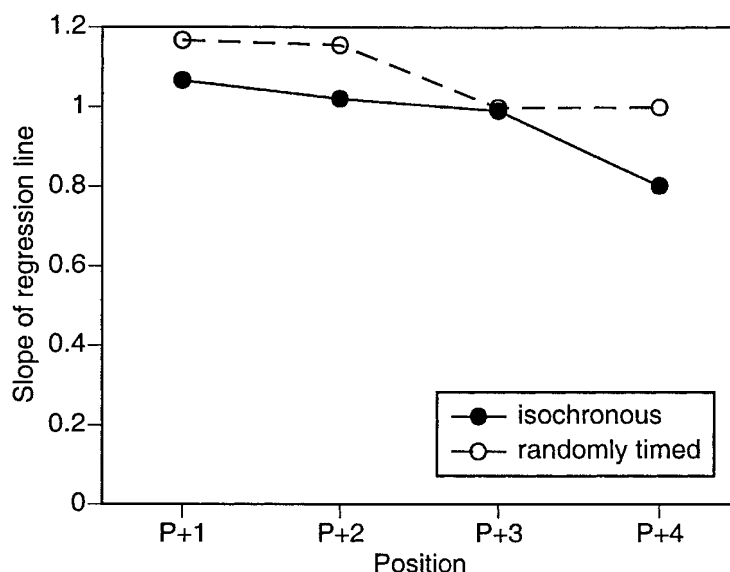

Figure 6. Average extent of phase correction following isochronous and randomly timed precursors in Experiment 3. The hesitation occurred in position P. A slope value of 1 indicates complete phase correction.

sor sequence reduces listeners' perceptual sensitivity to deviations from temporal regularity in a subsequent test sequence. This precursor effect occurred regardless of the type of precursor or test sequence used (clicks or music). Quite different results were obtained in Experiments 2 and 3 with regard to the variability of finger taps synchronized with a test sequence. Even though the participants were required to tap in synchrony with the precursors, so that precursor variability entailed concurrent perceptual *and* motor variability, this contextual variability had little or no effect on the variability of taps synchronized with a subsequent test sequence. Likewise, the speed of phase correction was not reduced by precursor variability.

In Experiment 2, a small precursor effect on tapping variability was found, but it was due to a brief increase in *systematic* deviations from regularity, not to an increase in the quasi-random variability attributed (in large part) to basic timekeeping processes. This suggests that the expressively timed precursor increased the participants' sensitivity to structural musical factors, such as meter, which in turn led to increased systematic modulations of the period of the underlying timekeeper (see also Repp, 1999b, 2002). This interpretation is consistent with the fact that the precursor effect wore off rapidly during exposure to the deadpan musical test sequence and with the absence of any precursor effect in Experiment 3, in which nonmusical materials were used. In other words, the small precursor effect observed in Experiment 2 was specific to music.

Taken together, the results show that the perceptual task of detecting deviations from temporal regularity in a sequence is more context sensitive than the sensorimotor task of synchronizing finger taps with a sequence. As was pointed out in the introduction, this finding can be interpreted in two ways. One is that different basic timekeepers underlie the two activities and that the perceptual timekeeper increases its variability after exposure to contextual timing variability, whereas the motor timekeeper does not. It would then remain mysterious why one timekeeper should be more context sensitive than the other. Conceptually and functionally, all timekeepers or oscillators seem similar, regardless of the task (e.g., Large & Jones, 1999; Pressing, 1999). Therefore, if a perceptual timekeeper is affected by perceptual variability, a motor timekeeper should be affected in a similar way by motor variability. This was apparently not the case. Moreover, there was some evidence in Experiment 3 for the presence of correlated timekeeper noise in concurrent perception and motor control.[14] Therefore, the hypothesis that similar perceptual and motor tasks rely on the same (or overlapping) basic timekeeping processes need not be rejected and, in fact, remains plausible.

The other possibility, which appears to be supported by the present findings, is that the contextual variability affected processes that were specific to the perceptual task. The most obvious difference between a detection task and a synchronization task is that detection requires conscious awareness and judgment of temporal differences, whereas synchronization relies largely on automatic and subconscious processes of temporal prediction and error correc-

tion, which may also be involved in perception (Large & Jones, 1999) but are not context sensitive. The context-sensitive processes seem to be located at a different level that governs the access of temporal differences to consciousness. Conscious awareness of temporal differences is limited by a variable detection threshold, and it is this threshold that seems to be raised following exposure to temporal variability. No such threshold seems to limit the use of temporal information for the control of action (Repp, 2000a).

Large and Jones (1999) attributed perceptual effects of contextual timing variability to a widening of the expectancy window of a slowly adapting attentional oscillator. Although similar oscillatory processes may underlie perception of timing and synchronized action, the expectancy window may be specific to perception. Indeed, its function is to set a detection threshold for deviations from expectancies. An expectancy window is formally equivalent to a probability distribution of temporal expectancies. Thus, it may be considered both as a memory representation of past variability (including contextual variability) and as a process that generates specific expectations. The present results thus are quite consistent with the theory of Large and Jones. What they contribute is evidence suggesting that a memory for past variability plays no role in sensorimotor synchronization. (However, *systematic* timing patterns can be learned and predicted in synchronization; see Repp, 1999a, 2002.) Experiments 2 and 3 also showed that contextual timing variability does not interfere with phase correction following a hesitation in a sequence, and this is consistent with the hypothesis that phase correction is a low-level automatic process (Repp, 2000a).

The results of Experiments 1 and 3 suggest that the precursor effect on perception does not abate in the course of an isochronous (or nearly isochronous) test sequence lasting up to 20 sec. Such a long-lasting effect is consistent with effects of session context reported by Large and Jones (1999). However, in the framework of their adaptive oscillator model, it is surprising that adaptation to the absence of external variability is so slow.

One unexpected finding in Experiment 1 was that the music precursors interfered more with detection performance than did the click precursors. This may have occurred because the music precursors attracted more attention, because they were more similar to the test sequences, or because variation in other parameters (pitch, intensity, asynchronies) somehow augmented the effect of temporal variability.

It should be mentioned here that the present study had an additional purpose, discussion of which has been relegated to an electronic appendix (see note 2). Briefly, it concerned the effect of expressive *intensity* variation in musical test excerpts on perception of deviations from regularity and on variability in sensorimotor synchronization. Intensity variations were absent in the materials of Experiments 1 and 2, but they had been present in earlier experiments (Repp, 1999b, 2000a), which were sufficiently similar otherwise to make comparisons across studies

meaningful. These comparisons revealed that intensity variations had a significant effect on perception, but not on synchronization. Thus, this local nontemporal variation, like the more remote temporal variation in the precursors, seemed to affect temporal processes specific to perception. Some other dissociations between perception of timing and synchronized action have been reported by Repp (2000a).

In conclusion, the perception–action dissociations reported here are likely to be the consequence of the differential involvement of conscious awareness and judgment in perceptual and sensorimotor tasks. There are many recent demonstrations of dissociations between (conscious) perception and action, especially in tasks based on visual and spatial information (e.g., Creem & Proffitt, 1998; Gentilucci, Chieffi, Daprati, Saetti, & Toni, 1996; Haffenden & Goodale, 1998; Klotz & Neumann, 1999; Rumiati & Humphreys, 1998). The present research contributes to the rapidly mounting evidence that action is often based on sensory information that does not reach awareness, or before it reaches awareness (see Neumann, 1990; Repp, 2000a). This seems to hold for auditory and temporal information, just as it does for visual and spatial information. Moreover, the present results suggest that temporally varying context has its effect not on a basic timekeeper, but on processes that bring temporal differences to awareness.

REFERENCES

Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology, 41*, 254-311.

Beek, P. J., Peper, C. E., & Daffertshofer, A. (2000). Timekeepers versus nonlinear oscillators: How the approaches differ. In P. Desain & L. Windsor (Eds.), *Rhythm perception and production* (pp. 9-33). Lisse, The Netherlands: Swets & Zeitlinger.

Creem, S. H., & Proffitt, D. R. (1998). Two memories for geographical slant: Separation and interdependence of action and awareness. *Psychonomic Bulletin & Review, 5*, 22-36.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America, 98*, 2524-2531.

Gentilucci, M., Chieffi, S., Daprati, E., Saetti, M. C., & Toni, I. (1996). Visual illusion and action. *Neuropsychologia, 34*, 369-376.

Haffenden, A., & Goodale, M. A. (1998). The effect of pictorial illusion on prehension and perception. *Journal of Cognitive Neuroscience, 10*, 122-136.

Ivry, R. B. (1997). Cerebellar timing systems. *International Review of Neurobiology, 41*, 555-573.

Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception & Performance, 21*, 3-18.

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience, 1*, 136-152.

Ivry, R. B., & Richardson, T. C. (2002). Temporal control and coordination: The multiple timer model. *Brain & Cognition, 48*, 117-132.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*, 323-355.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responding to time. *Psychological Review, 96*, 459-491.

Jones, M. R., & Yee, W. (1997). Sensitivity to time change: The role of context and skill. Journal of Experimental Psychology: Human Perception & Performance, 23, 693-709.

Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. B. (1985).

Do perception and motor production share common timing mechanisms: A correlational analysis. *Acta Psychologica, 60*, 173-191.

Klotz, W., & Neumann, O. (1999). Motor activation without conscious discrimination in metacontrast masking. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 976-992.

Kolers, P. A., & Brewster, J. M. (1985). Rhythms and responses. *Journal of Experimental Psychology: Human Perception & Performance, 11*, 150-167.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How we track time-varying events. *Psychological Review, 106*, 119-159.

Mates, J. (1994). A model of synchronization of motor acts to a stimulus sequence: I. Timing and error corrections. *Biological Cybernetics, 70*, 463-473.

Meegan, D. V., Aslin, R. N., & Jacobs, R. A. (2000). Motor timing learned without motor training. *Nature Neuroscience, 3*, 860-862.

Michon, J. A. (1967). *Timing in temporal tracking.* Assen, The Netherlands: van Gorcum.

Neumann, O. (1990). Direct parameter specification and the concept of perception. *Psychological Research, 52*, 207-215.

Pressing, J. (1999). The referential dynamics of cognition and action. *Psychological Review, 106*, 714-747.

Repp, B. H. (1994). On determining the basic tempo of an expressive music performance. *Psychology of Music, 22*, 157-167.

Repp, B. H. (1998a). The detectability of local deviations from a typical expressive timing pattern. *Music Perception, 15*, 265-290.

Repp, B. H. (1998b). A microcosm of musical expression: I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *Journal of the Acoustical Society of America, 104*, 1085-1100.

Repp, B. H. (1998c). Obligatory "expectations" of expressive timing induced by perception of musical structure. *Psychological Research, 61*, 33-43.

Repp, B. H. (1998d). Variations on a theme by Chopin: Relations between perception and production of deviations from isochrony in music. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 791-811.

Repp, B. H. (1999a). Control of expressive and metronomic timing in pianists. *Journal of Motor Behavior, 31*, 145-164.

Repp, B. H. (1999b). Detecting deviations from metronomic timing in music: Effects of perceptual structure on the mental timekeeper. *Perception & Psychophysics, 61*, 529-548.

Repp, B. H. (1999c). Relationships between performance timing, perception of timing perturbations, and perceptual–motor synchronization in two Chopin preludes. *Australian Journal of Psychology, 51*, 188-203.

Repp, B. H. (2000a). Compensation for subliminal timing perturbations in perceptual–motor synchronization. *Psychological Research, 63*, 106-128.

Repp, B. H. (2000b). Pattern typicality and dimensional interactions in pianists' imitation of expressive timing and dynamics. *Music Perception, 18*, 173-211.

Repp, B. H. (2001). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception & Performance, 27*, 600-621.

Repp, B. H. (2002). The embodiment of musical structure: Effects of musical context on sensorimotor synchronization with complex timing patterns. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance XIX* (pp. 245-265). Oxford: Oxford University Press.

Robertson, S. D., Zelaznik, H. N., Lantero, D. A., Bojczyk, K. G., Spencer, R. M., Doffin, J. G., & Schneidt, T. (1999). Correlations for timing consistency among tapping and drawing tasks: Evidence against a single timing process for motor control. *Journal of Experimental Psychology: Human Perception & Performance, 25*, 1316-1330.

Rumiati, R. I., & Humphreys, G. W. (1998). Recognition by action: Dissociating visual and semantic routes to action in normal observers. *Journal of Experimental Psychology: Human Perception & Performance, 24*, 631-647.

Schubotz, R. I., Friederici, A. D., & von Cramon, D. Y. (2000). Time

perception and motor timing: A common cortical and subcortical basis revealed by fMRI. *NeuroImage, 11,* 1-12.

SMITH, O. W. (1957). Relationship of rhythm discrimination to motor rhythm performance. *Journal of Applied Psychology, 41,* 365-369.

THAUT, M. H., TIAN, B., & AZIMI-SADJADI, M. R. (1998). Rhythmic finger tapping to cosine-wave modulated metronome sequences: Evidence of subliminal entrainment. *Human Movement Science, 17,* 839-863.

TREISMAN, M., FAULKNER, A., & NAISH, P. L. N. (1992). On the relation between time perception and the timing of motor action: Evidence for a temporal oscillator controlling the timing of movement. *Quarterly Journal of Experimental Psychology, 45A,* 235-263.

TURVEY, M. T. (1977). Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing* (pp. 211-265). Hillsdale, NJ: Erlbaum.

VIVIANI, P., & STUCCHI, N. (1989). The effect of movement velocity on form perception: Geometric illusions in dynamic displays. *Perception & Psychophysics, 46,* 266-274.

VIVIANI, P., & STUCCHI, N. (1992a). Biological movements look uniform: Evidence of motor–perceptual interactions. *Journal of Experimental Psychology: Human Perception & Performance, 18,* 603-623.

VIVIANI, P., & STUCCHI, N. (1992b). Motor–perceptual interactions. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior II* (pp. 229-248). Amsterdam: Elsevier.

VORBERG, D., & WING, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action* (Vol. 2, pp. 181-262). London: Academic Press.

WALLACE, S. A. (1996). Dynamic pattern perspective of rhythmic movement: An introduction. In H. N. Zelaznik (Ed.), *Advances in motor learning and control* (pp. 155-194). Champaign, IL: Human Kinetics.

WING, A. M., & KRISTOFFERSON, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception & Psychophysics, 14,* 5-12.

YEE, W., HOLLERAN, S., & JONES, M. R. (1994). Sensitivity to event timing in regular and irregular sequences: Influences of musical skill. *Perception & Psychophysics, 56,* 461-471.

ZELAZNIK, H. N., SPENCER, R. M., & DOFFIN, J. G. (2000). Temporal precision in tapping and circle drawing movements at preferred rates is not correlated: Further evidence against timing as a general-purpose ability. *Journal of Motor Behavior, 32,* 193-199.

NOTES

1. For a discussion of the differences between timekeepers and oscillators, see Beek, Peper, and Daffertshofer (2000). The author is not committed to a particular theoretical approach and will use both metaphors interchangeably.

2. Owing to a peculiarity of the MIDI software used (MAX), all real-time intervals were 2.4% shorter than specified, recorded, and reported in this paper. The actual baseline IOIs thus were 488 msec long. Nominally simultaneous MIDI note onsets were ordered in the MIDI instructions according to pitch from high to low, so that the most important tones (those constituting the top line) were not affected by MIDI transmission delays (cf. Repp, 1999b, Appendix). The motivation for making the key depression velocities uniform (which was detrimental to the aesthetic quality of the test excerpts but is uncontroversial from a scientific viewpoint) is spelled out in an electronic appendix to this paper, which can be found at www.haskins.yale.edu/haskins/STAFF/repp.html.

3. Although the difference in mean IOI durations suggests that the tempo of the precursor was slower than that of the test excerpts, this conclusion may not be justified. Expressive timing implies lengthening of IOIs without compensatory shortening, and this results in an asymmetric distribution of IOI durations. The perceived tempo may be closer to the lower edge of the IOI distribution than to the mean IOI (see Repp, 1998a; but also Repp, 1994). Thus, it is believed that precursors and test

excerpts were reasonably well matched in terms of basic tempo. Note also that tempo matching was not important, because exactly the same precursors were used in Experiments 1 and 2.

4. The final chord in musical precursors had a nominal duration of about 2.5 sec; in test excerpts, it lasted 2 sec. The silent interval was measured from the nominal end of the chord (the last MIDI instruction). In click precursors, however, the last MIDI instruction occurred only 100 msec after the last event onset. Thus, the interval from the onset of the last event in a precursor to the beginning of a test excerpt was longer for music precursors (about 4.5 sec) than for click precursors (2.1 sec). This may have introduced a bias toward obtaining stronger context effects with click precursors, but in fact the opposite result was obtained.

5. What is called *sensitivity* here is actually a form of perceptual bias or direction-specific sensitivity, as was discussed at length in Repp (1998d). However, it cannot be distinguished here from sensitivity that is not direction specific, because only lengthened IOIs were presented. In contrast to a perceptual bias, a simple response bias is the tendency to press the response key more or less often.

6. For a discussion of these variations, see Repp (1998d, 1999b) and the electronic appendix (note 2).

7. Two clear outliers (individual data for one block) were replaced with the mean value of the other two blocks in the same condition for that participant. One standard deviation, almost twice as large as the others, owing to phase drift, occurred in one block having expressively timed precursors. The other exceptionally large standard deviation occurred in tapping to the expressively timed precursor itself.

8. A small part of that difference, amounting to less than 20 msec, was due to the fact that the electronic contact of the response key occurred during its downward trajectory, before its bottom contact.

9. It is noteworthy that this tuning-in was observed even when the tempo was constant from one sequence to the next, as in the isochronous precursor condition. The first tap always tended to occur relatively late, so that the subsequent ITIs had to be shortened to achieve subjective synchrony.

10. Hesitation magnitude was not included as a variable, because its main effect is trivial and because a significant interaction with condition would reflect a floor effect at small Δt values.

11. This was done for reasons of convenience; there was no deeper theoretical reason. Shifts are asynchronies computed with reference to a hypothetical isochronous sequence. Prior to a perturbation, shifts are equal to asynchronies. At the perturbation point, the asynchrony is expected to change by $-\Delta t$, but the shift is expected to remain the same. Subsequently, asynchronies are expected to return to the average preperturbation asynchrony, whereas shifts are expected to approach Δt. An analysis in terms of shifts is equivalent to one in terms of asynchronies, although the standard deviations will be larger and depend more on Δt.

12. Note that the between-subjects standard errors in both figures are not directly relevant to the significance of differences in a repeated measures design.

13. Standard error bars have been omitted because there were unexpectedly large individual differences in phase correction strategies. Figure 6 suggests complete phase correction on the tap following a hesitation, but some individual participants showed much more gradual correction, whereas others showed large overcorrection. It would lead too far to discuss these differences here.

14. There is a possible alternative interpretation of this finding: A larger negative sensorimotor asynchrony may have increased the probability of a positive detection judgment. Even though people are not good at judging cross-modal asynchronies (tactile/kinesthetic taps vs. auditory tones) in sensorimotor synchronization (Repp, 2000a), an effect of subliminal sensorimotor asynchronies on perceptual judgments cannot be ruled out.